

NORTHWESTERN UNIVERSITY

Insights into Earthquake Processes and Hazards Using Statistical  
Methods

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Earth and Planetary Sciences

By

James Scott Neely

EVANSTON, ILLINOIS

December 2022

© Copyright by James Scott Neely 2022

All Rights Reserved

## ABSTRACT

Insights into Earthquake Processes and Hazards Using Statistical Methods

James Scott Neely

To better prepare for earthquakes, we need to know how large they will be, how strong the shaking will be, and how often they will occur. To answer these questions, seismologists look to past earthquakes to better understand future hazards. Earthquakes, however, are complex physical phenomena that occur on timescales much larger than our instrumental records so our past observations may only be providing a partial picture of earthquake behavior. Therefore it is critical to know the limitations of past observations, assess the uncertainty inherent in our earthquake models, and ensure that our models reflect our understanding of the processes that generate earthquakes. In this thesis, I present novel implementations of various statistical methods to help answer these fundamental earthquake questions.

Using probabilistic simulations, I show that the largest earthquakes in the eastern North America record may not be the largest possible earthquakes that can occur. In this region, the records are simply too short relative to the frequency of large earthquakes to exclude the possibility of larger earthquakes. However, using a similar probability

approach, I demonstrate that observed global variations in earthquake magnitude by fault geometry likely do reflect real differences and are not an artifact of short catalog lengths. I show that continental normal fault earthquakes do have smaller maximum magnitudes than other fault geometries and propose that the smaller maximum magnitudes reflect the weakness of continental lithosphere in extension.

In assessing potential earthquake hazards, we need to know not only how big earthquakes might be but also how strong the resulting shaking will be. Earthquake stress drop—the change in stress along a fault due to an earthquake—is a commonly estimated earthquake parameter which is thought to control the amplitude of high-frequency shaking that damages buildings and structures. I show that two of the most commonly used methods can produce drastically different estimates for the same earthquakes. There is significant, unaccounted for uncertainty in these estimates. As a result stress drop trends that appear using one method are unobservable using the other method.

Lastly, seismic hazard analysis requires knowing how often large earthquakes will occur. Current earthquake recurrence models make many simplifying assumptions that ignore the complexities of the processes that drive earthquakes. Here I present the analytical equations for the Long-Term Fault Memory based on Salditch et al.'s (2020) numerical design which produces earthquake probability estimates based on the specific sequence of past earthquakes. I derive the equations for two different versions of this model and apply them to real earthquake records. My analysis shows the specific earthquake sequence can significantly raise the estimated likelihood of an earthquake and may provide more accurate assessments of earthquake hazard. Together, these different analysis show how

statistical models can be applied in a variety of ways to provide insight into fundamental earthquake questions.

## Acknowledgements

I first want to thank my advisor Seth Stein and my committee Bruce Spencer and Emile Okal. Seth, thank you for giving me the freedom (and trusting me) to pursue the questions and problems that interested me. Bruce, thank you for providing me with statistical insights that I never would have thought of. Emile, thank you for sharing your wealth of knowledge on historical earthquakes and seismology.

Thank you to Kevin Furlong who took a chance on a student who knew next to nothing about geology. Thank you to Yihe Huang who graciously supported my research while I was deciding to pursue a PhD. And thank you to all the graduate students along the way who helped me become a better scientist.

Most importantly I want to thank my family. Thank you parents for always supporting my pursuits. And my wife Alex, thank you for supporting me on this long graduate school journey. And lastly, Rowan, thank you for being the best part of my day.

## Contents

ABSTRACT	3
Acknowledgements	6
List of Tables	11
List of Figures	12
Chapter 1. Thesis Overview	16
1.1. Introduction	16
1.2. Chapter 2: Have We Seen the Largest Earthquakes in Eastern North America?	16
1.3. Chapter 3: Large Uncertainties in Earthquake Stress Drop Estimates and Their Tectonic Consequences	17
1.4. Chapter 4: Why Do Continental Normal Fault Earthquakes Have Smaller Maximum Magnitudes?	18
1.5. Chapter 5: A More Realistic Earthquake Probability Model Using Long-Term Fault Memory	19
1.6. Chapter 6: The Generalized Long-Term Fault Memory Model and Applications to Paleoseismic Records	20
Chapter 2. Have We Seen the Largest Earthquakes in Eastern North America?	21

2.1. Summary	21
2.2. Introduction	22
2.3. Methods	27
2.4. Results and Analysis	32
2.5. Discussion	40
2.6. Conclusion	45
Chapter 3. Large Uncertainties in Earthquake Stress Drop Estimates and Their Tectonic Consequences	47
3.1. Summary	47
3.2. Introduction	48
3.3. Estimating Earthquake Stress Drop	48
3.4. Previously Reported Differences in Earthquake Stress Drop	51
3.5. Large Uncertainty in Stress Drop Estimates	53
3.6. Possible Sources of Earthquake Stress Drop Uncertainty	56
3.7. Observed Stress Drop Trends Vary Depending on Analysis	62
3.8. Discussion	67
Chapter 4. Why Do Continental Normal Fault Earthquakes Have Smaller Maximum Magnitudes?	71
4.1. Summary	71
4.2. Introduction	72
4.3. Data Set and Earthquake Classifications	74
4.4. Global Distribution of Large Normal Fault Earthquakes	76

4.5.	Variations in Earthquake Magnitude Distribution with Fault Geometry	77
4.6.	Smaller Normal Fault $M_{max}$ Not Due to Classification	80
4.7.	Smaller Normal Fault $M_{max}$ Not Due to catalog Length	82
4.8.	Reconciling GCMT $M_{max}$ with Early Instrumental Earthquakes	84
4.9.	Fault Dimensions Not a Limitation on Normal Fault Earthquake Magnitude	87
4.10.	Lithosphere Yield Stress Controls $M_{max}$	93
4.11.	Conclusion	100
Chapter 5.	A More Realistic Earthquake Probability Model Using Long-Term Fault Memory	101
5.1.	Summary	101
5.2.	Introduction	102
5.3.	Limitations of Current Earthquake Probability Methods	103
5.4.	Calculating Earthquake Probabilities with Long-Term Fault Memory	108
5.5.	Inter-Event Time Order Matters for LTFM	114
5.6.	Earthquake Forecast Immediately after the 1812 Pallett Creek Earthquake	118
5.7.	Assumptions About Earthquake Probability and Accumulated strain	120
5.8.	Conclusion	123
Chapter 6.	The Generalized Long-Term Fault Memory Model and Applications to Paleoseismic Records	125
6.1.	Introduction	125
6.2.	The Mechanics of Probability Models	128
6.3.	The Generalized LTFM	130

	10
6.4. Fitting Parameters and Choosing the Probability Function $f$	136
6.5. GLTFM Applied to Real Paleoseismic Records	138
6.6. Identifying the Best Model	142
6.7. Conclusion	148
Appendix A. Derivations for Markov LTFM Formulation	149
A.1. Fitting the LTFM to Data	149
A.2. The Model	151
A.3. Transition Probability and Stationary Distribution	156
A.4. Conditional Probability of an Earthquake Occurring in $T$ Epochs	158
A.5. Unconditional Probability of an Earthquake Occurring in $T$ Epochs	159
A.6. Conditional Probabilities Over States Given Recent Earthquake	160
A.7. Probability that the Next Earthquake Occurs $T$ or More Epochs After the Last Earthquake	162
A.8. Conditional Distribution over States Given Date of Last Earthquake	163
A.9. Conditional Probabilities Given Dates of Past 2 Earthquakes	165
A.10. Conditional Probabilities Based on the Past $m$ Earthquakes	169
A.11. Waiting Time Until Next Earthquake	173
A.12. Waiting Time Between Renewals of System	176
Appendix. References	179

## List of Tables

3.1	Wilcoxon rank-sum test (two-tailed) for statistical significance of tectonic environment stress drop patterns	66
4.1	Numbers of earthquakes $M_w \geq 5.8$ in the GCMT database classified by tectonic environment and fault geometry.	75
4.2	Lithology parameters used in yield stress envelope calculations. Values from Burov (2011).	96
4.3	Earth structure and geotherm calculation parameters.	96
5.1	Dates of the most recent 10 earthquakes along the Pallett Creek section of the San Andreas fault	103
A.1	LTFM notation glossary	152

## List of Figures

2.1	Seismicity of the eastern North America continental margin	25
2.2	Frequency-magnitude relationships calculated for southern North America and northern North America datasets.	29
2.3	Recurrence time cumulative distribution function, magnitude cumulative distribution function, and sample synthetic catalog	32
2.4	Histograms of $M_{max}^a$ from simulated catalogs	35
2.5	$b$ value results for the four sets of 10,000 sub-catalogs for the southern North America margin	36
2.6	$b$ value results for the four sets of 10,000 sub-catalogs for the northern North America margin	37
2.7	Distributions of LSQ estimates of $b$ and $M_{max}^a$	38
2.8	Distributions of MLE estimates of $b$ and $M_{max}^a$	39
2.9	Distributions of estimated $M_{max}^p$ recurrence	40
2.10	Percent of simulations yielding $M_{max}^a = 7.6$ within the parameter space for 300 and 750 year-long sub-catalogs	41
2.11	Mean recurrence times for earthquakes greater than a given magnitude	42

		13
2.12	Percentage of sub-catalogs with $M_{max}^a$ observed based on 10,000, 150 year-long sub-catalogs	43
3.1	Earthquake stress drop estimates depend on estimates of characteristic rupture dimension from either rupture duration or corner frequency	50
3.2	Correlation plots of AS09 frequency-domain and time-domain estimates of stress drop and rupture duration for individual earthquakes	55
3.3	Distributions of stress drop estimates in $\log_{10}$ space	57
3.4	Correlation plots of AS09 frequency-domain and time-domain estimates of stress drops for earthquakes grouped by magnitude	59
3.5	Correlation plots of frequency-domain and time-domain estimates of stress drop and rupture duration	61
3.6	Stress drop estimates for the 894 earthquakes in this study	62
3.7	Stress drop distributions for different tectonic environments	65
3.8	Corner frequency and inverse rupture duration as a function of seismic moment	68
4.1	Distribution of $M_w \geq 6.5$ normal fault earthquakes in the GCMT catalog	77
4.2	Depth versus magnitude plot for normal fault earthquakes	78
4.3	Tapered Gutenberg-Richter distributions for shallow continental, oceanic, and convergent earthquakes	79

