Earthquake Supercycles and Long-Term Fault Memory

Leah Salditch[a,b,*], Seth Stein[a,b], James Neely[a], Bruce D. Spencer[b,c], Edward M. Brooks[a,b], Amotz Agnon[d], Mian Liu[e]

[a] Department of Earth and Planetary Sciences, Northwestern University, 2145 Sheridan Rd., Evanston, IL 60208 U.S.A.
[b] Institute for Policy Research, Northwestern University, 2040 Sheridan Rd., Evanston, IL 60208 U.S.A.
[c] Department of Statistics, Northwestern University, 2006 Sheridan Rd., Evanston, IL 60208 U.S.A.
[d] Institute of Earth Sciences, Hebrew University, Sderot Magnes, Jerusalem 9190401, Israel
[e] Department of Geological Sciences, University of Missouri, 101 Geological Sciences Bldg., Columbia, MO 65211, U.S.A
[*] Corresponding Author, Leah@earth.northwestern.edu

Abstract

Long records often show large earthquakes occurring in supercycles, sequences of temporal clusters of seismicity, cumulative displacement, and cumulative strain release separated by intervals of lower levels of these measures. Supercycles and associated earthquake clusters are challenging to characterize via the traditionally used aperiodicity, which measures the extent that a sequence differs from perfectly periodic. Supercycles are not well described by commonly used models of earthquake recurrence. In the Poisson model, the probability of a large earthquake is constant with time, so the fault has no memory. In a seismic cycle/renewal model, the probability is quasi-periodic, dropping to zero after a large earthquake, then increasing with time, so the probability of a large earthquake depends only on the time since the past one, and the fault has only "short-term memory." We describe supercycles with a Long-Term Fault Memory (LTFM) model, where the probability of a large earthquake reflects the accumulated strain rather than elapsed time. The probability increases with accumulated strain (and time) until an earthquake happens, after which it decreases, but not necessarily to zero. Hence, the probability of an earthquake can depend on the earthquake history over multiple prior cycles. We use LTFM to simulate paleoseismic records from plate boundaries and intraplate areas. Simulations suggest that over timescales corresponding to the duration of paleoseismic records, the distribution of

earthquake recurrence times can appear strongly periodic, weakly periodic, Poissonian, or bursty. Thus, a given paleoseismic window may not capture long-term trends in seismicity. This effect is significant for earthquake hazard assessment because whether an earthquake history is assumed to contain clusters can be more important than the probability density function chosen to describe the recurrence times. In such cases, probability estimates of the next earthquake will depend crucially on whether the cluster is treated as ongoing or over.

Keywords

Earthquake; Supercycle; Cluster; Aperiodicity.

1. Introduction

Since the 1906 San Francisco earthquake, the dominant paradigm in earthquake seismology has been the earthquake cycle, in which strain accumulates between large earthquakes due to interseismic motion between the two sides of a locked fault and is released by coseismic slip on the fault when an earthquake occurs (Reid, 1910). Over time, this process should give rise to approximately periodic earthquakes and a steady accumulation of cumulative displacement (Figure 1a). The fact that earthquake sequences are only approximately periodic prompted a refinement of the model with "time-predictable" recurrence in which a specific strain level must accumulate for an earthquake, but the strain release in the earthquake is variable (Shimazaki and Nakata, 1980).

However, long earthquake records often show more complex behavior (Figure 1b). Wallace (1987) found that faults and groups of faults in the Western U.S.'s Great Basin often showed "*grouping, …a series of displacement events, each being followed by a period of quiescence. Slip rates during a group of events along a segment of fault, thus, could be considerably greater than the long-term average slip rate. During quiescent periods, the slip rate would be lower than the average rate and might even be zero... Additionally, if grouping is real, the concept that accumulated elastic strain is released at some regular interval by a single displacement event in a seismic cycle should be reexamined. Perhaps strain that has accumulated at a more or less constant rate is released in a stuttering, spasmodic manner in a group of displacement events."* Subsequent investigation (Friedrich et al., 2003) supported this analysis, finding that "*seismic strain release may be clustered on the 10-kyr timescale... with comparatively low, uniform strain accumulation rates on the 100-kyr timescale.*" They suggested calling the conventional earthquake cycles "Reid-type" behavior and the longer period variations "Wallace-type" behavior.

Wallace (1987) further noted that if this behavior is *"common, as these preliminary analyses suggest, care must be exercised in evaluating seismic hazard potentials. It is crucial to determine the timing and distribution of individual faulting events because long-term average slip rates may give grossly incorrect assessments of the hazard potential."* For example, the estimated probability of a future large earthquake can depend crucially on whether a cluster is treated as ongoing or over.

Such variations in earthquake behavior on timescales longer than individual cycles are often termed "supercycles," following Sieh et al.'s (2008) observation from corals near the Sumatra trench that "*because each of the three past episodes of emergence consists of two or more discrete events, we refer to the broad periods of strain accumulation and relief as supercycles rather than merely cycles"* and Goldfinger et al.'s (2013) analysis showing that large Cascadia subduction zone earthquakes reflect "*strain supercycles that transcend individual seismic cycling*."

Conceptually, the history of strain accumulation and release is the underlying process that gives rise to patterns in the resulting earthquakes and cumulative displacement. In the schematic example of Figure 1b, supercycles appear as patterns longer than individual earthquake cycles in the earthquake history, cumulative displacement, and cumulative strain records. The fullest picture is given by the strain record. This infers the strain by combining data about strain release via slip in earthquakes over time with the interseismic strain accumulation inferred from the slip between earthquakes taken from present-day geodetic, long-term geological, or other data. The cumulative displacement record shows the dates of earthquakes and the coseismic slip in each, whereas the earthquake history gives only the dates. The displacement record can be viewed as the time derivative of the strain record, and the earthquake history can be viewed as the time derivative of the displacement record, with each differentiation involving a loss of information. Conversely, constructing a displacement record requires supplementing an earthquake history with coseismic slip data, and constructing the strain record then involves also including data or assumptions about the interseismic strain accumulation. Hence the earthquake history has the least uncertainty, and the displacement and strain records have progressively larger uncertainties.

Supercycles are difficult to define precisely. One approach is to use major minima in the cumulative strain, which often mark the beginning of intervals during which few large earthquakes and hence little cumulative slip occurs. However, identifying major minima is often challenging and non-unique, especially given the assumptions needed to construct a strain history. Moreover, because data about interseismic strain accumulation and the slip in individual earthquakes are often unavailable, supercycles most often are inferred from an earthquake history that shows temporal clusters of seismicity, separated by intervals of lower seismicity or gaps without large earthquakes.

In this paper we use the term "supercycles" broadly, to describe long-term variability shown by aspects of the earthquake record that are difficult to reconcile with commonly used models of earthquake recurrence. The observation of supercycles, especially at plate boundaries and in plate boundary zones, is intriguing because plate boundaries are being loaded by steady plate motion.

We first review some proposed examples of supercycles on various faults, and show that these arise in the full range of tectonic environments - at plate boundaries, within plate boundary zones, and in plate interiors. We discuss the fact that supercycles and the sometimes-resulting earthquake clusters are not described by commonly used models of earthquake recurrence. We then introduce a model of Long-Term Fault Memory (LTFM), in which the probability of a large earthquake reflects the accumulated strain, and use it to explore many aspects of supercycles. Finally, we discuss the challenges supercycles pose for earthquake hazard assessment.

2. Examples of Supercycles

Supercycles and/or clustering have been observed in many tectonic environments (Figure 2). The best data come from earthquake histories at plate boundaries, because the relatively rapid plate motion (typically > 5 mm/yr) gives shorter and hence easier-to-observe cycles. In some cases, the slip and strain history also show evidence for supercycles.

Weldon et al. (2004) used the dates and offset in paleoearthquakes since 500 CE across the San Andreas fault near Wrightwood, California, together with the interseismic slip rate observed from present-day geodesy and long-term geological rates, to reconstruct the history of strain accumulation and release (Figure 2a). They argue that "*it is hard to escape the conclusion that strain accumulated over many earthquake cycles was responsible for the flurry of large slip* *events.*" Nearby, at Pallett Creek, Sieh et al. (1989) find that paleoearthquakes occurred in clusters within which they were "*separated by periods of several decades, but the clusters are separated by dormant periods of two to three centuries."* To the south, where the San Jacinto fault takes up some of the motion between the Pacific and North America plates, Rockwell et al. (2015) find that "*for much of the past 4,000 years the fault ruptured in a quasi-periodic fashion. In the past 1,000 years, in contrast, a flurry or cluster of four earthquakes occurred in a 150-year period, and the overall recurrence interval is much shorter.*"

Sieh et al. (2008) analyzed relative sea level changes recorded by corals from Sumatra, which show interseismic subsidence and coseismic uplift (Figure 2b). They infer that *"this 700-kilometer-long section of the Sunda megathrust has generated broadly similar sequences of great earthquakes about every two centuries for at least the past 700 years... Because each of the three past episodes of emergence consists of two or more discrete events, we refer to the broad periods of strain accumulation and relief as supercycles rather than merely cycles."* To the north along the subduction zone, Rubin et al. (2017) studied a 4500 year sequence of at least 11 tsunami deposits and find that "*the average time period between tsunamis is about 450 years with intervals ranging from a long, dormant period of over 2,000 years, to multiple tsunamis within the span of a century... these variable recurrence intervals suggest that long dormant periods may follow Sunda megathrust ruptures as large as that of the 2004 Indian Ocean tsunami."*

The dates and volumes of turbidite deposits, assumed to have been generated by great earthquakes on the Cascadia megathrust (Adams, 1990), show evidence for supercycles. Using these to infer the history of strain energy accumulation and release (Figure 2c), Goldfinger et al. (2013) find that "*the resulting sawtooth pattern reveals what we interpret as a complex pattern of long-term energy cycling on the Cascadia megathrust... Overall, what is suggested by this pattern is that some events release less energy, whereas others release more energy than available from plate convergence (slip deficit) and may have borrowed stored energy from previous cycles."* Although an additional event has been identified (Goldfinger et al., 2017), the inferred strain energy history would not be substantially altered. Kelsey et al.'s (2005) analysis of coastal deposits that record local tsunamis and seismic shaking finds that "*over the 4600 yr period when Bradley Lake was an optimum tsunami recorder, tsunamis from Cascadia plate-boundary earthquakes came in clusters."*

A somewhat different style of supercycles has been proposed for the Japan Trench off Tohoku (Figure 2*d*). Satake (2015) proposed that *"The 2011 Tohoku earthquake source includes the Miyagi-oki region, where M ~ 7.5 earthquakes repeated with average interval of 37 years. The typical slip of such large earthquakes is approximately 2 m, meaning that the cumulative coseismic slip is about 6 m per century. Because the subduction rate of the Pacific plate is approximately 8 m per century, 2 m slips may remain unreleased. Such a difference was previously interpreted as aseismic slip, but can be accumulated at the plate interface and cause a large coseismic slip of approximately 15 m with a recurrence interval of approximately 700 years... Such [a] supercycle model can explain the unusually large slip of the 2011 Tohoku earthquake. The term 'supercycle' was first used for a seismic cycle consisting of a series of large events, but often used for long-term cycle imposed on shorter cycles ('superimposed cycle')."*

The Sumatra and Tohoku records have interesting similarities and differences. In both, supercycles reflect infrequent events that have slip much greater than typical events. However, the Sumatra earthquake history has long gaps separating clusters, whereas for Tohoku smaller earthquakes occur frequently between the largest events, so the supercycles in the strain record do not appear in the earthquake history as gaps and clusters. It is worth noting that a Sumatra-type record could result if the detection limit in a paleoseismic record is too high to record the smaller events, or a Tohoku-type record could result if the recent rate of smaller events could not be extrapolated into the past.

In other areas, supercycles have been inferred from the earthquake history, even though the strain history requires data on the slip in individual events. Agnon's (2014) analysis of a long record of seismites, sediment records of earthquake shaking (Marco et al., 1996), along the Dead Sea Transform in Israel (Figure 2e) finds *"a pattern of long quiescence periods between quasi-periodic clusters. During each cluster of seismicity the recurrence interval is quite uniform, varying among clusters between 200 and 1,400 years. Quiescence periods may linger 3–10,000 years."* Further north on this transform, Wechsler et al. (2014) find that "*the interevent time of surface-rupturing earthquakes varies by a factor of two to four during the past 2 ka at our site, and the fault's behavior is not time predictable.*"

Supercycles and/or clustering have also been identified in plate boundary zones, where diffuse deformation is spread over multiple faults with long-term slip rates typically slower than on the primary plate boundary faults, and in continental interiors, which typically deform at < 1 mm/yr. As noted earlier, paleoseismic data from faults and groups of faults in the Western U.S.'s Great Basin (Wallace, 1987), part of the broad boundary zone that accommodates motion between the Pacific and North American plates, often show "*clustered strain release and uniform, low strain accumulation*" (Friedrich et al., 2003), shown schematically in Figure 2f.

Topographic data within the Australian plate, where erosion is very slow, provide some of the best evidence available of how continental intraplate faults slip over time, shown schematically in Figure 2g. Clark et al. (2012) found that *"a common characteristic of morphogenic earthquake occurrence in Australia appears to be temporal clustering. Periods of earthquake activity comprising a finite number of large events are separated by much longer periods of seismic quiescence, at the scale of a single fault and of proximal faults. In several instances there is evidence for deformation at scales of several hundred kilometers switching on and off over the last several million years."* As result, *"assigning an 'active/inactive' label to a fault in a slowly deforming area based upon the occurrence (or non-occurrence) of an event in the last few thousands to tens of thousands of years is not a useful indicator of future seismic potential"* (Clark et al., 2011) and *"it is debatable whether a 'recurrence interval' on individual faults applies" (*Clark, 2003).

These examples illustrate that long-term variability in earthquake behavior is a common effect, although the specifics vary between different areas. Hence in this paper, we take the view that observations of clustering likely reflect supercycles.

3. Earthquake recurrence models

The most easily studied aspect of supercycles is that they often - but need not always - cause variability in earthquake recurrence interval times, notably temporal clusters (Figure 2), which have important consequences for hazard estimation. As a result, many studies focus on possible clusters in a fault's earthquake history and their implications for the recurrence of future large earthquakes.

Neither of the commonly used classes of models for the recurrence of large earthquakes (Stein and Wysession, 2009) includes the possible effect of supercycles. The models are posed in terms of the conditional probability of an earthquake in a time period, based on a conceptual model of earthquake recurrence. The parameters for an area are inferred from its history of large earthquakes and the rate of smaller earthquakes. The models do not predict actual event timing, due to their stochastic nature.