		14
4.4	Plot of magnitude versus P-axis plunge for shallow continental, oceanic, and convergent earthquakes	81
4.5	Probability that for 135 earthquakes, $M_w$ 7.1 is the largest observed earthquake	83
4.6	Probability of observing at least 1, 2, or 3 shallow continental normal fault earthquakes $\geq M_w$ 7.3	86
4.7	Histogram of active continental normal fault lengths from the GEM-Active Fault database	89
4.8	Mapped surface rupture length versus magnitude for large historical earthquakes	90
4.9	Yield-stress envelope (YSE) for oceanic and continental lithosphere in compression and extension	95
4.10	P-axis plunge versus moment magnitude for oceanic earthquakes with outer rise earthquakes (those earthquakes within 100 km of the trench) removed	98
4.11	Schematic plot of stress within homogenous material for extension and bending	99
5.1	Estimating earthquake probabilities	104
5.2	Pallett Creek earthquake probabilities	107
5.3	Long-term fault memory (LTFM) as a hidden Markov process	108
5.4	LTFM simulations with different input parameters	112

5.5	LTFM applied to sample records	115
5.6	Probability estimates immediately after 1812 Pallett Creek earthquake	119
5.7	Probability estimates immediately after 1812 Pallett Creek earthquake	121
5.8	Comparisons of Weibull and LTFM	123
6.1	Application of memoryless and short-term memory models	131
6.2	GLTFM calculation schematic	133
6.3	GLTFM applied to sample paleoseismic record	135
6.4	Impact of probability accumulation rate on earthquake probability	139
6.5	GLTFM applied to the Pallett Creek paleoseismic record	143
6.6	GLTFM applied to the Hayward fault paleoseismic record	144
6.7	GLTFM applied to the Nankai Trough (section A-B) subduction zone paleoseismic record	145
6.8	GLTFM applied to the Cascadia subduction zone paleoseismic record	146
6.9	GLTFM applied to the Alpine fault paleoseismic record	147

## CHAPTER 1

### **Thesis Overview**

#### **1.1. Introduction**

My thesis has been guided by three fundamental earthquake questions: (1) How large will earthquakes be? (2) How often will large earthquakes occur? and (3) How strong will the shaking be? The answers to these questions have important implications for both our understanding of the physical processes that drive earthquakes and the hazards that they pose. In this thesis I apply multiple statistical methods in novel ways to address unanswered seismological questions. I move beyond standard statistical seismology that focuses primarily on descriptive statistics to quantify past earthquake observations and instead focus on what statistical analyses of past earthquakes can and cannot tell us about future earthquake behavior.

#### **1.2. Chapter 2: Have We Seen the Largest Earthquakes in Eastern North America?**

Eastern North America is not a particularly seismically active region, however, large earthquakes have been observed in the past. Historical records suggest that over the last 300 years the largest earthquakes along the southern margin reach magnitude 7. Along the northern margin, the largest earthquake in the past 100 years was in the mid magnitude 7 range. However, are these historical records long enough to constrain the maximum possible earthquake magnitudes for the region or are larger earthquakes possible and simply

haven't been observed yet because they are so infrequent? In this chapter I use recent, instrumentally recorded earthquakes to simulate thousands of synthetic catalogs that are the same length as our historical records. These simulations reveal that our historical seismicity records along eastern North America are too short to reliably contain larger and more infrequent earthquakes. These results do not mean that larger earthquakes can occur along the coastal margin of eastern North America. Rather, the current historical records are too short to eliminate the possibility of such larger earthquakes.

This chapter is published as Neely, J.S., Stein, S., Merino, M., Adams, J. (2018), Have we seen the largest earthquakes in eastern North America? *Physics of the Earth and Planetary Interiors* 284, 17-27. doi:10.1016/j.pepi.2018.09.005

### **1.3. Chapter 3: Large Uncertainties in Earthquake Stress Drop Estimates and Their Tectonic Consequences**

In this chapter, I quantify the uncertainties of the earthquake stress drop estimates and the impact of the uncertainty on tectonic analyses. Earthquake stress drop, the stress change on the fault due to an earthquake, is a commonly calculated parameter that provides a window into rupture dynamics, including earthquake fracture energy, ground motions, and rupture duration. Earthquake stress drop is also thought to control the high frequency shaking that damages structures. Seismologists use stress drop estimates of past, more frequent small magnitude earthquakes to estimate the stress drops, and potential shaking, of a region's future larger earthquakes. However, stress drop estimates can vary by orders of magnitude, and different studies often produce conflicting results

suggesting large stress drop uncertainties. In this chapter, I quantify stress drop estimate uncertainty by comparing two independent stress drop estimates using different but theoretically equivalent methods for the same set of earthquakes. I find that there is essentially no correlation between the two different estimates. The independent stress drop estimates for the same earthquake can vary by as much as two to three orders of magnitude, and tectonic stress drop trends that are apparent using one method are absent using the other. These results suggest that stress drop estimate uncertainties are quite large and many reported stress drop trends may just be a product of uncertainty and not underlying tectonic differences.

This chapter is published as Neely, J.S., Stein, S., Spencer, B.D. (2020), Large uncertainties in earthquake stress-drop estimates and their tectonic consequences, *Seismological Research Letters* 91(4), 2320-2329. doi:10.1785/0220200004

#### **1.4. Chapter 4: Why Do Continental Normal Fault Earthquakes Have Smaller Maximum Magnitudes?**

In this chapter, I examine the assumption that continental normal fault earthquakes have smaller maximum magnitudes than other fault geometries and propose a possible explanation for why. The hazard posed by large normal faults, such as the Wasatch Fault in Utah, depends on whether they can rupture their full length in one very large earthquake or if earthquakes are confined to small segments. I investigate whether these observations reflect real differences between fault geometries or are simply an artifact of the relatively short earthquake records. I apply statistical methods to show that the smaller maximum magnitude trend is likely real, and if larger normal fault earthquakes were possible, we

would have likely observed them. I then analyze possible physical mechanisms for the observed differences. My analysis suggests that the length and structure of normal faults is not what limits the size of normal fault earthquakes. Rather, the relative weakness of lithosphere strength in extension is likely the primary culprit for the smaller normal fault earthquakes.

This chapter is published as Neely, J.S., Stein S. (2021), Why do continental normal fault earthquakes have smaller maximum magnitudes? *Tectonophysics* 809. doi:10.1016/j.tecto.2021.228854

### **1.5. Chapter 5: A More Realistic Earthquake Probability Model Using Long-Term Fault Memory**

Paleoseismic records of earthquake recurrence indicate a complex pattern of strain accumulation and release with earthquake clusters followed by long gaps until the next earthquake. These records suggest that faults contain long-term memory where the time until the next earthquake is influenced by the specific sequence of preceding earthquakes. These records suggest that earthquakes may only release some of the accumulated strain along a fault leaving residual strain which impacts the timing for the next earthquake. However, the most commonly used earthquake probability models ignore this fundamental aspect of the earthquake process. Salditch et al. (2020) recently developed a new numerical model called the Long-Term Fault Memory (LTFM) model that is specifically designed to replicate the complex temporal patterns observed in paleoseismic records. In this chapter, I reformulate the LTFM as a hidden Markov model. I derive the necessary equations (see Appendix) to analytically calculate earthquake probabilities of interest and

demonstrate the advantages of this model compared to existing ones. I apply the Markov LTFM to the paleoseismic record for the southern San Andreas fault and show how it is the only earthquake probability model to accurately forecast the occurrence of the 1857 Fort Tejon earthquake so soon after the preceding earthquake in 1812.

This chapter is in revision as Neely, J.S., Salditch, L., Spencer, B.D., Stein S., A more realistic model for the probability of large earthquakes. *Bulletin of the Seismological Society of America*. In Revision.

## **1.6. Chapter 6: The Generalized Long-Term Fault Memory Model and Applications to Paleoseismic Records**

In this chapter, I present the Generalized Long-Term Fault Memory Model (GLTFM), an updated version of the LTFM that is more versatile and easier to implement. The GLTFM allows for the simple calculation of earthquake probabilities based on the specific sequence of observed earthquakes. The GLTFM identifies which earthquakes in a sequence likely left residual strain along the fault and it produces estimates of earthquake probability based on this residual strain. I apply the GLTFM to paleoseismic sequences and compare its probability estimates to other commonly used models. The GLTFM indicates a significantly higher probability of an earthquake (compared to existing models) along the Hayward fault in the San Francisco Bay and the southern San Andreas fault due to residual strain after the most recent earthquakes along those respective faults. The GLTFM is a powerful new tool in seismic hazard analysis that will yield more accurate assessments of earthquake likelihood.