One model treats earthquake occurrence as a Poisson process, in which the probability of a large earthquake is constant with time (Figure 3a). This probability depends on the mean recurrence interval μ, such that the probability of at least one event in a time interval t that is short compared to μ is t/μ. In this model the occurrence of a large earthquake does not reduce the probability of another. Hence the fault has no "memory," the dates of previous earthquakes have no effect on when the next will occur, and any clusters resulting from short intervals between events arise purely by chance. As the earthquake record's length increases, the standard deviation of the recurrence intervals approaches the mean. Equality of the mean and standard deviation of inter-event times is a property of a Poisson process, but - as shown later in the paper - other stochastic processes can also have this property. Because the Poisson model is the simplest recurrence model, it is traditionally used in earthquake hazard modeling and provides a null hypothesis against which other models can be tested (Rundle and Jackson, 1977; Smalley et al., 1987; Kagan and Jackson, 1991; Michael, 1997; Biasi et al., 2002).

An alternative class of probability models is based on the concept of an earthquake cycle (Figure 1), in which strain accumulates between large earthquakes and is completely released when one occurs (Reid, 1910; Savage and Burford, 1973; Sykes and Nishenko, 1984; Matthews et al., 2002; Field et al., 2015). In these models, the probability of a large earthquake increases with time until one occurs, at which point the probability drops to zero and the cycle begins again (Figure 3b). This assumption corresponds to the fault releasing all the strain accumulated on it in each cycle, so strain would not accumulate on timescales longer than individual cycles.

The length of time between earthquakes is described by one of a number of probability distributions (Gaussian, lognormal, Weibull, Brownian passage, etc.) for the recurrence times. The fault "remembers" only the last event, when the probability was renewed - reset to zero - so recurrence times in successive cycles are independent. Because the probability of a large earthquake depends only on the time since the past one, the fault has only "short-term memory." Renewal models are increasingly used in earthquake hazard analysis (WGCEP, 2003). The probability distributions describing the recurrence intervals are peaked around the average expected interval, so much longer or shorter intervals are rare, and earthquakes should occur quasi-periodically rather than in clusters. Thus as an earthquake record length increases, the standard deviation of the observed recurrence times should become small relative to their mean.

Hence clusters in an earthquake record could have various causes, each of which is likely to apply in some cases. First, they could be apparent clusters, artifacts of the limits of the paleoseismic record such as missing events or errors in earthquake dating (Weldon et al., 2005; Akciz et al., 2010). Second, if recurrence is described by Poisson or earthquake cycle models, clusters could result by chance when short recurrence intervals arise. Third, clusters could result from interactions between nearby faults or fault segments (Ward, 1992; Goes, 1996; Rundle et al., 2006; Dolan et al., 2016).

However, the fact that strain accumulation and/or clusters are observed on many fault systems has led to proposals that they are, at least in part, a real effect due to intrinsic properties of the faulting process (Ben-Zion et al., 1999). Hence in this paper, we take the view that observations of clustering likely reflect supercycles. We thus explore the possibility that faults have "long-term memory," such that the occurrence of large earthquakes depends on earthquake history over multiple previous earthquake cycles (Figure 3c).

Faults having long-term memory would have important consequences. Weldon et al. (2004) point out that *"resetting of the clock during each earthquake not only is conceptually important but also forms the practical basis for all earthquake forecasting because earthquake recurrence is statistically modeled as a renewal process (Cornell and Winterstein, 1988). In a renewal process, intervals between earthquakes must be unrelated so their variability can be expressed by (and conditional probabilities calculated from) independent random variables. Thus, if the next earthquake depends upon the strain history prior to that earthquake cycle, both our understanding of Earth and our forecasts of earthquake hazard must be modified... there can be little doubt that the simple renewal model of an elastic rebound driven seismic cycle will need to be expanded to accommodate variations that span multiple seismic cycles.*"

4. Characterizing Earthquake Sequences

4.1 Aperiodicity

In discussing long-term fault memory, it is useful to consider how earthquake sequences are characterized. A common characterization uses the aperiodicity, which measures the extent that a sequence differs from perfectly periodic. Aperiodicity, also termed the coefficient of variation (CV), is defined by $\alpha = \sigma/\mu$ where μ is the mean of the recurrence intervals (interevent times) and σ is their standard deviation (Kagan and Jackson, 1991; Goes, 1996; Vere-Jones 1970). An aperiodicity of zero corresponds to a perfectly periodic sequence, because $\sigma = 0$. An aperiodicity of one could correspond to a sequence produced by an ideal Poisson process with $\sigma = \mu$ but could also arise from other stochastic processes. $\alpha > 1$ corresponds to a "bursty" sequence that is so strongly clustered that $\sigma > \mu$. Because the entire range between perfectly periodic and perfectly Poissonian is termed "quasiperiodic," we divide it into the portion with $\alpha < 0.5$ as "strongly periodic" or "weakly aperiodic" - closer to purely periodic than purely Poissonian - and that with $1 > \alpha > 0.5$ as "weakly periodic" or "strongly aperiodic" - closer to purely Poissonian than purely periodic (Figure 4). Although sequences with $\alpha > 1$ are often termed "clustered," we use the term "bursty" because sequences with $\alpha < 1$ can be quite clustered, as discussed shortly.

A related characterization uses the burstiness parameter

$$B = (\alpha - 1)/(\alpha + 1) = (\sigma - \mu)/(\sigma + \mu)$$

(Goh and Barabasi, 2008). An ideal periodic sequence has B = -1, a perfectly Poisson sequence has B = 0, and bursty sequences have 0 < B < 1. Goh and Barabasi (2008) also characterize sequences by a memory parameter

$$M = \sum_{i=1}^{N-1} (\tau_i - \mu_1)(\tau_{i+1} - \mu_2)/\sigma_1\sigma_2$$

where N is the number of recurrence intervals τ_i, μ_1 and σ_1 are the mean and standard deviations of τ_i (i = 1, 2,...,N-1), and μ_2 and σ_2 are the mean and standard deviations of τ_{i+1} (i = 1, 2,..., N-1). M ranges from -1 to 1, with $M > 0$ when short interevent times are generally followed by short ones, and long interevent times are generally followed by long ones. $M < 0$ when short interevent times are generally followed by long ones, and vice versa. These arise because M is a normalized form of the autocorrelation of lag one, i.e. the crosscorrelation between the series of interevent times and that series shifted by one.

Figure 4 shows the aperiodicities for the earthquake sequences in Figure 2. The Wrightwood and Cascadia (Figures 2a and 2c) sequences have α = 0.47 and 0.51, so the time series alone do not indicate the supercycle behavior shown by the strain records. In contrast, the Sumatra and Dead Sea transform (Figures 2b and 2e) sequences have α = 1.05 and 1.6, indicating the supercycle behavior. The Great Basin and Australia sequences (Figures 2f and 2g) were described schematically without specific dates, so the aperiodicity illustrated is also schematic. Also shown is the global result from Goes (1996), who compiled 52 earthquake sequences from the San Andreas fault and the Middle America, Alaska, Chile, and Japan trenches. She found aperiodicities varying from 0.0 to 1.7, with "a large average aperiodicity" of 0.72 ± 0.36 that she interpreted as showing that earthquake recurrence is more irregular "than often assumed in hazard analysis."

These examples illustrate some of the issues in using aperiodicity to characterize sequences:

i) *Sequences with the same aperiodicity can be quite different.* Because the aperiodicity depends only on the mean and standard deviation of the interevent times, it does not depend on the order of events. Thus quite different sequences can have the same aperiodicity (Cowie et al., 2012). Figure 5a shows a sequence of paleoearthquakes composed of clusters of events several decades apart, separated by gaps of two to three centuries. The sequence has $\alpha = 0.79$, showing strong aperiodicity. Grouping the short-interval events together (Figure 5b) does not change α, but we would probably view the sequence as showing a change from longer recurrence times in the past to more recent short recurrence times. The memory parameter illustrates the difference, in that the more clustered sequence has a negative value, $M = -0.28$, whereas the grouped sequence has $M = 0.70$. This difference between the two sequences can also be seen in the interevent time plots shown to the right of each sequence. In these, major gaps appear as interevent times longer than the mean, which is shown by a horizontal line. In the first sequence, short and long intervals generally alternate, giving clusters and negative values of M. In the second sequence, short and long intervals are grouped, giving a positive memory.

ii) *Sequences with "quasiperiodic" aperiodicity can be quite clustered.* Earthquake sequences that we would consider clustered can fall below the nominal burstiness criterion of $\alpha > 1$. Figure 6b shows that lengthening the three major gaps in Figure 6a by 100 years increases the aperiodicity from 0.79 to 0.92, making the clustering stronger and the weak periodicity even weaker. Lengthening the gaps by 300 years (Figure 6c) increases the aperiodicity to 1.08. In all three panels we assume that observations begin at the earliest observed event (at the right side of the time axis showing years before present), so no gap is observed prior to the earliest event. This example illustrates that a sequence must be very strongly clustered to be bursty.

iii) *Sequences with aperiodicity close to 1 need not result from a Poisson process.* Earthquake records with aperiodicity close to 1 could resemble those that would be generated by a Poisson process. However, other stochastic processes, including the Long-Term Fault Memory process discussed later in this paper, can also generate earthquake records with interevent times whose mean and standard deviation are similar. Hence given the evidence in some areas of an underlying process involving strain supercycles, we think it useful to consider such sequences as clustered in many senses. In particular, considering clustering in such cases means that estimates of the probability that the next earthquake will occur within a given time window will depend crucially on whether the cluster is treated as ongoing or over.

iv) Aperiodicity can vary within an earthquake record. In particular, it is likely to be underestimated by short records. Because a short record is likely to contain events with recurrence times shorter than the mean of a longer record, shorter sequences underestimate aperiodicity (Ellsworth et al., 1999; Mucciarelli, 2007). This effect is seen in both synthetic catalogs (Ward, 1992) and earthquake records (Goes, 1996). Parsons (2008a) used Monte Carlo simulations to estimate the parameters of a parent distribution of recurrence times most likely to yield an observed time series. For example, an observed 1800-year-long earthquake record on the South Hayward fault with mean recurrence of 180 years and aperiodicity 0.48 is most likely to have arisen from a parent distribution with mean recurrence of 210 years and aperiodicity 0.6 (Parsons, 2008b).

4.2 Cluster Analysis

Another way to characterize earthquake sequences is through clustering. The statistical literature provides several criteria for defining a cluster and how many exist in a sequence. Categorizing clusters could facilitate definition of a supercycle, for example one cluster plus one gap. Hence we briefly review different clustering methods which either assign events to a cluster or choose the number of clusters. Clustering algorithms are broadly classified as either partitioning or hierarchical. To illustrate, we use Sieh et al.'s (1989) record from Pallett Creek, California (Figure 7a).

Partitioning methods such as the popular *k*-means algorithm are used to divide a sequence of observations, forming a given number of clusters, *k*, each observation assigned to one cluster.

Other methods, discussed later, are used to determine the number of clusters for a given sequence. In our application, the observations in a sequence are the dates, in years, of n earthquakes in an earthquake record and the clusters are defined as time intervals encompassing the range of dates. In the Pallett Creek record, n = 10. A k-means algorithm starts by guessing k cluster centers, which are averages of dates. The process then alternates two steps: 1) The closest cluster center is identified for each observation, measured by time in years between earthquakes and cluster centers, and the observation is assigned to that cluster. 2) Each cluster center is recalculated as the average date of its members (Hastie et al., 2009). This process repeats until it minimizes the sum or total within-cluster sum of squares (TWSS) of distances from cluster centers, i.e., it minimizes the sum of within-cluster variances of clusters i=1,…, k multiplied by the number of observations in that cluster, n_i (Hartigan, 2006). Commonly, this analysis is performed for a range of k and different aspects of the resulting cluster assignments are assessed to determine the number of clusters. The choice of k will strike a balance between too many clusters and not enough. The methods for choosing k do not always agree, as discussed next.

Some methods, such as the Elbow method, examine the graph of TWSS versus k (Figure 7b) and choose the value of k corresponding to a kink in the plot resembling a bent elbow (Tibshirani et al., 2001). Increasing k beyond this value conveys a lesser reduction in TWSS. By this method, Pallett Creek has 4 clusters. The Silhouette method compares the tightness (length of clusters) and separation (distance between clusters) to determine whether the cluster lengths are small compared to the distances between-clusters (Rousseeuw, 1987). Each observation receives a silhouette value, ranging from −1 to +1, indicating the extent to which the observation is well matched to its assigned cluster and poorly matched to the others. The number of clusters k is chosen to maximize the average values for all observations; again k = 4 for Pallett Creek (Figure 7c). The Gap method plots two curves that are functions of k, the logarithm of TWSS and its expected value under a uniform distribution of earthquake dates within the record (Tibshirani et al., 2001). The Gap statistic is the distance (gap) between the curves. The chosen value for k has the maximum Gap statistic, which again is k = 4 (Figure 7d).

Hierarchical methods do not rely on advance specification of the number of clusters, k, but rather create clusters for all k = 1, …, n possibilities. This process is illustrated by dendrogram plots (tree diagrams), showing the order in which different clusters are merged through connecting branches (Figure 7e). The vertical axis shows the cumulative difference in dates (in years) between cluster centers being merged. Hierarchical methods are of two kinds: 1) Divisive, in which all observations start in a single cluster, $k = 1$, and are iteratively separated until $k = n$. 2) Agglomerative nesting (termed AGNES), in which all observations start in their own cluster $(k = n)$, with the closest clusters (defined here by years between cluster centers) iteratively joined until $k = 1$ (Kaufman and Rousseeuw, 1990). AGNES may be better at identifying small clusters, while divisive methods may be better at identifying large clusters, although this choice makes no difference for our example. A popular AGNES algorithm, Ward's (1963) method, minimizes the within-cluster sum of squares using an update formula which assigns a new cluster's height on the vertical axis as the cumulative distance between the cluster centers being merged at that step and each step below it (Murtagh and Legendre, 2014). We show Ward's method because it is intended for interval-scaled data such as the dates of earthquakes (Kaufman and Rousseeuw,1990). Figure 7e shows the tree resulting from applying Ward's method to Pallett Creek. Clusters that merge at high levels on the vertical axis (indicating large distances between cluster centers being merged at that step) relative to the level of the clusters within them can be interpreted as a 'natural' number of clusters (Hastie et al., 2009). This determination is subjective, so in this example one could reasonably choose 2 or 4 clusters (Figure 7e). The four clusters {1,2,3},{4,5,6},{7,8},{9,10} are the same as obtained from k-means with $k = 4$.

Goldfinger's (2012) hierarchical clustering analysis on the 10,000-year-long Cascadia earthquake record found either four or five clusters, using AGNES with complete linkage (furthest neighbor) method. Furthest neighbor defines the distance between two clusters as the distance between two observations, one in each cluster, that are farthest away from one another (Tibrishani et al., 2001b). The two clusters with the shortest distance between them are merged at each step. Applying the complete linkage method to the Pallett Creek record yields the same tree as shown in Figure 7e using Ward's method. Goldfinger performed several tests of the statistical significance of the clusters with most resulting in a rejection of an underlying Poisson distribution. He cautions, however, "*there is no requirement that physical systems pass statistical tests*" (Goldfinger et al., 2012).

Hierarchical methods are complementary to partitioning methods such as *k*-means. For example, one can use the cluster centers from Ward's method as the initial cluster centers in *k* -means. As discussed above, results from *k*-means for chosen *k* can be compared to the results of AGNES. Our results are moderately robust to slight changes in dates, as illustrated by comparing the slightly differing dates of Pallett Creek from Sieh et al. (1989), Biasi et al. (2002), and Scharer et al. (2011). The tree diagrams are the same, because they largely reflect only the events' order. Differences in the *k*-means evaluations are shown in Table 1. The Gap statistic for Scharer et al.'s dates yields 1 cluster, and the next best number is 4, with the difference between their statistics being quite small compared to the differences between other numbers of clusters. A similar situation occurs in the silhouette for the Sieh et al. dates (Figure 7c) where one could argue for 2, 3, or 4 clusters because of the similar values. The Elbow method is the most stable between these different records and the Gap statistic is the least.

Table 1: Differences in the number of clusters indicated by three methods for records of earthquakes at Pallett Creek with slightly differing dates.

Record	Gap Statistic	Elbow	Silhouette
Sieh et al., 1989	4	4	4
Biasi et al., 2002	2	4	2
Scharer et al., 2011	1	4	4

5. Long-Term Fault Memory Model

To explore how earthquake supercycles and clusters arise, we use a simple Long-Term Fault Memory (LTFM) model, which is a modified version of the standard earthquake cycle model. In it, the probability of an earthquake reflects the accumulated strain. This increases steadily with time until an earthquake happens, after which it decreases, but not necessarily to zero (Figure 8). Hence, the probability of an earthquake depends on the earthquake history over multiple prior cycles. Clusters happen because after a gap, a period of quiescence, the probability can remain higher than the long-term average for several cycles. The model simulates large earthquakes releasing only part of the strain accumulated on the fault, in contrast to the standard model in which all of the accumulated strain is released.

LTFM is a simple model with only a few parameters. The annual probability *P(t)* grows with time at rate $dP/dt = A = 2/\tau^2$, simulating steady strain accumulation. τ is an initial mean recurrence interval, such that if no earthquake occurs during the initial time interval $t = \tau$, the average annual probability is $1/\tau$. If the probability is above a threshold value δ, which we typically set as zero, an earthquake can occur. When an earthquake occurs, the probability drops by $\Delta P = -R$, simulating a partial strain release. Hence on average R/A years of accumulated strain is released in an earthquake.

The accumulation parameter A controls the long-term seismicity rate, and the release parameter R controls the clustering. Small R yields long-term memory and more clusters, whereas in the limit large R gives the standard earthquake cycle model with only short-term memory because it forces the probability to zero after each earthquake. The probability is not allowed to go below 0 or to exceed 1.

We generate earthquake histories by using the Mersenne Twister pseudo-random number generator (Matsumoto and Nishimura, 1998), sampling from a uniform distribution between 0 and 1. If the value exceeds the probability for that year, no earthquake occurs and the probability increases by A for the next year. If the value is less than that year's probability, an earthquake occurs and the probability drops by R for the following year. Linearly increasing probabilities have been used by other authors, e.g., Pinedo and Shpilberg (1981).

The saw-tooth behavior of LTFM simulates the proposed long-term variations in stored elastic strain or strain energy (Figure 2). Supercycles and clusters arise because longer intervals between earthquakes generally begin at times of low probability, consistent with the pattern noted in terms of cumulative strain by Weldon et al. (2004). A lower probability corresponds to the fault having less memory of previous earthquakes. Thus, as the probability (i.e. cumulative strain) approaches zero, the corresponding supercycle can be viewed as approaching a renewal process.

Because LTFM is a stochastic model, the resulting earthquake sequences depend on both the model parameters and chance. As a result (Figure 9) sequences can appear strongly periodic, weakly periodic, Poissonian, or bursty. The four sequences in this example have the same probability (i.e. strain) accumulation rate (A = $2/125^2$) but different release parameters (R = $200A$, $175A$, $80A$, $50A$). As shown, the aperiodicity increases as R decreases. The strongly periodic sequence arises in a way similar to a standard earthquake cycle model because R is so large that the probability drops to zero after each earthquake, so the fault has no memory. The effects of fault memory increase for successively smaller values of R, making the sequences less periodic. However, A and R control only the overall sequence properties via the probability of earthquake occurrence, because when earthquakes occur is random. As a result, the aperiodicity varies between different portions of the sequence.

In some cases, we use two thresholds, $\delta_2 > \delta_1$ and corresponding probability drops, $R_2 > R_1$, to describe the earthquakes with larger and smaller strain changes implied by some records (Figure 2). Hence if $P(t) > \delta_i$, the probability drops by R_i. Using two probability thresholds and probability drops to describe both rare larger and more frequent smaller strain changes allows LTFM to simulate the range of observed supercycle behavior (Figure 10). The higher threshold and probability drop simulate infrequent events that have slip and strain release much greater than typical events, and so end a supercycle. Using two similar thresholds simulates a Sumatra-style earthquake history with long gaps separating clusters, because earthquakes can occur only late in a supercycle. This case would correspond to a very strong fault. Conversely, a low threshold for smaller earthquakes and a much higher one for larger earthquakes simulates a Tohoku-style record where smaller earthquakes occur frequently between the largest events, so the supercycles in the strain record do not appear in the earthquake history as gaps and clusters. The threshold and drop parameters can be chosen to simulate the very long gaps associated with intraplate and plate boundary zone earthquakes. In such situations, because strain accumulates slowly relative to plate boundaries, the lower threshold is quite low. Hence this threshold can be used in most applications with higher strain rates (e.g., Figure 10b), since it would have essentially the same effect as a zero threshold (e.g., Figure 9).

5.1 Example

To explore choosing LTFM parameters to match key aspects of an earthquake history, we simulated the record from Pallett Creek, California. Although recent studies have reestimated the dates (Biasi et al., 2002; Scharer et al., 2011), we used Sieh et al.'s (1989) dates because the resulting clusters provide a better test case. We ran the model 100 times for pairs of input parameters, R and τ, and averaged the mean and standard deviation of recurrence intervals for each pair. Contouring these averages identified regions of the model space, and hence ranges of the input parameters, that produce simulations with comparable mean and standard deviation to those observed in the paleoseismic record. We then searched these regions for parameters giving a memory parameter close to that observed.

A simulation with accumulation rate $A = 2/289^2$ and release parameter $R = 130A$ that gives clustering behavior similar to that observed is shown in Figure 11. The data have $\mu = 132$ yr, $\sigma = 105$ yr, $\alpha = 0.79$ and $M = -0.28$, and the simulation has $\mu = 136$ yr, $\sigma = 102$ yr, $\alpha = 0.75$ and $M = -0.33$, indicating weak periodicity. The event timing differs between the simulation and the observed record due to the model's stochastic nature. The longer intervals between earthquakes begin at times of low probability, consistent with the pattern noted in terms of cumulative strain by Weldon et al. (2004).

We used the LTFM model to explore the long-term variability of fault behavior by creating simulations much longer than paleoseismic records, and then sampling them for intervals corresponding to paleoseismic records. Figure 12 shows results for a 50,000 year long simulation using parameters appropriate for Pallet Creek. The mean and standard deviation of recurrence times averaged over a moving 1345-year window, corresponding to a paleoseismic record, are relatively stable over long time periods. This stability would be consistent with the idea of steady loading and unloading by plate motion and large earthquakes. However the mean and standard deviation of recurrence times vary somewhat. The aperiodicity shows that the simulated paleoseismic record sometimes appears strongly periodic (standard deviation small relative to the mean) implying a seismic cycle model, while at other times it looks weakly periodic, Poissonian (standard deviation similar to the mean), or bursty. This variability is illustrated by the earthquake history between model years 19,000 and 22,000. Hence the recurrence variability due to long-term fault memory can give rise to paleoseismic records that at different times appear to have different underlying statistical distributions. Thus a given paleoseismic or instrumental window may give a biased view of the long-term seismicity.

5.2 LTFM and intraplate earthquakes

Long-term fault memory may also be an important contributor to the space-time variability of continental intraplate earthquakes. Considerable recent attention (reviewed by Liu and Stein, 2016, Calais et al., 2016, and Stein et al., 2017a) has been directed to how and why earthquakes within continents behave differently in space and time from those on plate boundaries. Faults at plate boundaries are loaded at constant rates by relatively rapid and steady relative plate motion. Consequently, earthquakes concentrate along the plate boundary faults and show quasi-periodic (relative to intraplate earthquakes) occurrences, although the actual temporal patterns are often complicated. The spatial gaps that appear are filled in over time.

However, in mid-continents, the slower tectonic loading is shared by a complex system of interacting faults spread over a large region, such that a large earthquake on one fault could increase the loading rates on other faults in the system. Because the low tectonic loading rate is shared by many faults, individual faults may remain dormant for a long time before they accumulate enough strain for a short period of activity. The resulting earthquakes are therefore episodic, clustered, and spatially migrating (Li et al., 2009; Stein et al., 2009). These effects can be seen in many areas, including North China (Liu et al., 2011), Europe (Camelbeeck et al., 2007; 2014), and the central United States (Crone and Luza, 1990; Newman et al., 1999; Holbrook et al., 2006; Tuttle et al., 2006; Gold et al., 2018).

Topographic data from Australia, where erosion is very slow, provide some of the best evidence available of how intraplate faults slip over time. Figure 2f illustrates this pattern of clusters of activity separated by much longer and irregular intervals of quiescence. Liu and Stein (2016) note that the pattern of displacement accumulated over time is similar to the Devil's Staircase function, a fractal property of chaotic dynamic systems (Devaney et al., 1989; Turcotte, 1997). The apparent long-distance roaming of large mid-continental earthquakes also suggests dynamic system behavior. In such a system, change of any part of the system (such as rupture of a fault) could impact nonlinearly the behavior of the whole system.

Although this view of intraplate seismicity fits what is known in general terms, the specifics are still unclear. In particular, how effectively stress can be transferred to distant faults is unknown. We thus used the LTFM model to explore the possibility that long-term fault memory may also contribute to the space-time variability.

A noticeable difference between the clustering in Australia and that on plate boundaries is that in Australia the gap durations are more than ten times as long as the clusters, whereas on plate boundaries the gaps are only 2-3 times as long as the clusters. As shown in Figure 9, LTFM can describe this effect via assuming the level of strain accumulation required for an earthquake. A proposed alternative is that clusters of large intraplate earthquakes reflect the fault weakening after the first major event, so as to permit repeated failure (Li et al., 2009). Models have been proposed for how weakening and subsequent healing might occur (Sibson, 1992;

Lyakhovsky et al., 2001). In Lyakhovsky et al.'s model, as the rate ratio between loading and healing increases, behavior changes from regular to clustered. This is because healing tends to purge long-term memory**.** These models were developed with a view toward describing the evolution of fault properties over multiple earthquake cycles, i.e. a different type of long-term fault memory.

5.3 LTFM model discussion

Our results illustrate that a modified version of the standard earthquake cycle model can be used to simulate and explore key features of supercycles that are observed at many plate boundaries and in plate interiors. This is gratifying, given the model's simplicity. LTFM can be thought of as an idealized model like those used in many disciplines, including physics, astronomy, meteorology, biology, and economics, that allow investigations to focus on some key characteristics of a complex phenomenon and explore whether they can be explained by simple assumptions. Reutlinger et al. (2018) explain that *"we call such models 'toy models'—a term that is not meant to have belittling or derogatory connotations... First, models of this type are strongly idealized... Second, such models are extremely simple in that they represent a small number of causal factors (or, more generally, of explanatory factors) responsible for the target phenomenon. Third, these models refer to a target phenomenon."* A good example would be the simple analytical model of subduction zones that extracts key aspects of sophisticated numerical models and thus can be used how the temperature structure and resulting plate driving force depend on the age of the subducting plate and convergence rate (Stein and Wysession, 2009)

In this spirit, we have used a simple model that simulates general properties of supercycles. We plan to explore its possible applicability to paleoseismic records in other areas and in different tectonic regimes. For example, clusters have been observed in paleoseismic data in plate boundary zones, where diffuse deformation is spread over multiple faults and long-term slip rates are slower than on primary plate boundary faults (which typically move at > 10 mm/yr) but higher than in continental interiors (which typically deform at < 1 mm/yr) (Wallace, 1987; Rockwell et al., 2000; Friedrich et al., 2003; Oskin et al., 2008; Dolan et al., 2016; Gold et al., 2017). Some clusters seem to arise on individual faults, whereas others involve groups of faults. The Wasatch fault and adjacent faults show a strain release and slip pattern similar to that in Australia (Figure 2e) (Wallace, 1987; Friedrich et al., 2003). In the Eastern California shear zone, regional strain release appears to occur via "*distinct periods or bursts of seismic activity punctuated by periods of relative quiescence. Individual faults, however, appear to behave in a quasiperiodic fashion, with the clustering produced by the in-phase earthquake generation of the system of faults*" (Rockwell et al., 2000). Hence LTFM may be involved in plate boundary zone faults, but fault interactions and changes in loading across the zone may also contribute.

Additional features could be added to the model without overcomplicating it. Its current form allows for two classes of earthquakes causing different probability decreases, or strain releases. In some cases, only one may be needed, as motivated by observations that slip in large events on individual fault segments appears similar (Schwartz and Coppersmith, 1984) and Weldon et al.'s (2004) observation that on the area of the San Andreas they studied "*there appears to be no relationship between strain level and the size of earthquakes*." However, Goldfinger et al. (2013) note a "weak tendency" for clusters to terminate with an "outsized" event, as found for the Tohoku and Sumatra records (Figure 2). Moreover, some of the strain release may occur via slow slip events (Rogers and Dragert, 2003; Jiang et al., 2017) that may not appear in the paleoseismic record.