## CHAPTER 2

## Have We Seen the Largest Earthquakes in Eastern North America?

### 2.1. Summary

The assumed magnitude of the largest future earthquakes,  $M_{max}$ , is crucial in assessing seismic hazard, especially for critical facilities like nuclear power plants. Estimates are made using various methods and often prove too low, as for the 2011 Tohoku, Japan, earthquake. Estimating  $M_{max}$  is particularly challenging within tectonic plates, where large earthquakes are infrequent, vary in location and time, and often occur on previously unrecognized faults. For example, it is unclear whether the short historical record includes the largest possible earthquakes along the eastern continental margin of North America. We explore this issue by generating synthetic earthquake histories and sampling them over a few hundred years. Due to the short histories, the maximum magnitudes appearing most often in a sub-catalog,  $M_{max}^a$ , are often smaller than the maximum magnitude in the parent catalog,  $M_{max}^p$ , that can occur. Future earthquakes along the continental margin may thus be significantly larger than those observed to date. More generally, these simulations demonstrate that the largest earthquake in a catalog likely reflects a combination of catalog length, a region's earthquake productivity, and relative proportion of small to large events. For regions with low seismicity, small variations in  $b$  value, the ratio of large to small events, due to sampling has a significant impact on the expected recurrence

times of large magnitude earthquakes. Although the precise likelihood of observing  $M_{max}^p$  depends on the distribution of recurrence times, a catalog shorter than an earthquake's mean recurrence time will likely not contain an event of that size. As a result,  $M_{max}$  cannot always be reliably estimated from earthquake catalogs.

## 2.2. Introduction

The 2011 Virginia earthquake that shook much of the northeastern U.S. showed that earthquakes large enough to cause significant damage do occur in eastern North America (Wolin et al., 2012) (Figure 2.1). Assessing the hazard of such earthquakes poses major unresolved issues. Hazard maps, giving the maximum shaking expected in an area with a certain probability in some time period (Cornell, 1968), require assuming where and how often large earthquakes will occur and how large they will be. However, the recent Tohoku, Sumatra, and Wenchuan earthquakes illustrate that earthquakes much larger than previously expected occur in many places (Stein and Okal, 2007; Geller, 2011; Stein and Okal, 2011; Peresan and Panza, 2012; Wyss et al., 2012; Gulkan, 2013). Such surprises arise because parameters required to reliably estimate the hazards are often poorly known (Stein et al., 2012).

A crucial parameter is  $M_{max}$ , the magnitude of the largest earthquake expected on a fault or in an area (Stein et al., 2012). The Tohoku, Sumatra, and Wenchuan earthquakes were more damaging than expected because their magnitudes were much larger than the  $M_{max}$  assumed in hazard planning (Kanamori, 2011; Sagiya, 2011). Unfortunately, inferring  $M_{max}$  is difficult. Even where we know the long-term rate of motion across a plate boundary fault, or the deformation rate across an intraplate zone, neither predict

how strain will be released although some models, like UCERF3 (Field et al., 2017), provide detailed probabilistic earthquake rupture forecasts. Estimates from the expected fault dimensions often prove incorrect. Strain release can occur seismically or aseismically, and seismic strain release can occur via earthquakes with different magnitudes and rate distributions.

As a result, quite different  $M_{max}$  estimates can be made using different methodologies (Kijko, 2004; Wheeler, 2009; U.S Nuclear Regulatory Commission, 2012; Kagan and Jackson, 2013). Because all one can say with certainty is that  $M_{max}$  is at least as large as the largest earthquake in the available record, it was earlier practice to use that magnitude or add an ad hoc increment. However, because catalogs are often short relative to the average recurrence time of large earthquakes (McGuire, 1977; Stein and Newman, 2004; Bell et al., 2013), earthquakes larger than anticipated often occur. Long paleoseismic records, such as in Cascadia (Goldfinger, et al., 2017), containing multiple earthquake cycles likely do a better job of estimating  $M_{max}$  than shorter historical catalogs. Some studies identify faults and use relations between fault length and earthquake magnitude (Wells and Coppersmith, 1994) to infer  $M_{max}$ . Other approaches extrapolate from current catalogs (Kijko, 2004) or combine areas presumed to be geologically similar to sample more large earthquakes (U.S. Nuclear Regulatory Commission, 2012; Kagan and Jackson, 2013).

Estimating  $M_{max}$  is challenging at plate boundaries, where known plate motion rates can be compared to earthquake records on known faults to infer the slip in, and thus magnitude of, large earthquakes (McCaffrey, 2008). The situation is even more complicated within plates, where deformation rates are poorly known, large earthquakes are

rarer and variable in location and time, and often occur on previously unrecognized faults (Crone et al., 2003; Camelbeeck et al., 2007; Stein et al., 2009; Clark et al., 2011; Liu et al., 2011; Leonard et al., 2014). As a result, it is unclear whether apparent differences in  $M_{max}$  between various intraplate regions are real or artifacts of the short catalogs available (Vanneste et al., 2016).

Before continuing, we should note that  $M_{max}$  has slightly different meanings depending on the assumed frequency-magnitude distribution. Some distributions assume a “hard”  $M_{max}$  that the frequency-magnitude distribution truncates at or asymptotically approaches. Other distributions, assume a “soft”  $M_{max}$  where larger earthquakes are allowed but with a much lower frequency than predicted by the un-truncated Gutenberg-Richter relationship (Kagan, 2002). In a “soft”  $M_{max}$  distribution,  $M_{max}$  is a slight misnomer as some earthquakes are expected to exceed this threshold, although they would be exceedingly rare. Whether the use of a “hard” or “soft”  $M_{max}$  is more appropriate for hazard planning, is not addressed in this chapter.

We explore the problem of  $M_{max}$  estimation via earthquake catalog for eastern North America. Notable events along the southern North America margin include the 1755 Cape Ann (Massachusetts), 1886 Charleston, and 1929 Grand Banks earthquakes (Figure 2.1). Larger earthquakes are known along the northern margin, notably the 1933 Baffin Bay event. This passive continental margin, like others, is not inert since it experiences moderate levels of seismicity (Stein et al., 1979; Stein et al., 1989; Schulte and Mooney, 2005; Wolin et al., 2012).

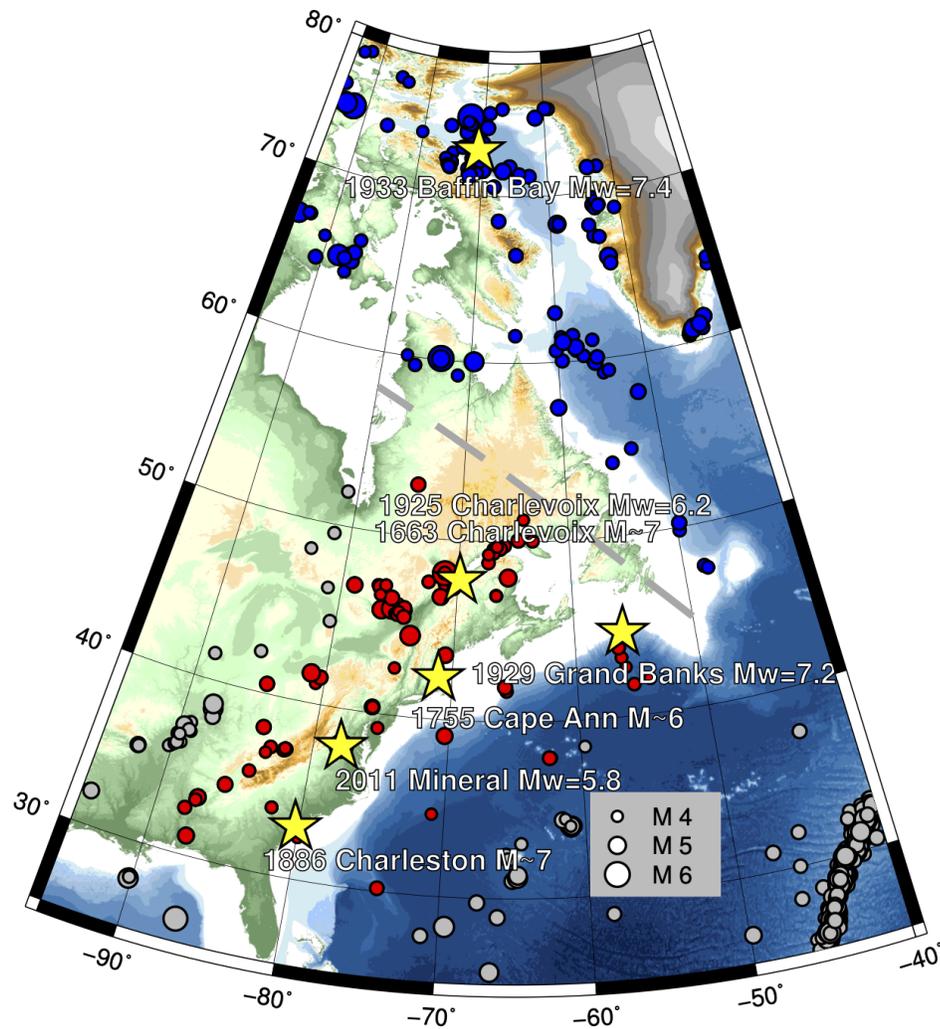


Figure 2.1. Seismicity of the eastern North America continental margin taken from the ANSS catalog from 1985 through 2017. Red and blue dots correspond to seismicity along the southern and northern North America margins, respectively. Grey dots correspond to inland and oceanic earthquakes not included in the analysis. Grey dashed line indicates boundary between southern and northern margins. Major historical events are also shown.

A challenge in assessing the earthquakes' hazard is that we know little about their causes, partly because they are relatively rare due to the slow deformation at such margins. Along plate boundaries, relative plate motion is the primary driver of seismicity. Geodynamic modeling, however, predicts that stresses from variations in topography and crustal structure across the margin, combined with sublithospheric mantle flow may have a strong influence on intraplate earthquakes (Ghosh et al., 2013). These intraplate earthquakes may reflect reactivation under these stresses of faults created by previous continental collision and breakup, given that passive margins are often reactivated (Johnston, 1989, Cloetingh et al., 2008). The complexity of the mechanisms driving earthquakes and the scarcity of large earthquakes in intraplate tectonic settings make it difficult to assess  $M_{max}$  for an individual fault. Many concepts (such as seismic cycles and recurrence intervals) that are applied to individual faults along plate boundaries are likely not appropriate for individual faults in intraplate settings (Calais et al., 2016; Liu and Stein, 2016; Clark et al., 2017). However, by considering seismicity to be homogeneously distributed over an area, we assume that the non-steady state processes that affect local and short-term seismicity are averaged out over large enough areas.  $M_{max}$  is therefore an areal  $M_{max}$  with an associated areal recurrence time that is not necessarily linked to an individual fault.

A crucial issue is how much to rely on past large earthquakes, as illustrated by successive Geological Survey of Canada hazard maps (Swafford and Stein, 2007; Adams, 2011; Wolin et al., 2012). The 1985 suite of maps concentrate hazard at the sites of the Grand Banks and Baffin Bay earthquakes, assuming that these recently active areas are especially hazardous. The 2005 (and 2015) maps have an additional ribbon of hazard

along the passive margin, assuming that similar earthquakes can occur anywhere along the margin.

The observed seismicity may be an imperfect sample of more uniform seismicity, as suggested by seismicity between the Grand Banks and Baffin Bay, some of which may be aftershocks of prehistoric earthquakes (Basham and Adams, 1983; Ebel et al., 2000; Stein and Liu, 2009; Wolin et al., 2012). Simulations with short catalogs yield apparent concentrations and gaps that are artifacts of the sampling (Swafford and Stein, 2007). Similarly, although seismicity in the eastern U.S is patchy, geological observations show evidence of slow long-term deformation (Pazzaglia et al., 2010) and the present seismicity occurs in areas that are not geologically or geomorphologically different from nearby areas that appear aseismic.

A related question is whether the larger earthquakes along the northern portions of the North America margin represent a real difference from the southern portions of the margin. The difference could be real, perhaps due to stresses associated with deglaciation (Stein et al., 1979; Stein et al., 1989; Mazzotti et al., 2005; Sella et al., 2007; Wolin et al., 2012) or to how intraplate stresses interact with the differently oriented margins, or might merely reflect the short earthquake record. We thus consider the two regions separately and explore their differences.

### 2.3. Methods

Absent reliable ways of assessing  $M_{max}$ , we use synthetic earthquake histories to explore what  $M_{max}$  values would be observed in a short catalog. We assume earthquakes satisfy a Gutenberg-Richter frequency-magnitude relation,  $\log_{10}(N) = a - bM$ , where  $N$

is the annual number of earthquakes with magnitude  $\geq M$ ,  $a$  defines the seismicity rate, and  $b$  is the slope of the line relating the rates of small and large earthquakes. The regional  $a$  and  $b$  values were estimated from the Advanced National Seismic System (ANSS) catalog for earthquakes with  $M \geq 4$  from 1985 through 2017. All earthquakes along the northeastern Canadian margin and near Hudson Strait (an area of weak extension during the opening of the Atlantic) were grouped as northern North America passive margin earthquakes (Figure 2.1). Southern North America passive margin earthquakes are all earthquakes within  $6^\circ$  of the margin south of Newfoundland. Earthquakes near the historical Grand Banks earthquake were assigned to the southern North America margin. The southern and northern margins contain 95 and 145 earthquakes greater than  $M=4$ , respectively.

A linear Gutenberg-Richter relationship can be fit to a set of frequency-magnitude data using either maximum likelihood (MLE) (Aki, 1965; Weichert, 1980; Shi and Bolt, 1982) or least squares (LSQ) estimates (Figure 2.2), each with advantages and disadvantages. In general, a least-squares fit characterizes large-magnitude occurrences better, but may not match the rate of the smaller ones well. It does well at identifying deviations from the linear Gutenberg-Richter distribution at larger magnitudes. Conversely, MLE weights numerous smaller magnitudes heavily and so gives more stable estimates of the distribution's  $b$  value (slope). Thus if one expects that the data come from a Gutenberg-Richter distribution, so that the deviations from the linear trend are artifacts of sampling or otherwise, MLE estimation is preferable. As a result, most seismic hazard analyses use MLE. Conversely, if one approaches the data without this expectation, LSQ can be viewed as a better characterization of the data themselves. Analyses seeking to estimate

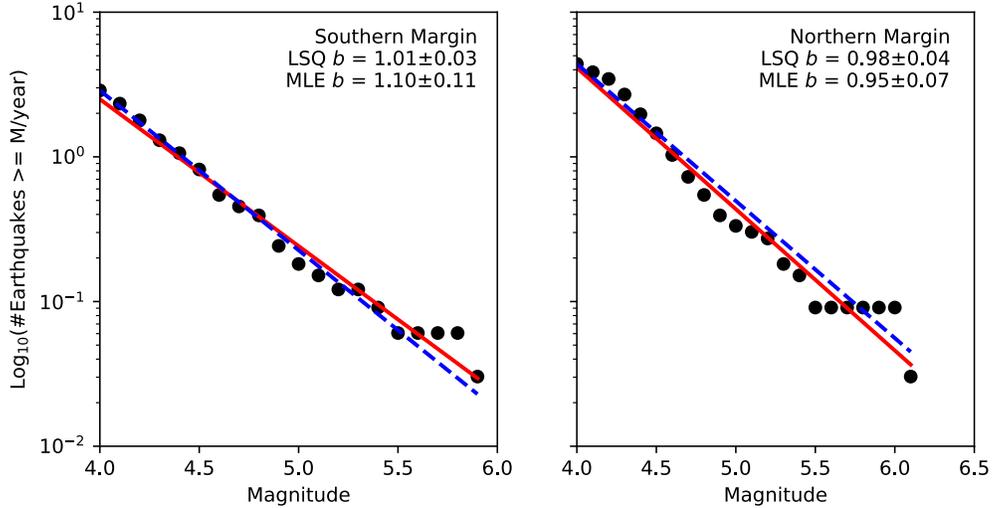


Figure 2.2. Frequency-magnitude relationships calculated for southern North America (left panel) and northern North America (right panel) datasets. Blue dashed line is for the maximum likelihood estimate (MLE) fit. Red line is for the least squares (LSQ) fit. The MLE results are used as inputs to the simulations. Uncertainties indicate  $1\sigma$ .

$b$  typically use MLE. The two methods may give different  $b$  values leading to different recurrence time predictions for large events.

Applying LSQ to relatively short catalogs may result in two biases. If the largest “observed” earthquake is smaller than expected from the parent distribution,  $b$  is biased upward (steeper slope), underestimating the rate of large earthquakes. On the other hand, if the largest earthquakes are “observed”, their recurrence interval and the region’s  $b$  are underestimated. Hence the short history causes us to either underestimate  $M_{max}$  when earthquakes of this size do not appear, or better estimate it but conclude that such earthquakes are more common than they really are (Stein and Newman, 2004).

Because deformation rates are relatively slow and seismicity may migrate between faults, earthquake recurrence times may not be well described by an earthquake cycle

model. Hence, we use a time-independent Poisson model for the entire region to describe the recurrence time between events. In a Poisson model, the probability  $P_T(\tau)$  of one earthquake greater than or equal to a given magnitude  $m$  occurring in the next  $\tau$  years depends on the rate parameter  $\lambda_m$  (Equation 2.1).

$$(2.1) \quad P_T(\tau) = 1 - e^{-\lambda_m \tau}$$

We calculate the rate parameter for a given magnitude using the  $b$  value from the MLE method and the  $a$  value projected from the measured  $M \geq 4$  recurrence rate, assumed to be the minimum magnitude of completeness (Equation 2.2).

$$(2.2) \quad \lambda_m = 10^{a-bm}$$

For the southern North America margin  $a = 4.88 \pm 0.44$  and  $b = 1.10 \pm 0.11$  whereas for the northern margin  $a = 4.43 \pm 0.27$  and  $b = 0.95 \pm 0.07$ . Although various hypotheses have been proposed for what the  $b$  value physically represents (Rundle, 1989; Todes et al., 2021), here we treat it as simply a way to parameterize the magnitude distribution curve.

To infer the time  $\tau$  until the next earthquake, we invert Equation 2.1 and set  $1 - P_T(\tau) = X$ . The value of  $X$  is sampled from a uniform random distribution  $U(0, 1]$ , where  $U$  is a continuous probability density function between 0 and 1 that has a constant probability density.

$$(2.3) \quad \tau = \frac{-1}{\lambda_m} \log X$$

Equation 2.3, where  $\log$  is the natural logarithm, indicates when the next earthquake happens but not its magnitude. To infer the magnitude, we follow a procedure similar to the recurrence interval sampling (Zhuang and Touati, 2015). The cumulative probability distribution of the linear Gutenberg-Richter relationship  $P_M(m)$  is a function of magnitude  $m$ , minimum magnitude  $m_{min}$ , and  $b$ .

$$(2.4) \quad P_M(m) = 1 - e^{-\log(10)b(m-m_{min})}$$

By substituting  $m' = m - m_{min}$ , setting  $1 - P_M(\tau) = Y$  (where  $Y$ 's value is uniformly randomly sampled from  $U(0, 1]$ ), and rewriting Equation 2.4, we can sample the magnitude cumulative distribution function to determine the magnitude of the earthquake in our synthetic catalog (Equation 2.5). Our methodology assumes a "hard"  $M_{max}$ , truncated Gutenberg-Richter distribution and discards any earthquakes that exceed the assigned  $M_{max}$  value. This assumption is made for simplicity.

$$(2.5) \quad m' = \frac{-1}{b \log 10} \log Y$$

Using the  $a$  and  $b$  values from the ANSS earthquake catalog as inputs, we generate a synthetic earthquake catalog by randomly sampling the resulting recurrence time  $P_T(\tau)$  and magnitude  $P_M(m)$  cumulative distribution functions (Figure 2.3).