Fault interactions could be introduced into the model by having multiple faults that affect the probability of large earthquakes on each other. In some situations these may increase clustering, and in others they may reduce it. This effect is likely to contribute to the variability in earthquake size often observed at subduction zones (Thatcher, 1990; Stein and Okal, 2007). One example is the trench segment that produced the *M*w ~9.6 1960 Chilean earthquake. Its rupture mode must be variable because the seismic-slip rate inferred assuming that the 1960 earthquake is this segment's characteristic earthquake exceeds the convergence rate. Hence Stein et al. (1986) proposed that either the characteristic earthquake is smaller than the 1960 event, the average recurrence interval is greater than observed in the past 400 years, or both. Recent paleoseismic studies support this analysis (Cisternas *et al.*, 2005). Paleoseismic studies also find evidence for variable size of thrust events, presumably due to the differences between multisegment and single-segment rupture, in areas including the Nankai Trough (Ando, 1975) and the Kuril trench (Nanayama et al., 2003).

Viewing supercycles as a result of long-term fault memory fits into a general framework in the literature of complex dynamic systems. Clustered events, described as "bursts," are observed in many disparate systems, from the firing system of a single neuron to the outgoing mobile phone sequence of an individual (Karsai et al., 2012). Such systems display "…*a bursty, intermittent nature, characterized by short timeframes of intense activity followed by long times of no or reduced activity*," (Goh and Barabasi, 2008). As a result, the system's state depends on its history, so it has long-term memory (Beran et al., 2013).

An additional point worth noting is that we generally limit our discussion to cases where the supercycle is shorter than the climatic forcing cycles such as global glaciation periods.

5.4 Mathematics of the LTFM model

The LTFM model is a stochastic process, specifically a Markov chain with a finite number of states at discrete times 0, 1, 2, . . . The states correspond to possible values of accumulated strain, reflected in the probability *P(t),* which are finite in number.[1] The probability that an earthquake occurs at time t, conditional on the full history of strain accumulation and release at all times prior to t, depends only on the most recent level of strain, i.e., at time $t-1$.[2] Given *P(t),* the probability does not otherwise depend on time. Thus, the history prior to t is fully captured by *P(t-1).* The process starts over each time accumulated strain is equal to the strain at time $t=0$ (or, for practical purposes, is close to that amount). The length of time until the process starts over can be interpreted as the length of a supercycle. The theory of Markov chains (Çinlar, 1975) allows us to directly specify the full probability distribution for the length of a supercycle, and hence its mean and standard deviation. The theory also allows us to specify the conditional probability of an earthquake at a time $t>s$ given the accumulated strain at current time *s.* The theory implies that the probability at a far future time t does not depend on the accumulated strain at time s and provides a formula for that probability. From this probability, the expected number of earthquakes in a distant time span of length T can be calculated, along with the approximate standard deviation.

---

[1] Possible values of P(t) have the form $\min\left\{(\alpha A-\beta R_1-\gamma(\lambda-1)R_1)^+,1\right\}$ with $\lambda = R_2 / R_1$, $(x)^+ = \max\left\{x,0\right\}$, and $\alpha,\ \beta,\ \gamma$ taking non-negative integer values.

[2] If the probability at time s is *P(s)* = C_s then the conditional probability of an earthquake at time t is equal to $\min\left\{(A+C_{t-1})\times\chi_{pos}(A+C_{t-1}-\delta_1),1\right\}$, with $\chi_{pos}(x)$ equal to 1 if $x>0$ and equal to 0 otherwise.

LTFM can also be posed in terms of the classic probability model of drawing balls from an urn. (Stein and Stein, 2013). If *e* balls are labeled "E" for earthquake, and *n* balls are labeled "N" for no earthquake, the probability of an earthquake is that of drawing an E-ball, which is the ratio of the number of E-balls to the total number of balls. If after drawing a ball, we replace it, the probability of an event is constant or time-independent in successive draws, because one happening does not change the probability of another happening. Thus an event is never "overdue" because one has not happened recently, and the fact that one happened recently does not make another less likely. LTFM corresponds an alternative, sampling such that the fraction of E-balls and the probability of another event change with time. We add *A* E-balls after a draw when an earthquake does not occur, and remove *R* E-balls when an earthquake occurs. This makes the probability of an event increase with time until one happens, after which it decreases and then grows again. Events are not independent, because one happening changes the probability of another.

6. Implications for hazard assessment

Advances in understanding supercycles would be important for seismic hazard assessment. Such assessments depend heavily on assumptions about the magnitude and recurrence rate of future large earthquakes (Stein et al., 2012), both of which are often more variable than assumed. A larger assumed aperiodicity will cause cumulative or conditional probabilities to decrease, all else fixed (Ward, 1992).

Current earthquake probability estimates depend on assuming a probability density function for the recurrence intervals with input parameters inferred from the available earthquake history. Figure 13 illustrates the resulting uncertainties for Cascadia. Figure 13a shows the effects of additional paleoseismic data. Goldfinger et al.'s (2012) chronology yielded a mean recurrence interval of 530 yrs and a standard deviation of 271 yrs for the entire 10,000 year record, and a mean recurrence interval of 326 yrs and a standard deviation of 88 yrs for the most recent cluster. Including a newly-identified event in a revised chronology (Goldfinger et al., 2017) has a small effect on the 10,000-year record's parameters, changing the mean recurrence interval to 502 yrs and a standard deviation of 239 yrs. However, adding this event makes all events in the past 5,000 years part of the same cluster, with a recurrence interval of 452 yrs and a standard deviation of 142 yrs.

Whether to assume that a recent cluster is continuing or has ended can lead to quite different estimates of earthquake probabilities (Stein et al., 2017b). Figure 13b shows the different distribution of recurrence intervals corresponding to the two different chronologies and various probability density functions with parameters corresponding to the two chronologies. By far the largest difference arises from assuming either that the recent cluster continues, or that the cluster is over so the appropriate parameters are those for the entire record. Assuming that we are still in the cluster predicts higher probability than using the entire record. This effect is more important than the specific probability density function assumed. The corresponding effect appears from considering the conditional probability of a large earthquake in the next 50 years, which results from integrating the probability density functions (Figure 13c).

More generally, if faults have long-term memory, then individual earthquake cycles, and hence the recurrence times between successive large earthquakes, are not independent. Hence the renewal approach of modeling their probability as a function of time since the previous large earthquake can give misleading results. The problem is not that a renewal model is inappropriate, but rather that the renewal depends on release of nearly all accumulated strain, and that may occur at different times than large earthquakes. As shown in Figure 12, the recurrence variability due to long-term fault memory can cause short earthquake records to give a biased view of the long-term seismicity. As a result, further investigation of long-term earthquake recurrence variability is important both for understanding the nature and causes of supercycles and for improving hazard assessment.

Acknowledgements

This research was supported by Northwestern University's Institute for Policy Research and Buffet Center for International Studies. Liu acknowledges support of NSF grant 1519980.

References

Adams, J., 1990. Paleoseismicity of the Cascadia subduction zone: Evidence from turbidites off the Oregon-Washington margin, Tectonics 9(4), 569-583.

Agnon, A., 2014. Pre-instrumental earthquakes along the Dead Sea Rift, In: Z. Garfunkel et al. (eds.), Dead Sea Transform Fault System: Reviews, Modern Approaches in Solid Earth Sciences 6, doi: 10.1007/978-94-017-8872-4_8

Akciz, S. O., Ludwig, L. G., Arrowsmith, J. R., Zielke, O., 2010. Century-long average time intervals between earthquake ruptures of the San Andreas fault in the Carrizo Plain, California, Geology 38 (9), 787-790, doi: 10.1130/G30995.1.

Ando, M., 1975. Source mechanisms and tectonic significance of historical earthquakes along the Nankai Trough, Japan, Tectonophysics 27, 119–140.

Ben-Zion, Y., K. Dahmen, V. Lyakhovsky, D. Ertas, Agnon, A., 1999. Self-driven mode switching of earthquake activity on a fault system, Earth Planet. Sci. Lett. 172(1–2), 11–21.

Beran, J., Feng, Y., Ghosh, S., Kulik, R., 2013. Long-Memory Processes, Springer-Verlag Berlin Heidelberg, doi: 10.1007/978-3-642-35512-7_1.

Biasi, G. P., Weldon, R. J., Fumal, T. E., Seitz, G. G., 2002. Paleoseismic event dating and the conditional probability of large earthquakes on the southern San Andreas Fault, California, Bull. Seismol. Soc. Am. 92(7), 2761-2781.

Calais, E., Camelbeeck, T., Stein, S., Liu, M., Craig, T.J., 2016. A new paradigm for large earthquakes in stable continental plate interiors, Geophys. Res. Lett. 4320, 10-621.

Camelbeeck, T., Vanneste, K., Alexandre, P., Verbeeck, K., Petermans, T., Rosset, P., Everaerts, M., Warnant, R., Van Camp, M., 2007. Relevance of active faulting and seismicity studies to assess long term earthquake activity in northwest Europe, In: Stein, S., Mazzotti, S. (Eds.), Continental Intraplate Earthquakes: Science, Hazard, and Policy Issues, Geol. Soc. Amer. Special Paper 425, 193–224.

Camelbeeck, T., Alexandre, P., Sabbe, A., Knuts, E., Garcia Moreno, D., Lecocq, T., 2014. The impact of the earthquake activity in western Europe from the historical and architectural heritage, In: Talwani, P. (Ed.), Intraplate Earthquakes -Solid Earth Geophysics. Cambridge University Press.

Çinlar, E., 1975. Introduction to Stochastic Processes, Prentice Hall, Englewood Cliffs, N.J.

Cisternas, M., Atwater, B.F., Torrejón, F., Sawai, Y., Machuca, G., Lagos, M., Eipert, A., Youlton, C., Salgado, I., Kamataki, T., Shishikura, M., 2005. Predecessors of the giant 1960 Chile earthquake, Nature 437(7057), 404.

Clark, D., McCue, K., 2003. Australian paleoseismology: towards a better basis for seismic hazard estimation, Ann. Geophys. 46(5), 1087–1106.

Clark, D., McPherson, A., Collins, C., 2011. Australia's seismogenic neotectonic record, Geosci. Aust. Rec. 11.

Clark, D., McPherson, A., Van Dissen, R., 2012. Long-term behavior of Australian stable continental region (SCR) faults, Tectonophysics 566, 1–30.

Cornell, C.A., Winterstein, S.R., 1988. Temporal and magnitude dependence in earthquake recurrence models, Bull. Seismol. Soc. Am. 78(4), 1522-1537.

Cowie, P.A., Roberts, G.P., Bull, J.M., Visini, F., 2012. Relationships between fault geometry, slip rate variability and earthquake recurrence in extensional settings, Geophysical Journal International 189(1), 143-160.

Crone, A.J., Luza, K.V., 1990. Style and timing of Holocene surface faulting on the Meers fault, southwestern Oklahoma, Geol. Soc. Amer. Bull. 102, 1–17.

Devaney, R.L., Devaney, L., Devaney, L., 1989. An Introduction to Chaotic Dynamical Systems, Addison-Wesley, Reading.

Dolan, J.F., McAuliffe, L.J., Rhodes, E.J., McGill, S.F., Zinke, R., 2016. Extreme multi-millennial slip rate variations on the Garlock fault, California: Strain super-cycles, potentially time-variable fault strength implications for system-level earthquake occurrence, Earth Planet. Sci. Lett. 446, 123-136.

Ellsworth, W.L., Matthews, M.V., Nadeau, R.M., Nishenko, S.P., Reasenberg, P.A., Simpson, R.W., 1999. A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities, US Geol. Surv. Open-File Rept. 99, 522.

Field, E. H., et al., 2015. Long-term time-dependent probabilities for the Third Uniform California Earthquake Rupture Forecast (UCERF3), Bull. Seismol. Soc. Am. 105(2A), doi: 10.1785/0120140093.

Friedrich, A.M., Wernicke, B.P., Niemi, N.A., Bennett, R.A., Davis, J.L., 2003. Comparison of geodetic and geologic data from the Wasatch region, Utah, and implications for the spectral character of Earth deformation at periods of 10 to 10 million years, J. Geophys. Res.: Solid Earth 108(B4).

Goes, S. D.B., 1996. Irregular recurrence of large earthquakes: An analysis of historic and paleoseismic catalogs, J. Geophys. Res. 101(B3), 5739-5749.

Goh, K.I., Barabasi, A.L., 2008. Burstiness and memory in complex systems, Europhys. Lett. 81(4), 48002, doi: 10.1209/0295-5075/81/48002.

Gold, R.D., Cowgill, E., Arrowsmith, J.R., Friedrich, A.M., 2017. Pulsed strain release on the Altyn Tagh fault, northwest China, Earth Planet. Sci. Lett. 459, 291-300.

Gold, R. D., Duross, C. B., Delano, J., Jibson, R. W., Briggs, R. W., Mahan, S. A., Williams, R., 2018. Four major Holocene earthquakes on the Reelfoot Fault, New Madrid Seismic Zone, GSA Annual Meeting (abstract).

Goldfinger, C., Nelson, C. H., Morey, A., Johnson, J. E., Gutierrez- Pastor, J., Eriksson, A. T., Karabanov, E., Patton, J., Gracia, E., Enkin, R., Dallimore, A., Dunhill, G., Vallier, T., 2012. Turbidite event history: Methods and implications for Holocene Paleoseismicity of the Cascadia Subduction Zone, U.S. Geol. Survey Professional Paper 1661-F, 184.

Goldfinger, C., Ikeda, Y., Yeats, R.S., Ren, J., 2013. Superquakes and Supercycles, Seismol. Res. Lett. 84(1), 24-32.

Goldfinger, C., Galer, S., Beeson, J., Hamilton, T., Black, B., Romsos, C., Patton, J., Nelson, C.H., Hausmann, R., Morey, A., 2017. The importance of site selection, sediment supply, and hydrodynamics: A case study of submarine paleoseismology on the Northern Cascadia margin, Washington USA, Marine Geology 384, 4-46.

Hartigan, J.A., 2006. Classification – 1, Encyclopedia of Statistical Sciences, 1-9, John Wiley & Sons, Inc., doi: 10.1002/0471667196.ess0356.pub2.

Hastie, T., Tibshirani, R., and Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed, Springer, New York, doi: 10.1007/b94608.

Holbrook, J., Autin, W.J., Rittenour, T.M., Marshak, S., Goble, R.J., 2006. Stratigraphic evidence for millennial-scale temporal clustering of earthquakes on a continental-interior fault: Holocene Mississippi River floodplain deposits, New Madrid seismic zone, USA, Tectonophysics 420(3–4), 431–454.

Jiang, Y., Liu, Z., Davis, E.E., Schwartz, S.Y., Dixon, T.H., Voss, N., Malservisi, R. Protti, M., 2017. Strain release at the trench during shallow slow slip: The example of Nicoya Peninsula, Costa Rica, Geophys. Res. Lett. 44(10), 4846-4854.

Kagan, Y. Y., Jackson, D. D., 1991. Long-Term earthquake clustering, Geophys. J. International. 104, 117-133.

Karsai, M., Kimmo, K., Barabasi, A., Kertesz, J. 2012. Universal features of correlated bursty behavior, Scientific Reports 2, 397, doi: 10.1038/srep00397.

Kaufman, L. and Rousseeuw, P.J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis, Wiley-Interscience, New Jersey.

Kelsey, H.M., Nelson, A.R., Hemphill-Haley, E., Witter, R.C., 2005. Tsunami history of an Oregon coastal lake reveals a 4600 yr record of great earthquakes on the Cascadia subduction zone, Geological Society of America Bulletin 117(7-8), 1009-1032.

Li, Q., Liu, M., Stein, S., 2009. Spatial-temporal complexity of continental intraplate seismicity: Insights from geodynamic modeling and implications for seismic hazard estimation, Bull. Seismol. Soc. Am. 99, doi: http://dx.doi.org/10.1785/0120080005.

Liu, M., Stein, S., 2016. Mid-continental earthquakes: Spatiotemporal occurrences, causes, and hazards, Earth-Science Reviews 162, 364-386.

Liu, M., Stein, S., Wang, H., 2011. 2000 years of migrating earthquakes in north China: How earthquakes in midcontinents differ from those at plate boundaries, Lithosphere 3(2), 128–132.

Lyakhovsky, V., Ben-Zion, Y., Agnon, A., 2001. Earthquake cycle, fault zones, and seismicity patterns in a rheologically layered lithosphere, J. Geophys. Res.: Solid Earth 106(B3), 4103-4120.

Marco, S., Stein, M., Agnon, A., Ron, H., 1996. Long-term earthquake clustering: A 50,000-year paleoseismic record in the Dead Sea Graben, J. Geophys. Res.: Solid Earth 101(B3), 6179-6191.

Matsumoto, M., Nishimura, T., 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator, ACM Transactions on Modeling and Computer Simulation 8(1), 3-30, doi:10.1145/272991.272995.

Matthews, M. V., Ellsworth, W. L., Reasenberg, P. A., 2002. A Brownian model for recurrent earthquakes, Bull. Seismol. Soc. Am. 92(6), 2233-2250.

Michael, A. J., 1997. Testing prediction methods: Earthquake clustering versus the Poisson model, Geophys. Res. Lett. 24(15), 1891-1894.

Mucciarelli, M., 2007. A test for checking earthquake aperiodicity estimates from small samples, Natural Hazards and Earth System Science 7(3), 399-404.

Murtagh, F. and Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?, Journal of Classification 31, 274-295, doi: 10.1007/s00357-014-9161-z.

Nanayama, F., Satake, K., Furukawa, R., Shimokawa, K., Atwater, B.F., Shigeno, K., Yamaki, S., 2003. Unusually large earthquakes inferred from tsunami deposits along the Kuril trench, Nature 424(6949), 660.

Newman, A., Stein, S., Weber, J., Engeln, J., Mao, A., Dixon, T., 1999. Slow deformation and lower seismic hazard at the New Madrid seismic zone, Science 284(5414), 619–621.

Oskin, M., Perg, L., Shelef, E., Strane, M., Gurney, E., Singer, B., Zhang, X., 2008. Elevated shear zone loading rate during an earthquake cluster in eastern California, Geology 36(6), 507-510.

Parsons, T., 2008a. Monte Carlo method for determining earthquake recurrence parameters from short paleoseismic catalogs: Example calculations for California, Journal of Geophysical Research: Solid Earth 113(B3), doi:10.1029/2007JB004998.

Parsons, T., 2008b. Earthquake recurrence on the south Hayward fault is most consistent with a time dependent, renewal process, Geophysical Research Letters 35(21).

Pinedo, M. and Shpilberg, D., 1981. Stochastic models with memory for seismic risk, The Journal of Risk and Insurance 48(1), 46-58.

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

Reutlinger, A., Hangleiter, D., Hartmann, S. 2018. Understanding (with) toy models, British J. for Philosophy of Science 69, 1069–1099.

Rockwell, T.K., Lindvall, S., Herzberg, M., Murbach, D., Dawson, T., Berger, G., 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(5), 1200-1236.

Rockwell, T.K., Dawson, T.E., Ben-Horin, J.Y., Seitz, G., 2015. A 21-event, 4,000-year history of surface ruptures in the Anza seismic gap, San Jacinto Fault, and implications for long-term earthquake production on a major plate boundary fault, Pure and Applied Geophys. 172(5), 1143-1165.

Rogers, G., Dragert, H., 2003. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip, Science 300(5627), 1942–1943, doi:10.1126/science.1084783.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Journal of computational and Applied Mathematics 20, 53-65.

Rubin, C.M., Horton, B.P., Sieh, K., Pilarczyk, J.E., Daly, P., Ismail, N., Parnell, A.C., 2017. Highly variable recurrence of tsunamis in the 7,400 years before the 2004 Indian Ocean tsunami, Nature Communications 8, 16019.

Rundle, J. B., Jackson, D. D. 1977. Numerical simulation of earthquake sequences, Bull. Seismol. Soc. Am. 67(5), 1363-1377.

Rundle, J.B., Rundle, P.B., Donnellan, A., Li, P., Klein, W., Morein, G., Turcotte, D.L., Grant, L., 2006. Stress transfer in earthquakes, hazard estimation and ensemble forecasting: inferences from numerical simulations, Tectonophysics 413(1–2), 109–125.

Satake K., 2015. Geological and historical evidence of irregular recurrent earthquakes in Japan, Phil. Trans. R. Soc. A 373, 20140375, http://dx.doi.org/10.1098/rsta.2014.0375.

Savage, J. C., Burford, R. O., 1973. Geodetic determination of relative plate motion in central California, J. Geophys. Res. 78, 832–845.

Scharer, K.M., Biasi, G.P. and Weldon, R.J., 2011. A reevaluation of the Pallett Creek earthquake chronology based on new AMS radiocarbon dates, San Andreas fault, California, Journal of Geophysical Research: Solid Earth 116(B12).

Schwartz, D. P., Coppersmith, K. J., 1984. Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones, J. Geophys. Res. 89, 5681–5698.

Shimazaki, K., Nakata, T., 1980. Time-predictable recurrence model of large earthquakes, Geophs. Res. Lett. 7, 279–282.

Sibson, R.H., 1992. Implications of fault-valve behavior for rupture nucleation and recurrence, Tectonophysics 211(1-4), 283-293.

Sieh, K., Stuiver, M. and Brillinger, D. A., 1989. More precise chronology of earthquakes produced by the San Andreas Fault in Southern California, J. Geophys. Res. 94, 603–623.

Sieh K., Natawidjaja, D.H., Meltzner, A.J., Shen, C.C., Cheng, H., Li, K.S., Suwargadi, B.W., Galetzka, J., Philibosian, B., Edwards, R.L., 2008. Earthquake supercycles inferred from sea-level changes recorded in the corals of West Sumatra, Science 322, 1674–167.

Smalley, R.F., Chatelain, J.-L., Turcotte, D.L., Prevot, R., 1987. A fractal approach to the clustering of earthquakes: applications to the seismicity of the New Hebrides, Bull. Seismol. Soc. Am. 77(4), 1368-1381.

Stein, S., Engeln, J.F., DeMets, C., Gordon, R.G., Woods, D., Lundgren, P., Argus, D., Stein, C., Wiens, D.A., 1986. The Nazca-South America convergence rate and the recurrence of the Great 1960 Chilean Earthquake, Geophysical Research Letters 13(8), 713-716.

Stein, S., Okal, E.A., 2007. Ultralong period seismic study of the December 2004 Indian Ocean earthquake and implications for regional tectonics and the subduction process, Bulletin of the Seismological Society of America 97(1A), S279-S295.

Stein, S., Liu, M., Calais, E., Li, Q., 2009. Mid-continent earthquakes as a complex system, Seismol. Res. Lett. 80(4), 551–553.

Stein, S., Wysession, M., 2009. An introduction to seismology, earthquakes, and earth structure, John Wiley & Sons.

Stein, S., Geller, R.J. and Liu, M., 2012. Why earthquake hazard maps often fail and what to do about it, Tectonophysics 562, 1-25.

Stein, S., Stein, J.L., 2013. Shallow versus deep uncertainties in natural hazard assessments, EOS 94, 4, 133-140.

Stein, S., Liu, M., Camelbeeck, T., Merino, M., Landgraf, A., Hintersberger, E., Kübler, S., 2017a. Challenges in assessing seismic hazard in intraplate Europe, Geol. Soc. London Special Publications 432(1), 13-28.

Stein, S., Salditch, L., Brooks, E., Spencer, B., Campbell, M., 2017b. Is the coast toast? Exploring Cascadia earthquake probabilities, GSA Today 27, 6-7.

Sykes, L.R., Nishenko, S.P., 1984. Probabilities of occurrence of large plate rupturing earthquakes for the San Andreas, San Jacinto, and Imperial faults, California, 1983–2003, J. Geophys. Res.: Solid Earth 89(B7), 5905-5927.

Thatcher, W., 1990. Order and diversity in the modes of circum-Pacific earthquake recurrence, J. Geophys. Res. 95, 2609–2624.

Tibshirani, R., Walther, G. and Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63(2), 411-423.

Turcotte, D.L., 1997. Fractals and Chaos in Geology and Geophysics, Cambridge University Press.

Tuttle, M.P., Al-Shukri, H., Mahdi, H., 2006. Very large earthquakes centered southwest of the New Madrid seismic zone 5,000–7,000 years ago, Seismol. Res. Lett. 77(6), 755–770.

Vere-Jones, D., 1970. Stochastic models for earthquake occurrence, Journal of the Royal Statistical Society Series B 32(1), 1-62.

Wallace, R. E., 1987. Grouping and migration of surface faulting and variation in slip rates on faults in the Great Basin province, Bull. Seismol. Soc. Am. 77, 868–877.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association 58(301), 236-234, doi: 10.1080/01621459.1963.10500845.

Ward, S.N., 1992. An application of synthetic seismicity in earthquake statistics: The Middle America Trench, Journal of Geophysical Research: Solid Earth 97(B5), 6675-6682.

Weldon, R., Scharer, K., Fumal, T., Biasi, G., 2004. Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work, GSA Today 14(9), 4-10, doi: 10.1130/1052-51732004)014<4:WATECW>2.0.CO;2.

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

Wechsler, N., Rockwell, T.K., Klinger, Y., Stepancikova, P., Kanari, M., Marco, S., Agnon, A., 2014. A paleoseismic record of earthquakes for the Dead Sea Transform fault between the first and seventh centuries C.E.: Nonperiodic behavior of a plate boundary fault, Bull. Seis. Soc. Am. 104(3), doi: 10.1785/0120130304.

WGCEP, Working Group on California Earthquake Probabilities, 2003. Earthquake Probabilities in the San Francisco Bay Region, U.S. Geologic Survey, Denver, USGS Open File Report 03-214.

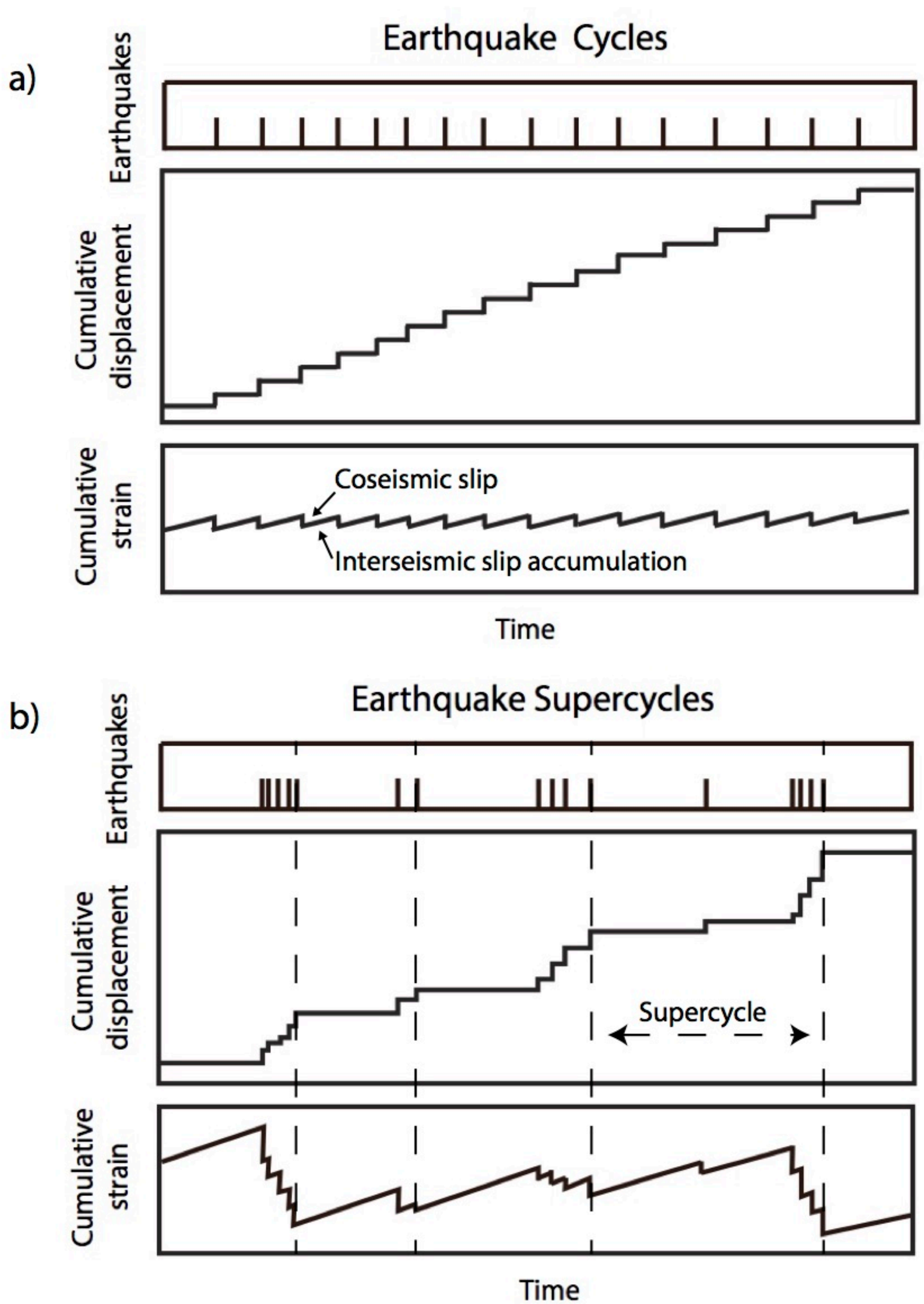

Figure 1: Schematic comparison of the histories of earthquake ocurrence, cumulative displacement, and cumulative strain for a fault without supercycles (a) and a fault with supercycles (b). Adapted from Wallace (1987) and Friedrich et al. (2003).