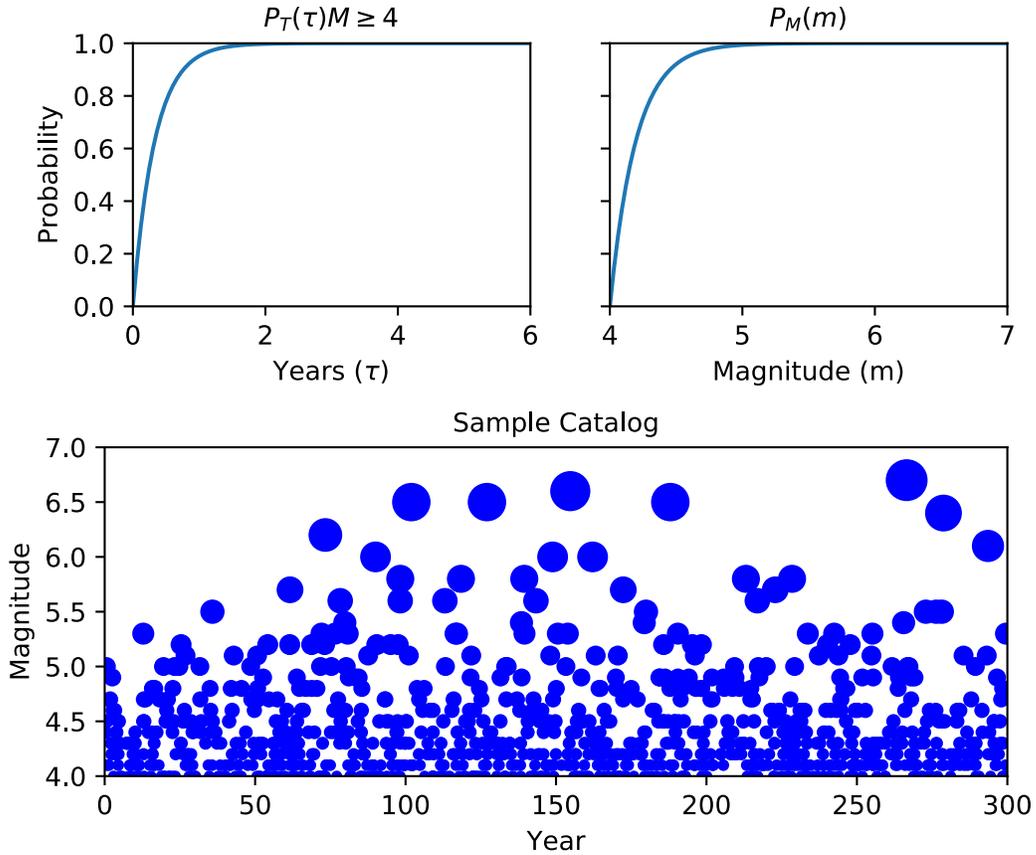


Figure 2.3. Recurrence time cumulative distribution function (top left), magnitude cumulative distribution function (top right), and sample synthetic catalog (bottom). We use  $a$  and  $b$  values from the ANSS catalog to generate the synthetic catalog.

## 2.4. Results and Analysis

For the southern North America margin, we use  $a = 4.88$  and  $b = 1.10$ , calculated from recent seismicity, corresponding to a  $M \geq 7$  earthquake on average every approximately 660 years. We generate four 10-million-year long synthetic catalogs, each with a different parent maximum magnitude,  $M_{max}^p$ , (7.0, 7.2, 7.4, and 7.6). During synthetic catalog

generation, if an earthquake’s magnitude exceeds the prescribed  $M_{max}^p$ , we remove that event from the catalog. We then create 10,000 random 300-year long sub-catalogs from the four parent catalogs, roughly corresponding to the length of the historic record for the southern North America margin.

For the northern North America margin we also generate four 10-million-year long catalogs but with  $M_{max}^p$  of 7.4, 7.6, 7.8, and 8.0, using  $a = 4.43$  and  $b = 0.95$ . The northern margin’s higher  $M_{max}^p$  values reflect the larger earthquakes observed along the northern portion of the North America margin. From these parent catalogs, we create 10,000 random 100-year long sub-catalogs, reflecting the observation record for the northern North America margin. We then compare the apparent maximum magnitude  $M_{max}^a$  “observed” in each sub-catalog to the parent distribution’s  $M_{max}^p$ .

The number of sub-catalogs whose apparent “observed”  $M_{max}^a$  matches the parent distribution’s  $M_{max}^p$  or exceeds 7.0 for the southern or 7.4 for the northern margins indicates the relative appropriateness of catalog length and assumed regional maximum magnitude based on the historical catalogs ( $M_{max}^h$ ). If relatively few sub-catalogs capture the parent distribution’s  $M_{max}^p$ , then it is unlikely that the real catalog observation window (300 and 100 years for the southern and northern margins, respectively) includes the largest magnitude events. Conversely, if  $M_{max}^a$  or a significant number of sub-catalogs exceeds the region’s assumed largest magnitude earthquakes ( $M_{max}^h$ ), then it is more likely that the catalog window length and assumed regional maximum magnitude are appropriate. In this scenario, the hypothetical larger magnitude events have a high probability of occurring, so their absence from the historical earthquake catalogs may indicate that such larger events do not occur.

Figure 2.4 shows apparent  $M_{max}^a$  values “observed” for the northern and southern margin sub-catalogs. The most common  $M_{max}^a$  value is approximately 6.7 for the northern and southern margins. For both regions, the  $M_{max}^p$  of the parent distribution is not the most common apparent  $M_{max}^a$  in the sub-catalogs. For example, along the northern margin  $M_{max}^p = 8.0$  occurs in only 2% of the sub-catalogs, and  $M_{max}^p = 7.6$  in the southern margin is present in only 2% of the sub-catalogs. For the northern margin  $M_{max}^p = 8.0$  simulation, 15% of the sub-catalogs have an apparent  $M_{max}^a$  that exceeds the commonly assumed regional  $M_{max}^h$  of 7.4. Likewise, 27% of the sub-catalogs for the southern margin  $M_{max}^p = 7.6$  simulations exceed the assumed  $M_{max}^h$  of 7.0.

Although relatively few of the sub-catalogs contain the  $M_{max}^p$  of the parent simulation, the parent  $b$  values are well recovered. The  $b$  values from the sub-catalogs given by both LSQ and MLE have Gaussian distributions (Figures 2.5, 2.6) with means similar to the input  $b$  value parameters. MLE yields tightly grouped  $b$  values with a standard deviation of approximately  $\pm 0.04$ . LSQ, however, yields  $b$  value uncertainties more than twice as large (approximately  $\pm 0.10$ ).

Figures 2.7 and 2.8 illustrate the combined uncertainties in  $M_{max}^a$  and  $b$  value found in the different sets of simulations shown in Figure 2.4. LSQ (Figure 2.7) has more scatter in  $b$  value and a strong tradeoff between  $b$  value and  $M_{max}^a$ . The LSQ method does a poor job recovering the parent distribution’s  $b$  value, which is estimated with large uncertainties, and  $M_{max}^p$ , which is generally biased low. The MLE method (Figure 2.8) better recovers the parent distribution parameters with little trade-off between  $b$  value and  $M_{max}^a$ , although  $M_{max}^a$  is still biased low.

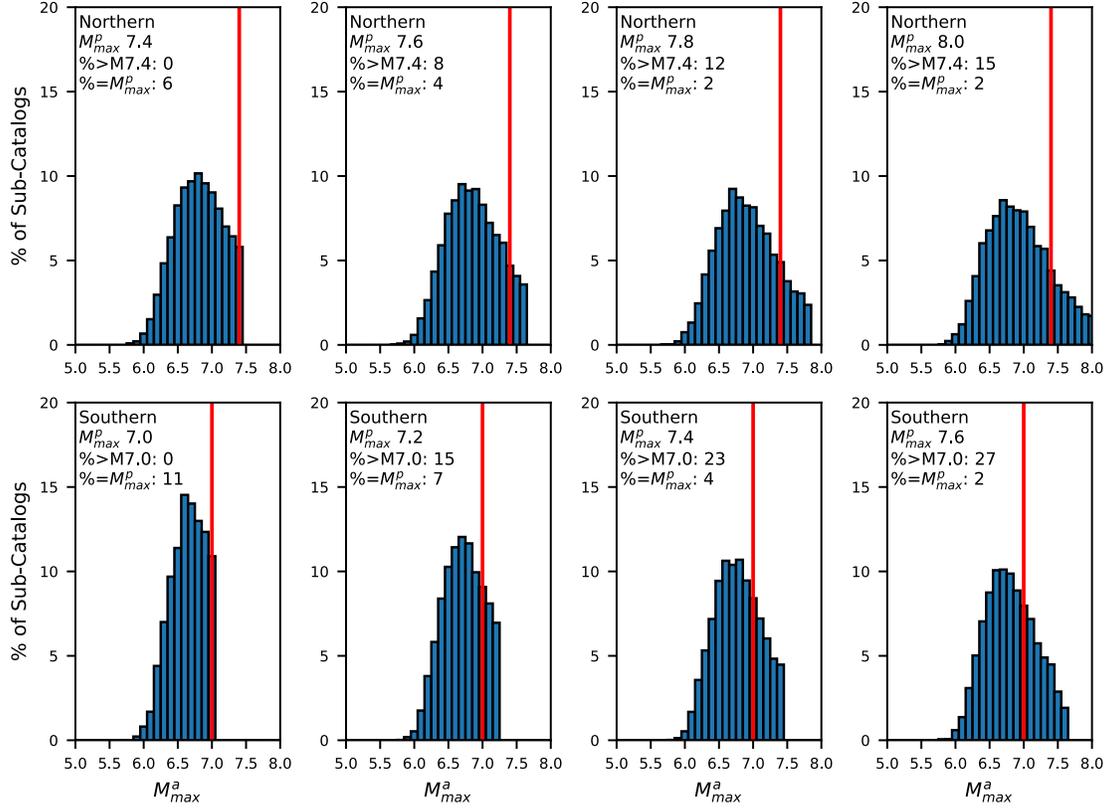


Figure 2.4. Histograms of  $M_{max}^a$  from simulated catalogs. Top: Results for four sets of 10,000 sub-catalogs for the northern North America margin, each with a different parent  $M_{max}^p$ . Panels show the percentage of simulations in which a given apparent  $M_{max}^a$  is observed. Red line represents the  $M_{max}^h$  of 7.4 inferred from the Baffin Bay earthquake. Bottom: Simulations for southern North America margin. Red line represents the inferred  $M_{max}^h$  of 7.0 corresponding approximately to the Charleston earthquake (Chapman et al., 2016), a key earthquake used for assessing seismic hazard along the North America margin (U.S. Nuclear Regulatory Commission, 2012).

The uncertainties in  $a$  and  $b$  values correspond to uncertainties in the estimated recurrence times. Simulations for  $M_{max}^p = 7.6$  in the southern margin and  $M_{max}^p = 8.0$  in the northern margin yield estimated recurrence times of these large and rare events that vary widely (Figure 2.9). The most likely recurrence times (calculated from the

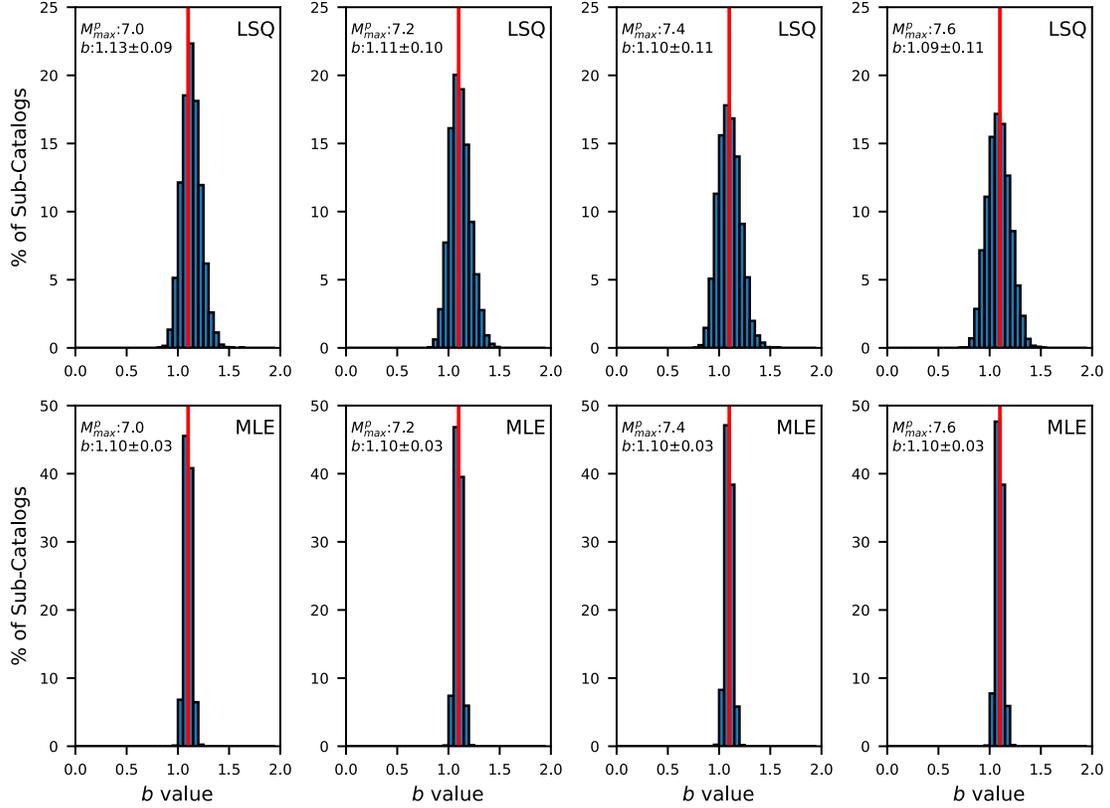


Figure 2.5.  $b$  value results for the four sets of 10,000 sub-catalogs for the southern North America margin, each drawn from a parent catalog with a different  $M_{max}^p$  and  $a = 4.88$ ,  $b = 1.10$  (red line). Mean  $b$  value and  $1\sigma$  uncertainties are listed for each simulation. Top: Panels show the percentage of simulations in which a given  $b$  value, calculated by the LSQ method, is observed. Bottom: Similar results using the MLE method to calculate  $b$  values.

sub-catalog simulations) for these largest events are biased and often shorter than the estimated recurrence times corresponding to the parent catalog parameters.

These uncertainties in the  $a$  and  $b$  values also have a significant impact on the likelihood of observing the parent distribution's  $M_{max}^p$ . We repeat the parent catalog generation process for a range of  $a$  and  $b$  values. For each  $a$  and  $b$  value combination, we generated

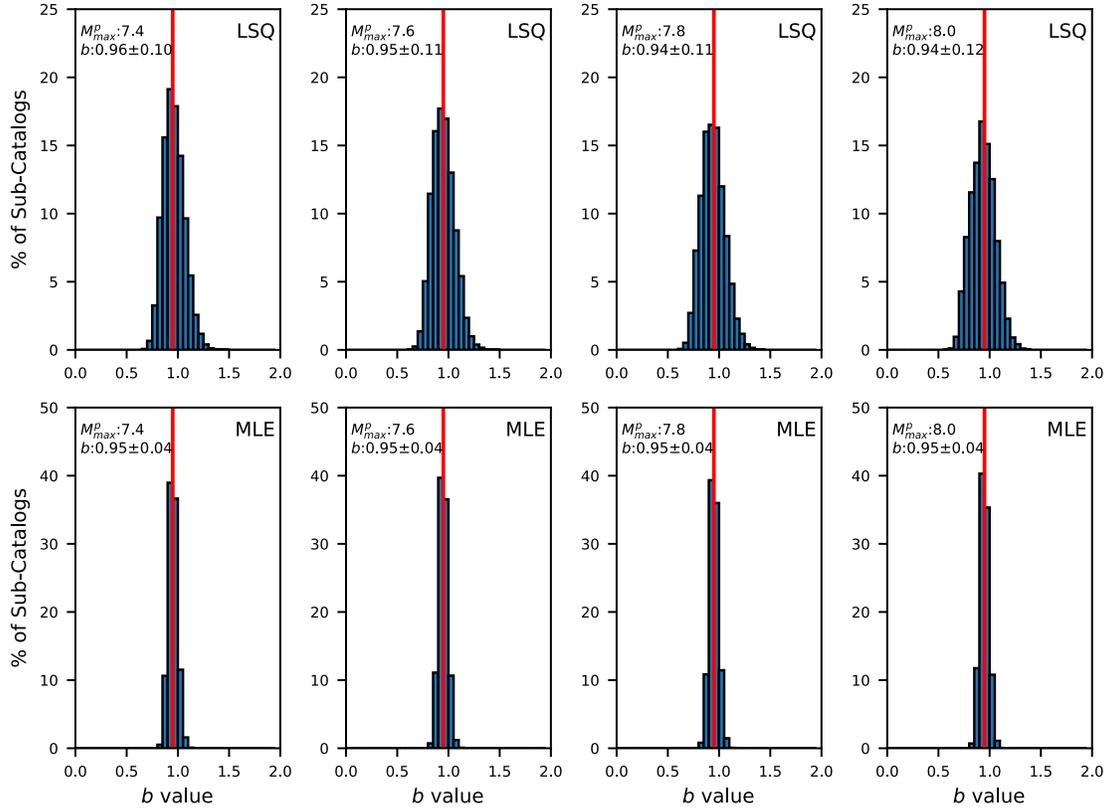


Figure 2.6.  $b$  value results for the four sets of 10,000 sub-catalogs for the northern North America margin, each drawn from a parent catalog with a different  $M_{max}^p$  and  $a = 4.43$ ,  $b = 0.95$  (red line). Mean  $b$  value and  $1\sigma$  uncertainties are listed for each simulation. Top: Panels show the percentage of simulations in which a given  $b$  value, calculated by the LSQ method, is observed. Bottom: Similar results using the MLE method to calculate  $b$  values.

a 10-million-year long catalog with  $M_{max}^p = 7.6$ . From each of these parent catalogs, we create 10,000 randomly sampled sub-catalogs for two different window lengths, 300 and 750 years (Figure 2.10). The color of the parameter space region corresponds to the percentage of sub-catalogs containing the parent  $M_{max}^p$ . For example, for a parent catalog generated using  $a = 4.4$  and  $b = 1.1$ , only 1% of the 300-year long sub-catalogs contain