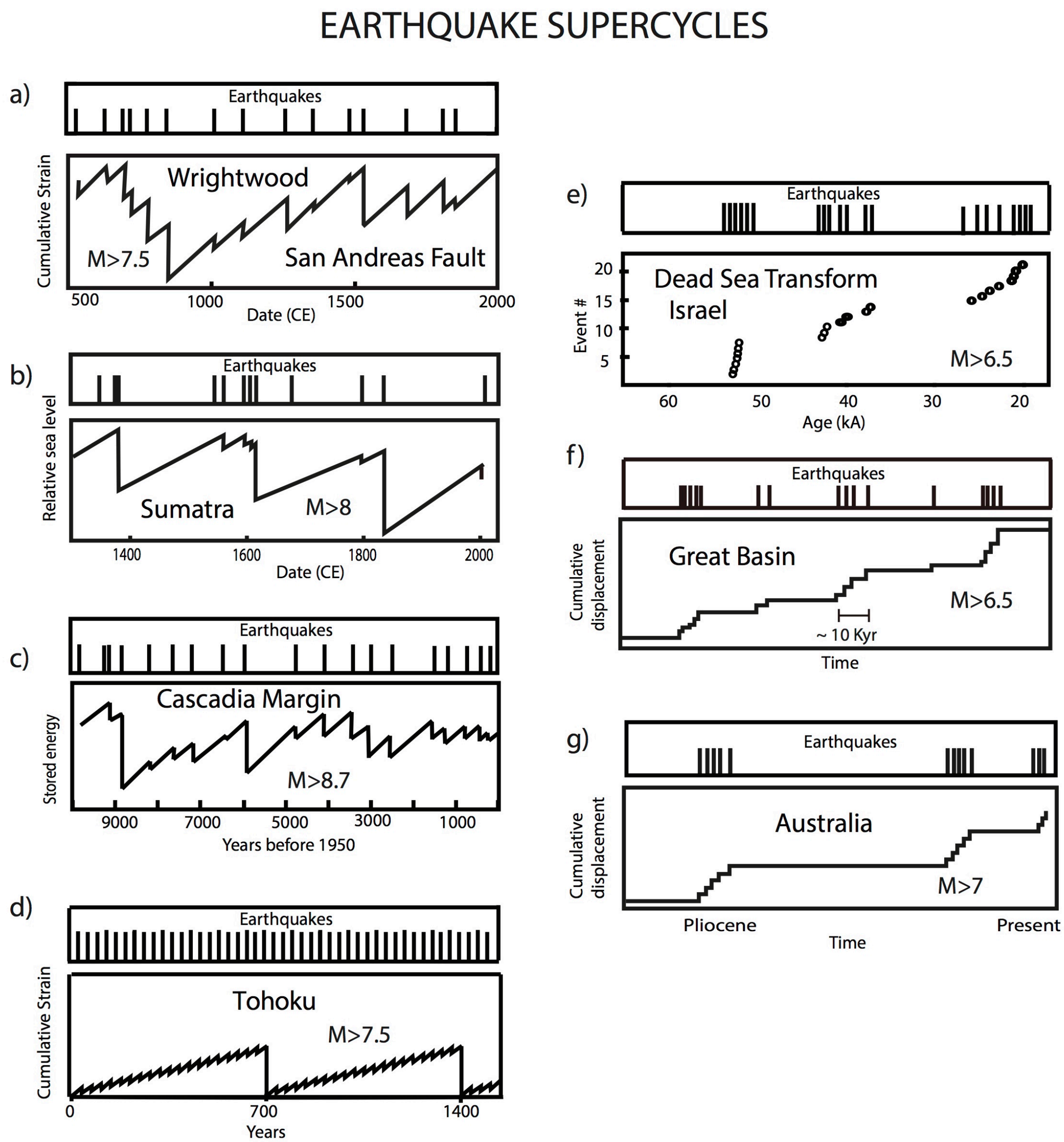

Figure 2: Examples of reported supercycles. a) Strain accumulation and release inferred from paleoseismic data across the San Andreas fault (Weldon et al., 2004). b) Supercycles on the Sumatra megathrust inferred from corals (Sieh et al., 2008). c) Long-term energy cycling inferred from turbidites on the Cascadia megathrust (Goldfinger et al., 2013). d) Schematic earthquake history for the Japan Trench off Tohoku (Satake, 2015). e) Earthquake history on the Dead Sea transform (Agnon, 2014). f) Schematic earthquake history for faults and groups of faults in the Western U.S.'s Great Basin (Wallace, 1987). g) Schematic earthquake history for faults in Australia (Clark et al., 2012).

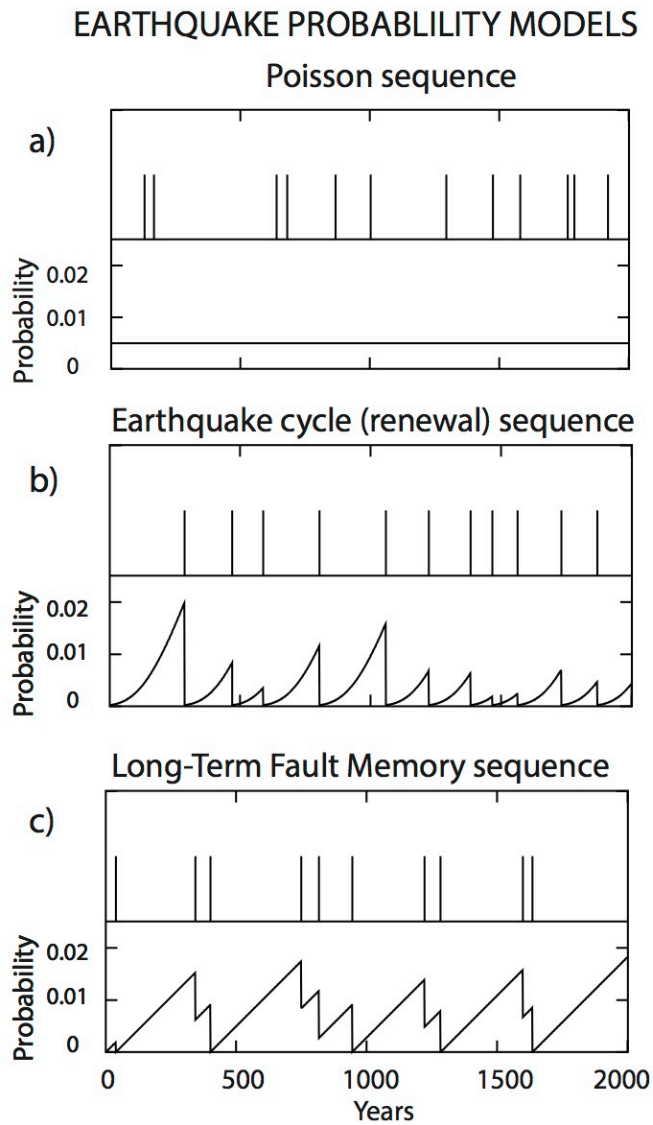

Figure 3: Comparison of earthquake recurrence models. a) Poisson process, in which the probability of a large earthquake is constant with time, so the fault has no memory and any clusters resulting from short intervals between events arise purely by chance. b) Earthquake cycle, in which the probability of a large earthquake increases with time until one occurs, at which point the probability drops to zero. The fault has only "short-term memory" because the probability of a large earthquake depends only on the time since the past one. c) Modified earthquake cycle in which after an earthquake the probability decreases, but not necessarily to zero. The fault has "long-term memory" because the probability depends on the earthquake history over previous cycles.

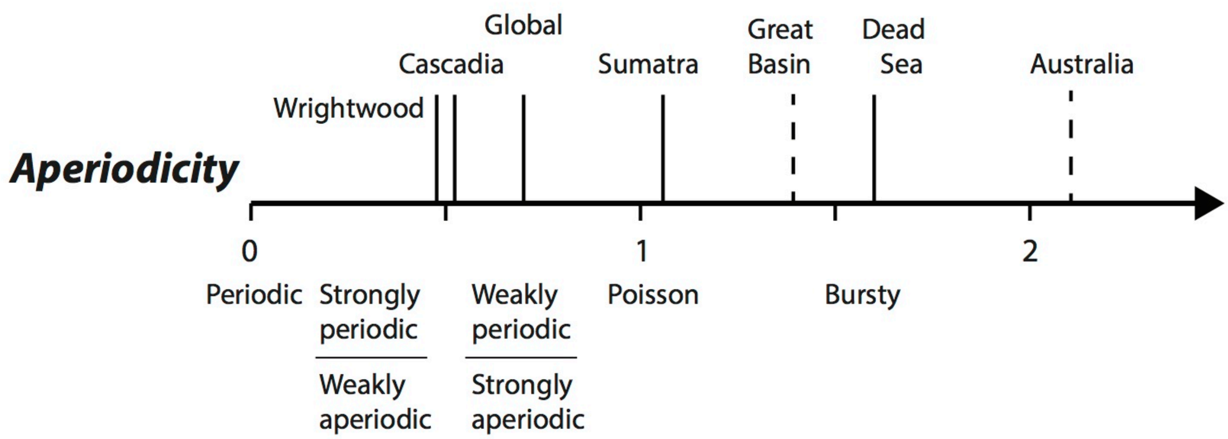

Figure 4: Illustration of characterizing earthquake sequences by their aperiodicity, which measures the extent that a sequence differs from perfectly periodic. Values are shown for examples in Figure 2. Solid bars show sequences with dates and dashed bars show schematic sequences with approximate aperiodicites. Also shown is the result from Goes' (1996) global compilation.

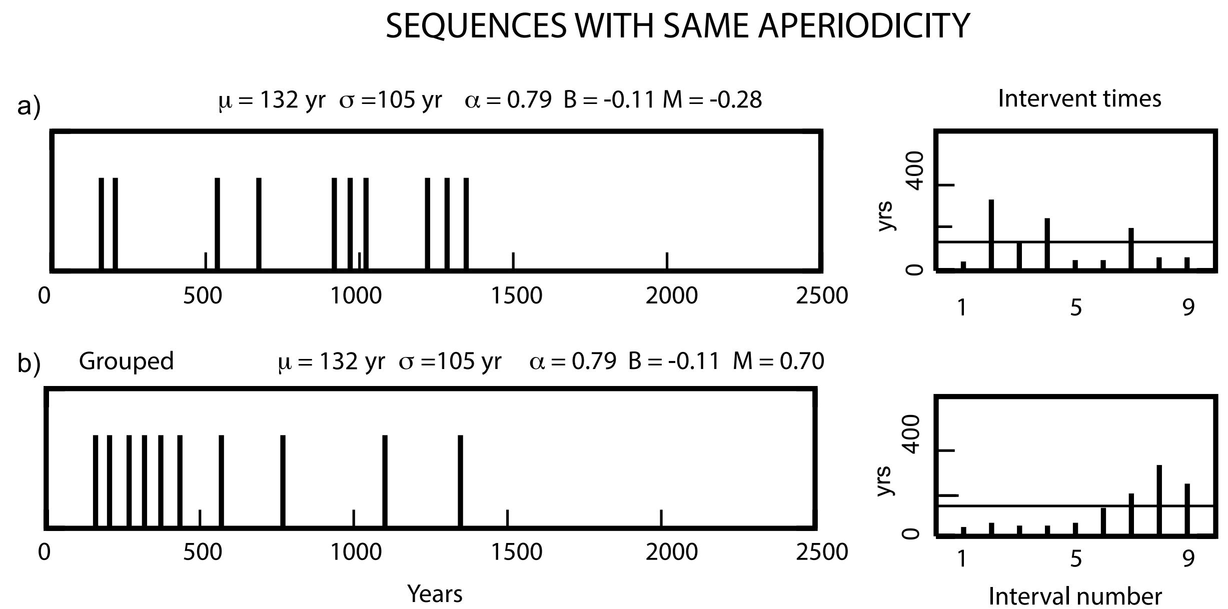

Figure 5: Comparison of two sequences with the same aperiodicity. a) Sequence with strong aperiodicity ($\alpha = 0.79$) showing clustering. (b) Same sequence with the short-interval events grouped together, which does not show clustering.