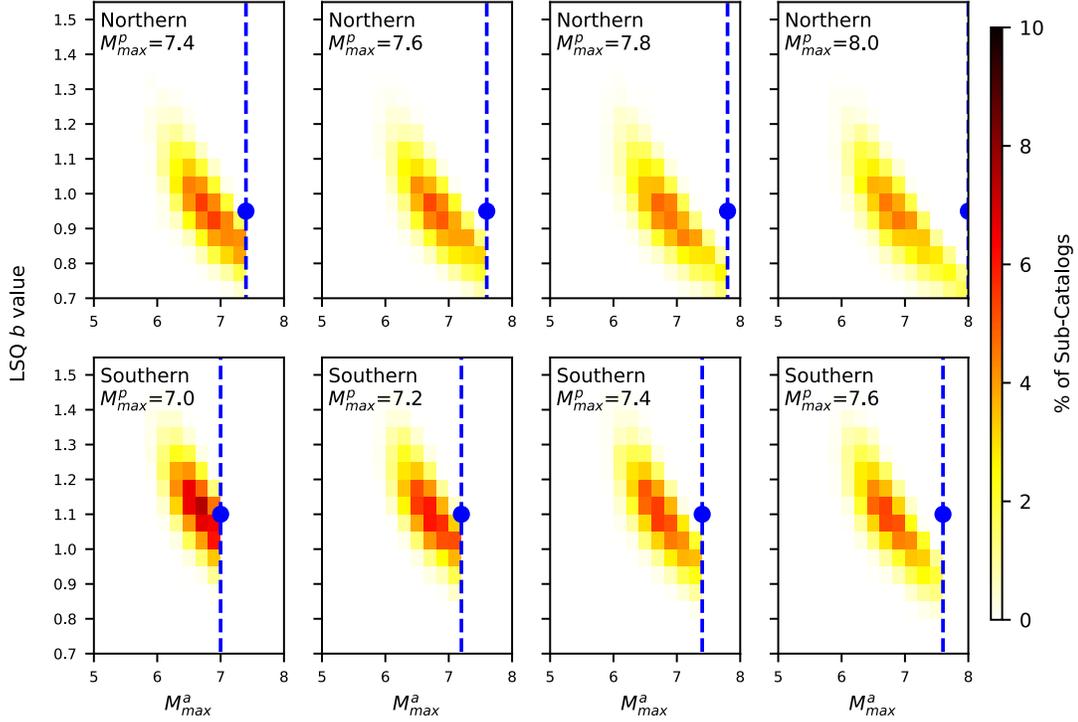


Figure 2.7. Distributions of LSQ estimates of  $b$  and  $M_{max}^a$  for the eight sets of simulations in Figure 2.4. Warmer colors indicate more sub-catalogs with those parameter values. Blue dot indicates parent distribution parameters. Dashed line shows  $M_{max}^P$ .

$M_{max}^P$ . For a parent catalog generated using  $a = 5.0$  and  $b = 0.95$ , 33% of the 300-year long sub-catalogs contain  $M_{max}^P$ . Within the parameter space, the northern and southern North America margin “true”  $a$  and  $b$  value parameters are shown with one standard deviation indicated (Figure 2.10).

Two clear patterns emerge. First, the  $b$  value more strongly influences the likelihood of observing the parent  $M_{max}^P$  than the  $a$  value. Small absolute variations in  $b$  value have a more significant impact than a similarly small change in  $a$  value. Second, as expected,

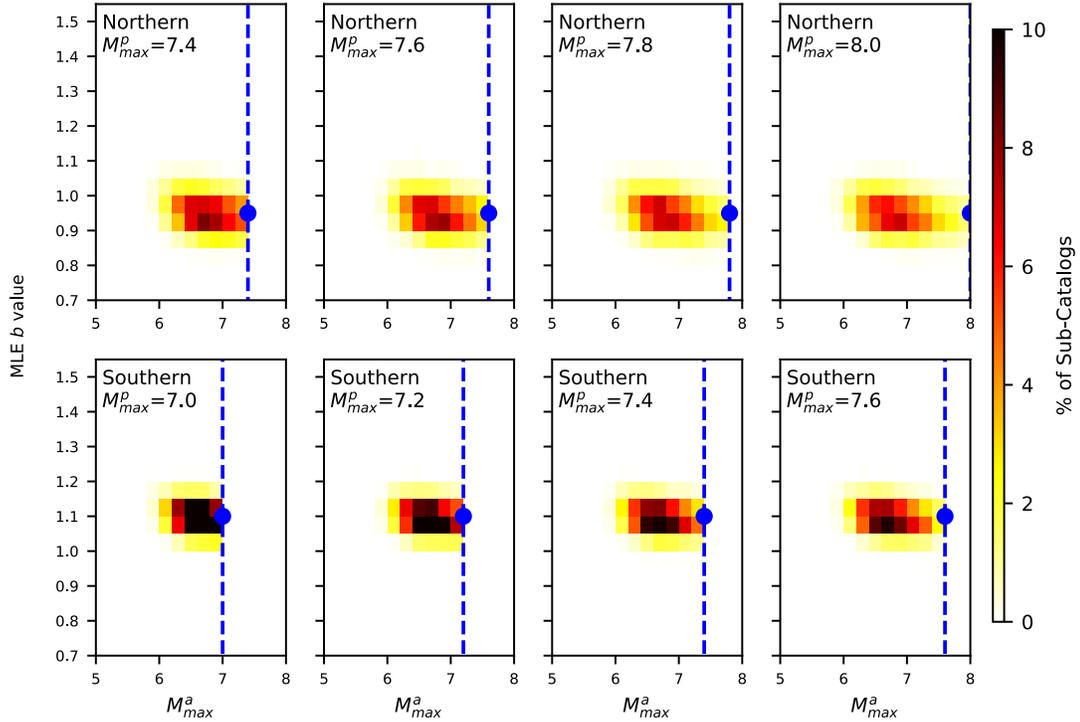


Figure 2.8. Distributions of MLE estimates of  $b$  and  $M_{max}^a$  for the eight sets of simulations in Figure 2.4. Warmer colors indicate more sub-catalogs with those parameter values. Blue dot indicates parent distribution parameters. Dashed line shows  $M_{max}^P$ .

increasing the length of the sub-catalog increases the likelihood of observing the parent  $M_{max}^P$ . However, the  $a$  and  $b$  values still largely control the likelihood of observing the parent  $M_{max}^P$ . Subtle changes in  $b$  value produce a wide range of expected mean recurrence times for large earthquakes (Figure 2.11). For  $2\sigma$  uncertainties (for the MLE  $b$  values), possible mean recurrence times span an order of magnitude for earthquakes  $M \geq 7$  along the northern (60-400 years) and the southern (150-3000 years) margins.

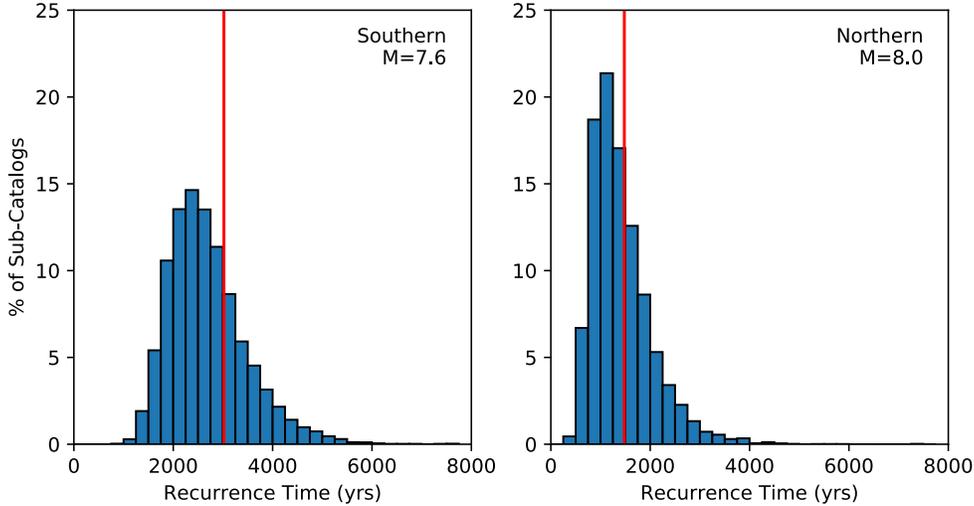


Figure 2.9. Distributions of estimated  $M_{max}^p$  recurrence. Left: Recurrence time distribution for an  $M = 7.6$  earthquake along the southern North America margin corresponding to the results of sub-catalogs with an  $M_{max}^p$  of 7.6. Right: Recurrence time distribution for a  $M = 8.0$  earthquake along the northern North America margin corresponding to the results of sub-catalogs with an  $M_{max}^p$  of 8.0. Red line indicates average recurrence time based on parent catalog parameters.

## 2.5. Discussion

The simulations suggest that future earthquakes along both margins may be larger than observed to date. The actual  $M_{max}^p$  for both margins may be the same, although if the lower  $b$  value in northern margin is real rather than a sampling artifact, large events would be more common there. A complexity is that some of the northern margin seismicity rate may reflect aftershocks of recent large earthquakes or of prehistoric earthquakes (Basham and Adams, 1983; Ebel et al., 2000; Stein and Liu, 2009; Wolin et al., 2012).

More generally, these simulations demonstrate that  $M_{max}$  cannot be reliably estimated from the available short earthquake catalogs. In both the southern and northern margin

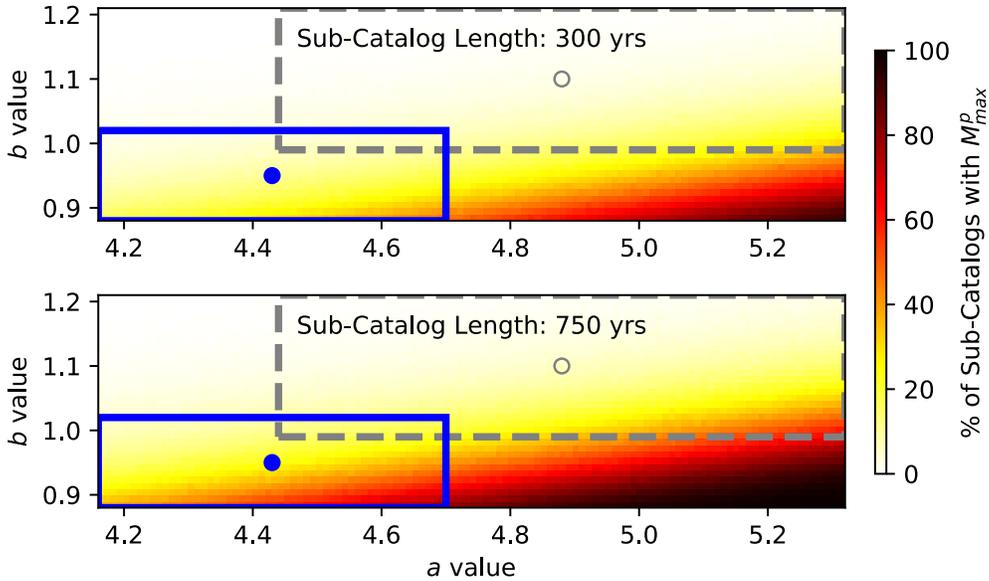


Figure 2.10. Percent of simulations yielding  $M_{max}^a = 7.6$  within the parameter space for 300 and 750 year-long sub-catalogs. For each  $a$  and  $b$  value combination shown above, we ran the synthetic catalog generation and sub-catalog sampling simulations. The color represents the percentage of sub-catalogs from a given  $a$  and  $b$  value combination that yield the parent distribution's  $M_{max}^p$  (7.6). Warmer (and black) colors indicate a greater percentage of sub-catalogs with  $M_{max}^a = M_{max}^p$ . The circles indicate  $a$  and  $b$  values used for the southern North America margin (grey open circle) and the northern North America (blue solid circle) simulations. The boxes indicate the  $1\sigma$  ranges for the southern margin (dashed grey) and northern margin (solid blue) parameters. Top: Results for 300 year-long sub-catalogs. Bottom: Results for 750 year-long sub-catalogs.  $b$  value variations have a greater impact than similarly sized  $a$  value variations in determining whether a sub-catalog will yield the parent distribution's  $M_{max}^p$ .