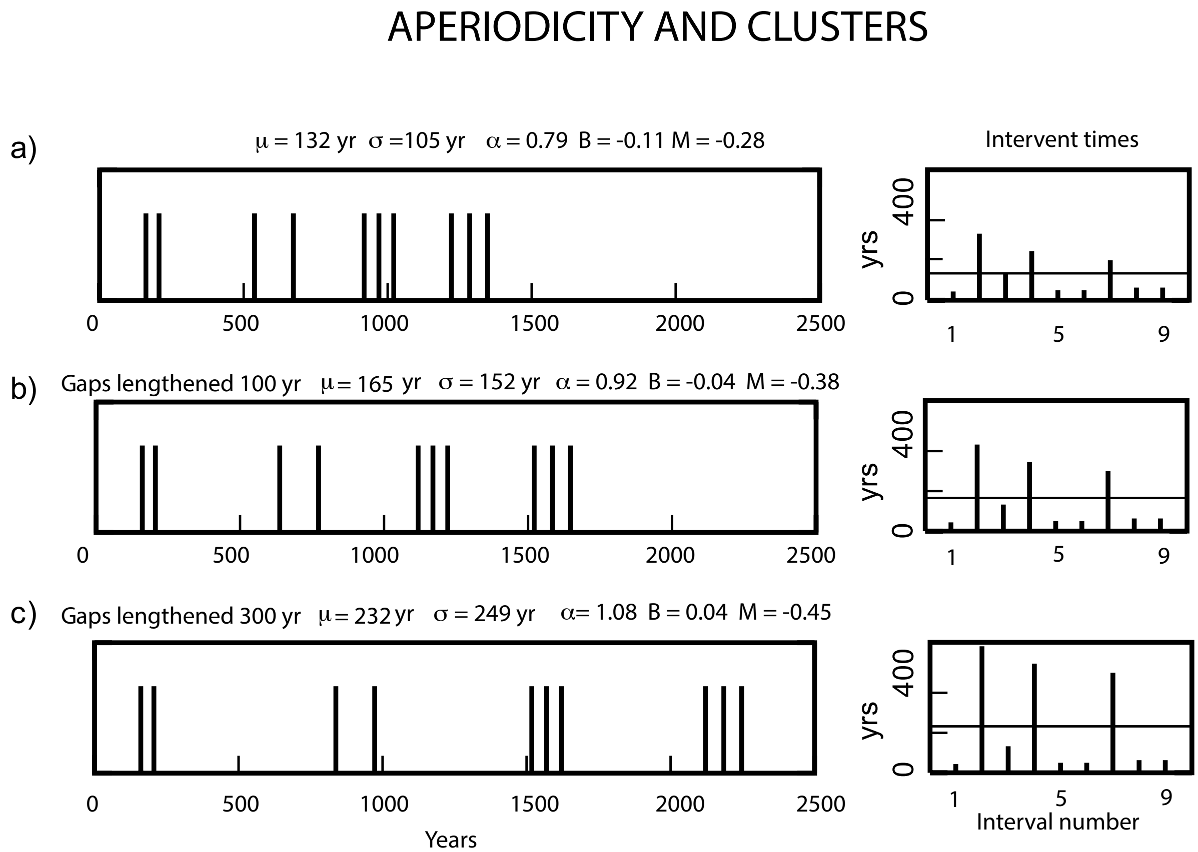

Figure 6: Illustration of the fact that "quasiperiodic" ($\alpha < 1$) sequences can be quite clustered. a) Initial sequence. b) Same sequence with major gaps lengthened by 100 years. c) Same sequence with major gaps lengthened by 300 years. Only c) has aperiodicity above the nominal burstiness criterion of $\alpha > 1$.

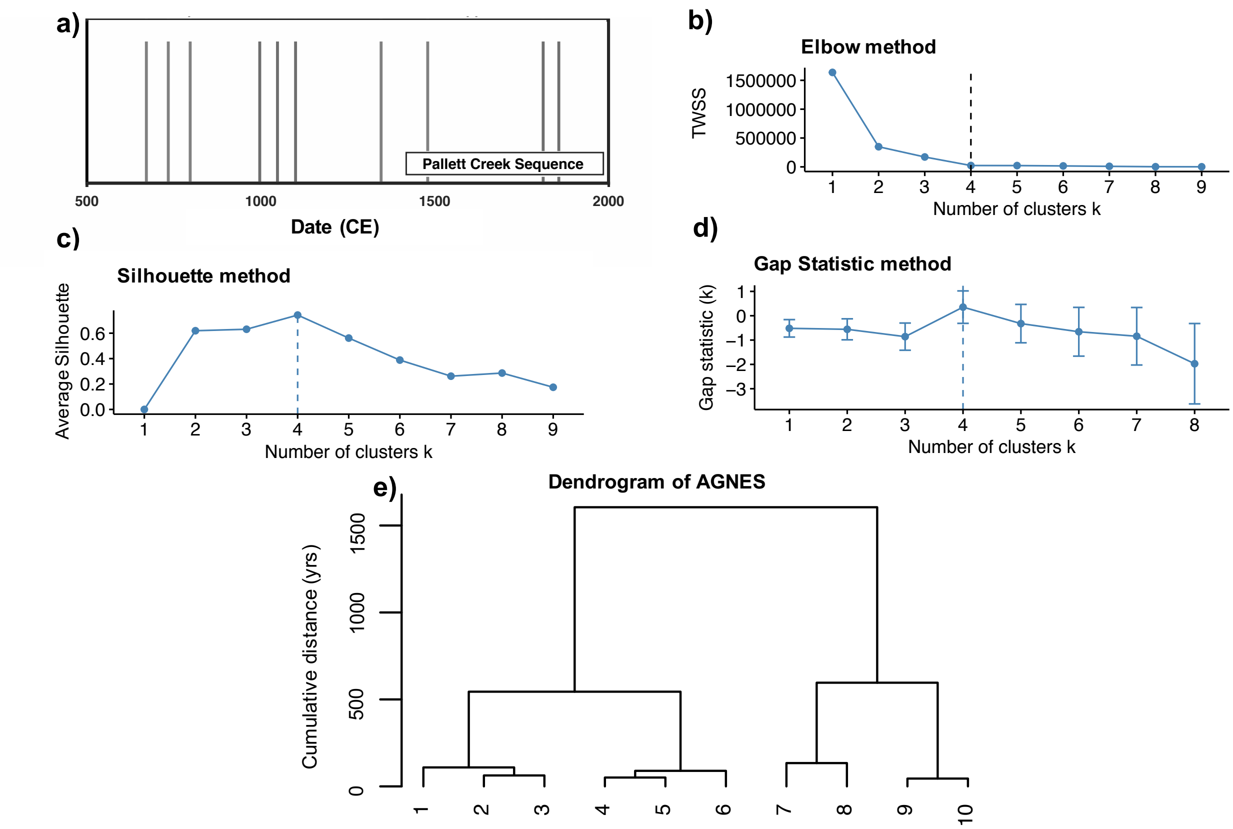

Figure 7: Results of different methods to determine the number of clusters in a) Pallett Creek record of Sieh et al., 1989, with event order corresponding to the figure in part e). b) Elbow method where number of clusters is the largest *k* before increasing *k* creates only minor improvements of TWSS. c) Silhouette method where maximum value indicates number of clusters. d) Gap statistic method where maximum value indicates number of clusters. e) Hierarchical clustering method using agglomerative nesting (AGNES) with Ward's method; vertical axis shows the cumulative length of time between cluster centers being merged at each step.

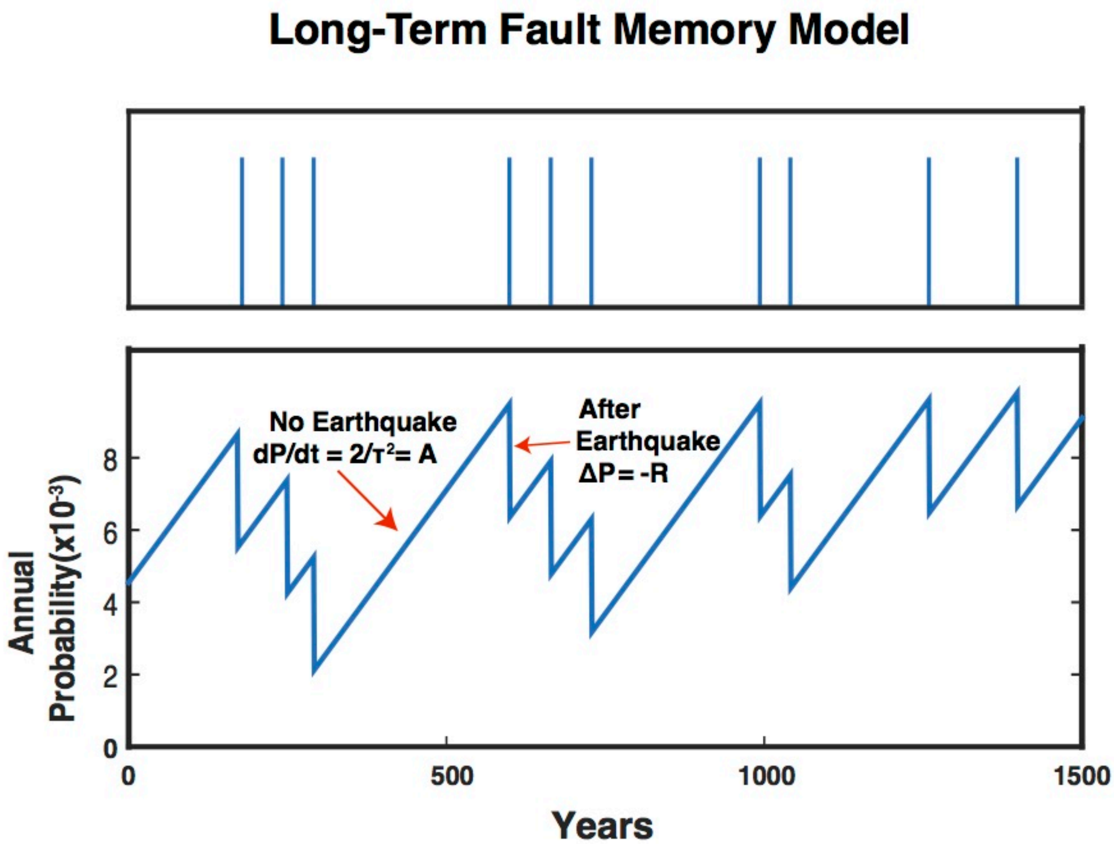

Figure 8: Long-Term Fault Memory model**.** **(**Top) Simulated earthquake history. (Bottom) Earthquake probability versus time.

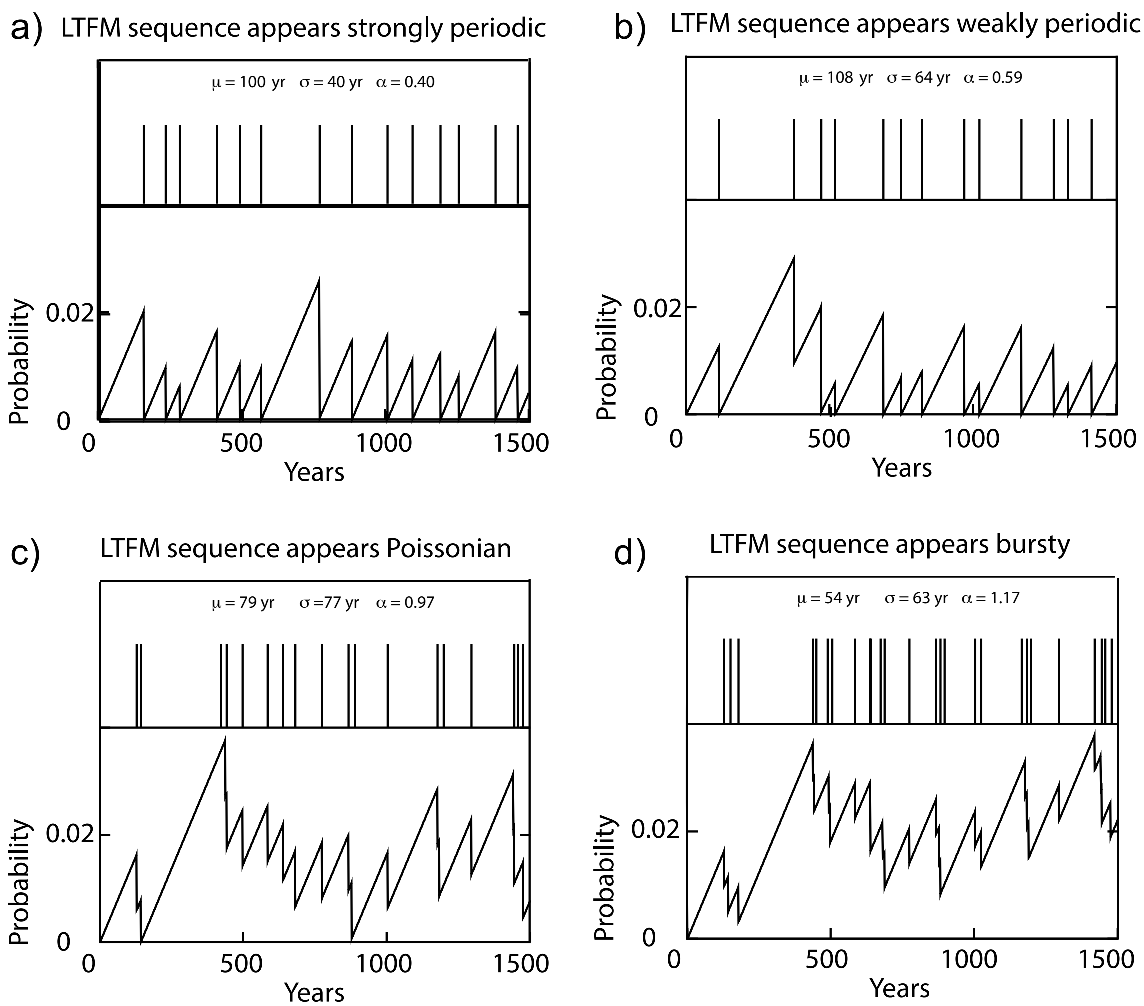

Figure 9: Sequences produced by the LTFM model can appear a) strongly periodic, b) weakly periodic, c) Poissonian, or d) bursty, depending on the model parameters. The four sequences shown have the same probability accumulation rate but different release parameters, so the aperiodicity increases as R decreases.

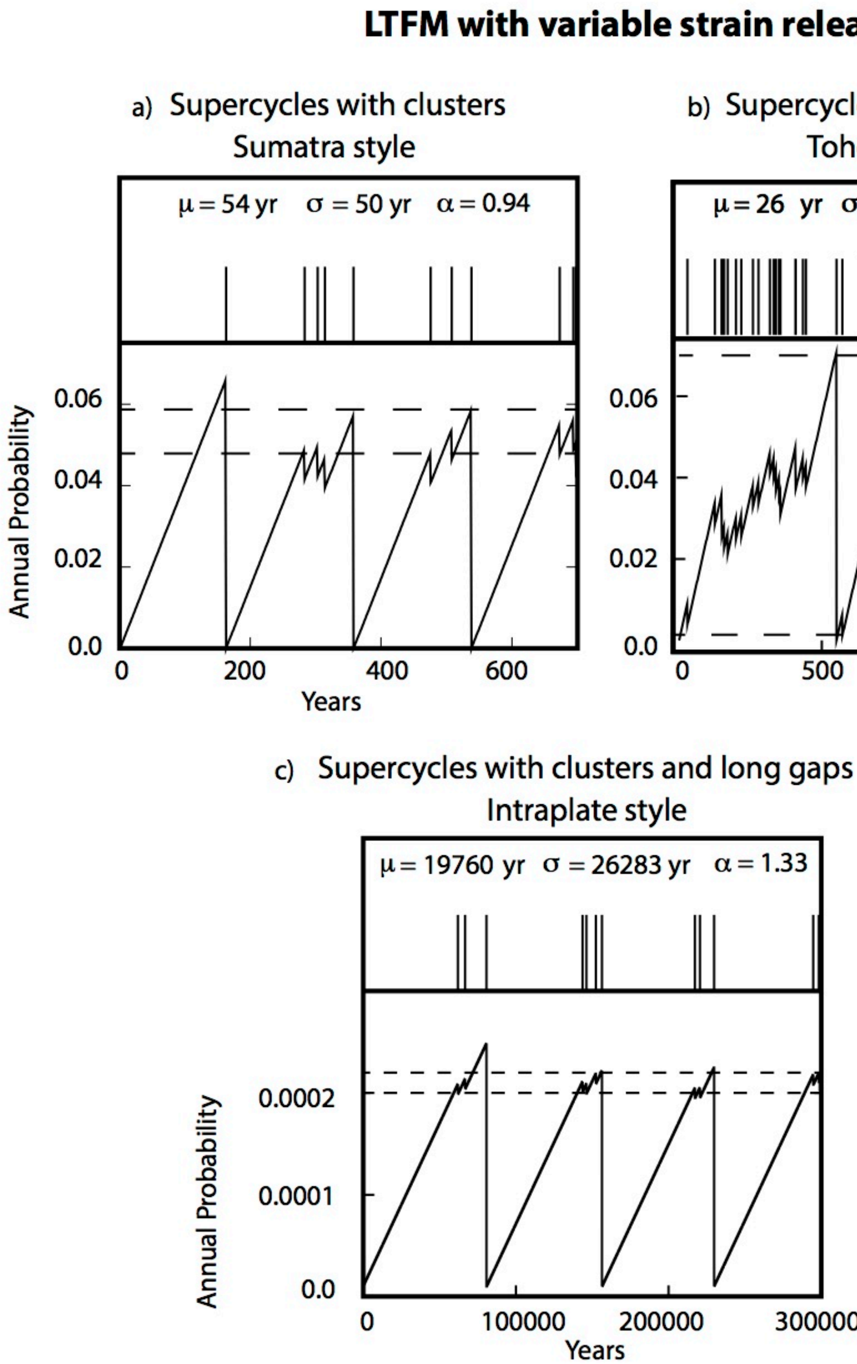

Figure 10: Using two probability thresholds (dashed lines) and probability drops to describe rare larger events and more frequent smaller events allows LTFM to simulate a wide range of observed supercycle behavior.

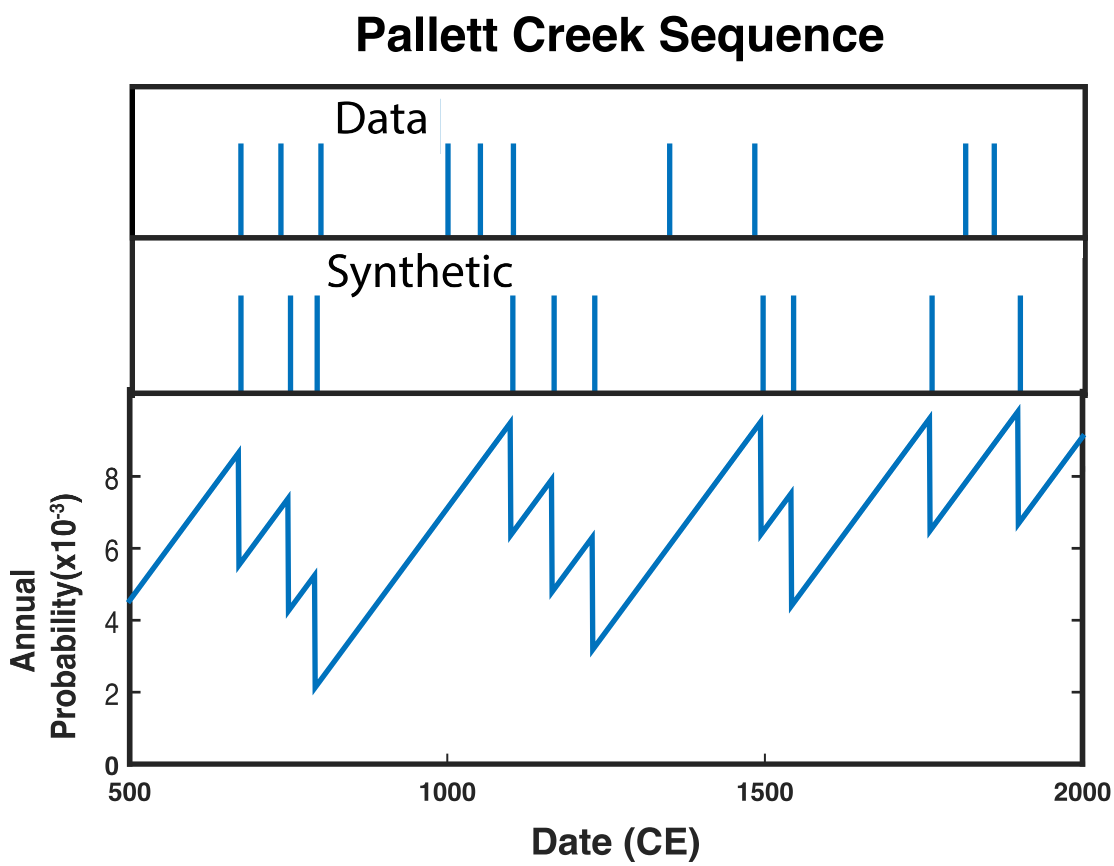

Figure 11: LTFM simulation for Pallett Creek, California. Top: Paleoseismic record (Sieh et al., 1989). Center and bottom: Simulation giving clustering similar to that observed. The event timing differs between the simulation and the observed record due to the model's stochastic nature.

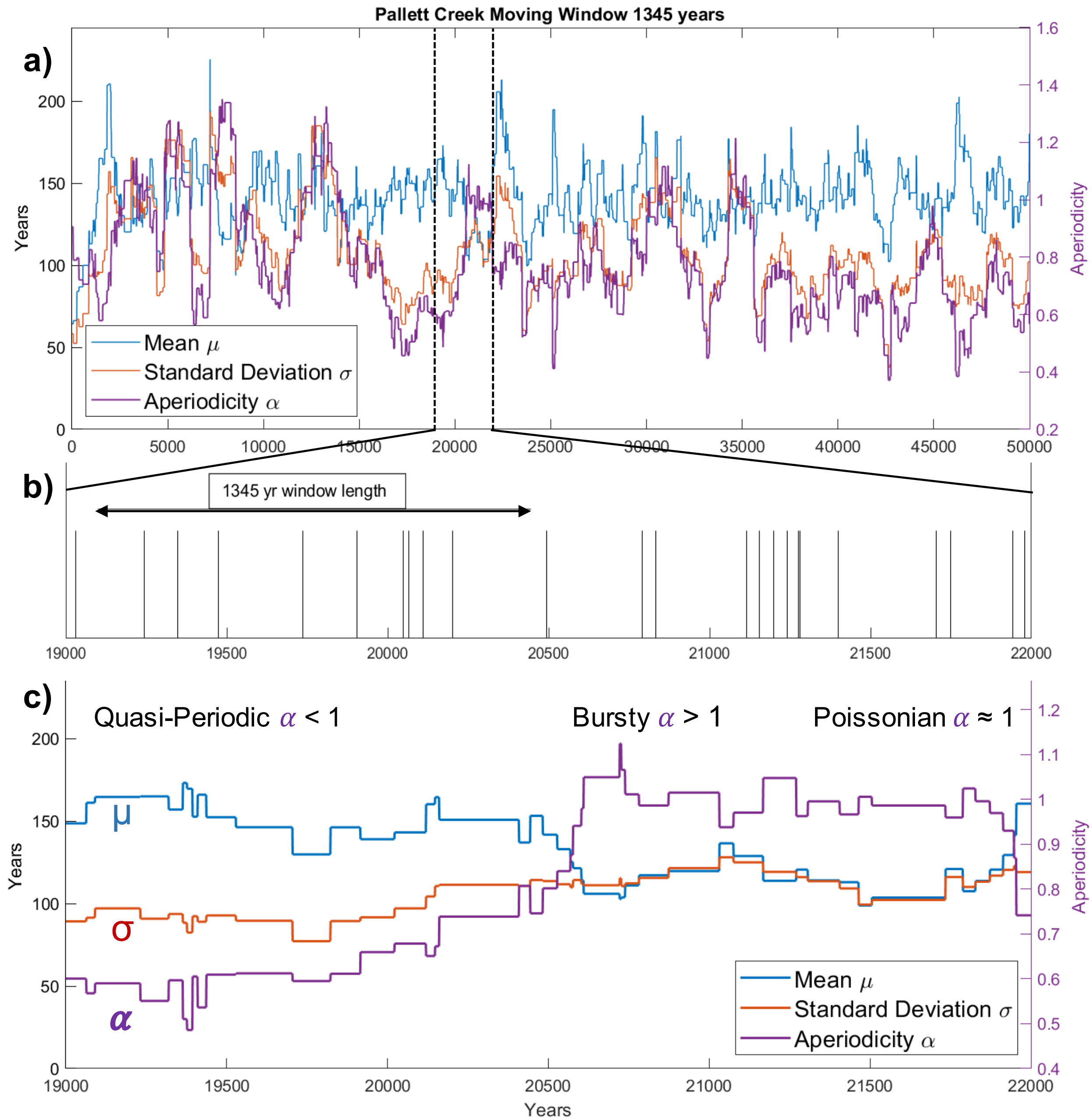

Figure 12: a) 50,000 year LTFM simulation using Pallett Creek parameters. The mean and standard deviation of recurrence times are averaged over a moving 1345-year window, corresponding to a paleoseismic record. b, c) 3,000 year section of simulation above between dashed lines in a). The aperiodicity shows that the simulated paleoseismic record sometimes appears strongly periodic ($\alpha < 0.5$) , while at other times it looks weakly periodic ($0.5 < \alpha < 1$), Poissonian ($\alpha \approx 1$), or bursty ($\alpha > 1$).

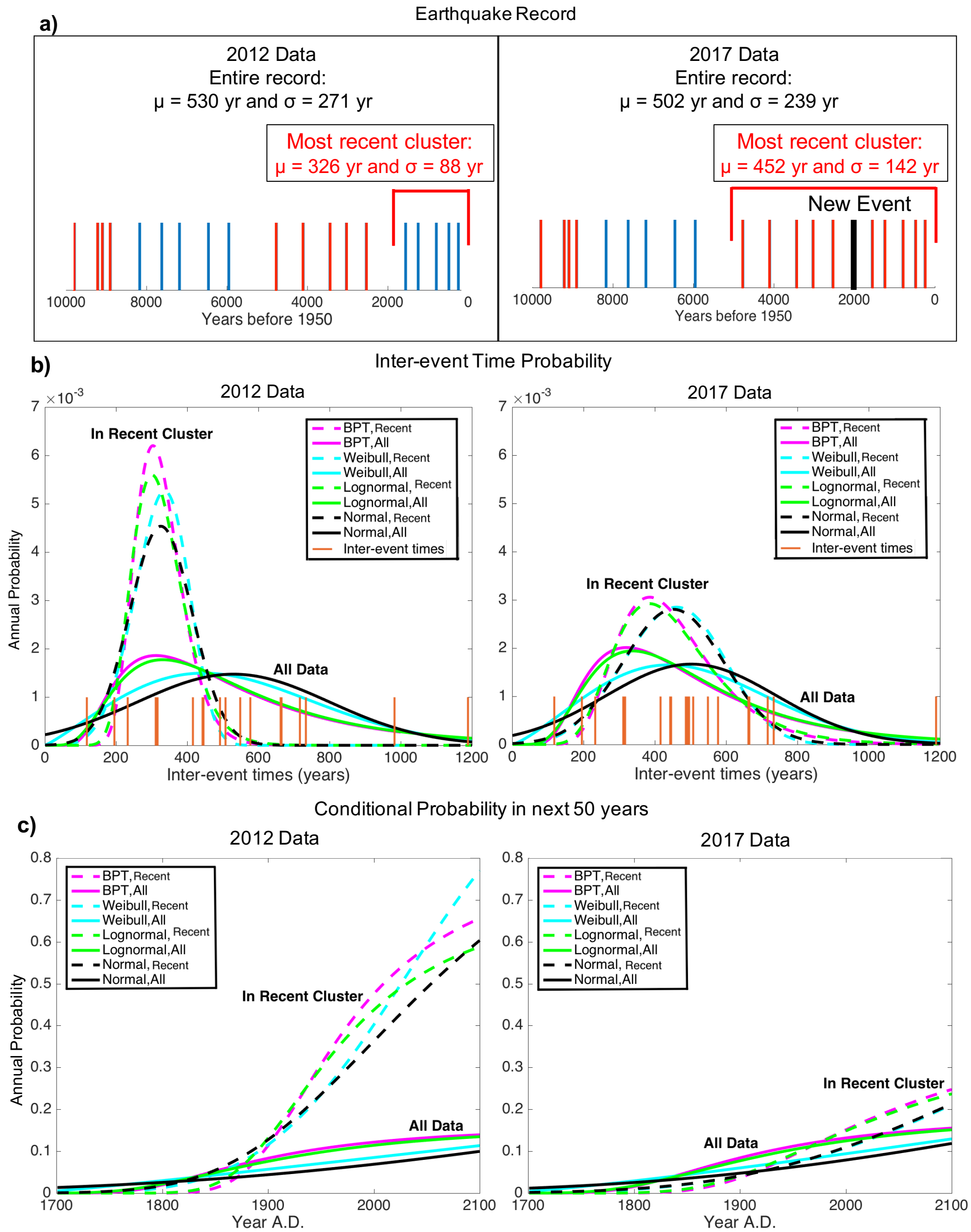

Figure 13: Illustration of earthquake probability issues for Cascadia due to a) differing paleoseismic records of Goldfinger et al., 2012 and Goldfinger et al., 2017, with its newly discovered event. Alternating red and blue events highlight the different clusters individual events are assigned to. b) Various probability density functions for inter-event times with parameters derived from the two chronologies in a). Orange sticks show the actual inter-event times in the corresponding records. Dashed lines use parameters of just the most recent cluster, corresponding to the assumption that the system is still in the recent cluster. Solid lines use the parameters of the entire record, corresponding to the assumption that the recent cluster has ended. c) Various conditional probabilities of an earthquake occurring in the next 50 years, using the same line designations in b). The largest difference in b) and c) arises from the recent cluster assumption, not in the specific density function assumed.