cases, the relatively low percentages of sub-catalogs yielding the parent catalog's maximum magnitudes indicate that these short catalogs are unlikely to include the largest earthquakes. Along these margins, the largest earthquake observed likely reflects the length of the history used, even if larger earthquakes occur. Although the precise fraction

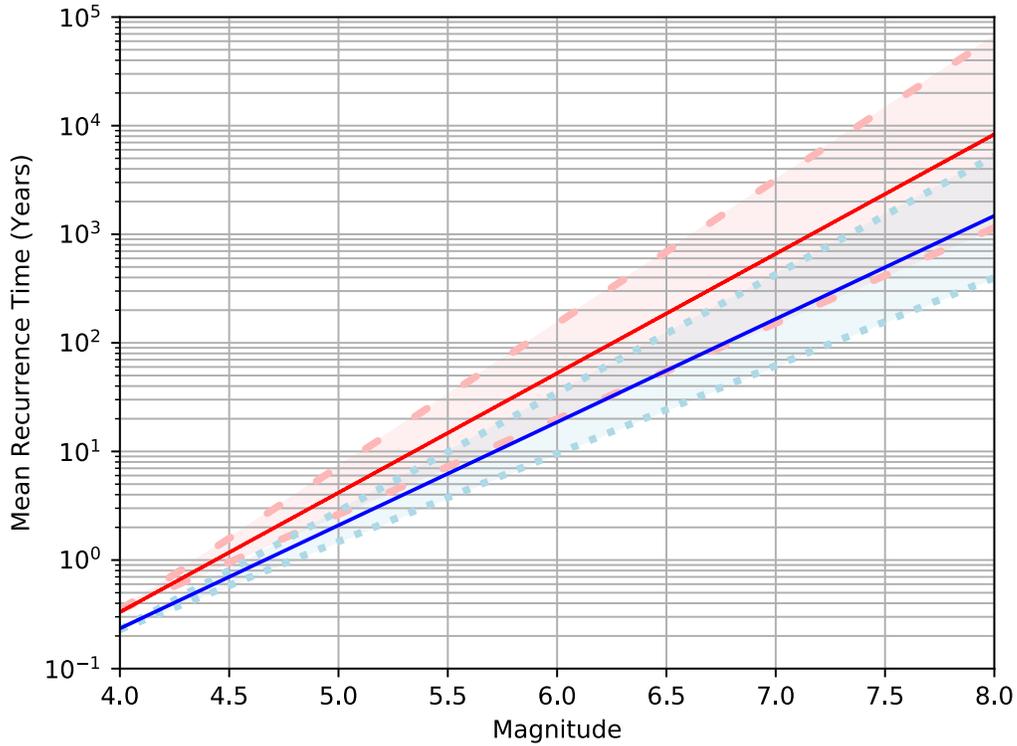


Figure 2.11. Mean recurrence times for earthquakes greater than a given magnitude. Recurrence times calculated from the inverse of Equation 2.2. Values for the northern margin (blue) and southern margin (red) are shown. The shaded envelopes indicate  $2\sigma$  uncertainties (dashed and dotted for the southern and northern margins, respectively). For earthquakes  $M > 5.5$ , there is significant overlap for expected mean recurrence times for the two regions.

depends on the distribution of recurrence times, a catalog shorter than an earthquake's mean recurrence time is unlikely to contain an event of that size (Figure 2.10).

In addition to failing to capture the true  $M_{max}$ , catalogs shorter than the mean recurrence time may not adequately capture the true  $a$  and  $b$  values. For regions with relatively infrequent large events, even small parameter uncertainties lead to recurrence

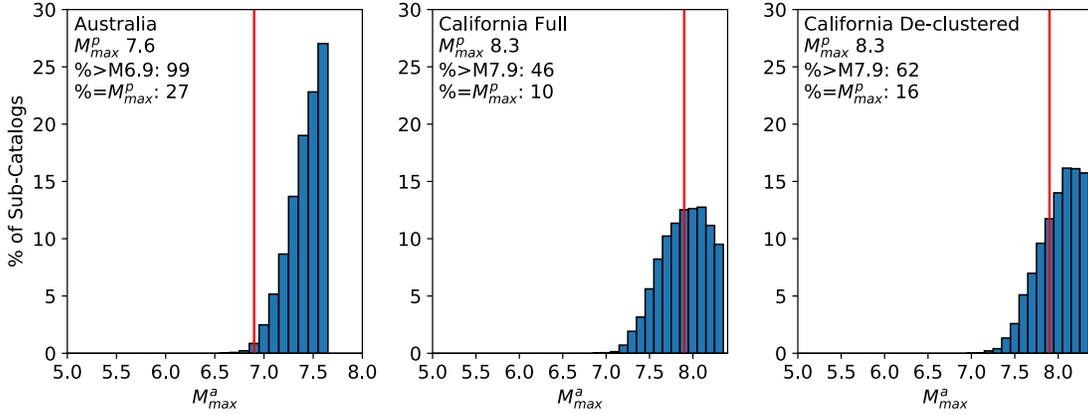


Figure 2.12. Percentage of sub-catalogs with  $M_{max}^a$  observed based on 10,000, 150 year-long sub-catalogs. Left: Results for Australia with  $M_{max}^p = 7.6$ . Red line represents the possible known  $M_{max}^h$  of 6.9, corresponding to the 1892  $M6.9$  West Tasman Sea earthquake. Middle: Results for the full California catalog with  $M_{max}^p = 8.3$ . Right: Results for a de-clustered California catalog with  $M_{max}^p = 8.3$ . For the California catalogs, the red line represents a possible  $M_{max}^h$  of 7.9 based on the 1859 Fort Tejon earthquake.

rate estimates varying by thousands of years (Figure 2.9). The calculated  $b$  value uncertainties for the northern and southern margins significantly impact whether the actual  $M_{max}$  is observed (Figure 2.10). Lower  $b$  values increase the likelihood of observing  $M_{max}$ . These small  $b$  value uncertainties also greatly affect the expected mean recurrence time for large magnitude events (Figure 2.11). Additionally, there is significant overlap in expected mean recurrence times for these two regions, indicating that despite differences in historical seismicity, these regions may be seismically similar. A well-constrained  $b$  value is therefore pivotal to calculating the likelihood of observing the actual  $M_{max}$  and estimating recurrence times for regions with relatively infrequent large events.

Similar analysis could be done for other regions. For example, we ran the simulation for the Australian continent and margin ( $a = 4.84 \pm 0.14$  and  $b = 0.90 \pm 0.03$ ) with a

sub-catalog length of 150 years and  $M_{max}^p = 7.6$  (Figure 2.12). Using the magnitudes in the Geoscience Australia catalog, we calculated the  $a$  and  $b$  values utilizing the 593 earthquakes with  $M \geq 4$  that have occurred since the beginning of 1985. The small parameter differences between Australia and the eastern North America continental margin have a significant impact on the likelihood of observing  $M_{max}^p$ . 27% of the Australia sub-catalogs contained  $M_{max}^p$ , and 99% contained an  $M_{max}^a > 6.9$ , corresponding to the 1892  $M = 6.9$  West Tasman Sea earthquake. These results indicate that, unlike in eastern North America, the activity rate in Australia may mean the historical record is much closer to being long enough to observe  $M_{max}$ .

An interesting subtle effect is that  $b$  value differences due to catalog processing also impact recurrence times for large magnitude events. For example, de-clustering the California earthquake catalog affects the probability of observing  $M_{max}$ . We ran simulations for a 150-year sub-catalog sampled from a parent catalog with  $a$  and  $b$  values from Felzer (2008) for a full ( $a = 5.93$   $b = 1.02$ ) and de-clustered ( $a = 4.83$   $b = 0.85$ ) California catalog with  $M_{max}^p = 8.3$  (Figure 2.12). These may be typical results, as declustering removes chiefly smaller earthquakes while leaving most of the larger ones. Declustering thus reduces the  $a$  value, but more importantly also the  $b$  value; it is the latter that is more important for determining whether  $M_{max}$  has been observed. Despite the full California catalog's higher earthquake productivity ( $a$  value), the de-clustered catalog's lower  $b$  value produces a shorter recurrence rate for large magnitude events. In the de-clustered catalog, 16% of the sub-catalogs capture  $M_{max}^p$ , and 62% contain  $M_{max}^a$  exceeding 7.9 (the size of 1857 Fort Tejon earthquake). Using the full catalog, 10% have  $M_{max}^p$  and only 46% of the sub-catalog  $M_{max}^a$  exceed 7.9. Based on these results, the de-clustered catalog analysis

suggests that a 150-year observation window may be long enough to observe  $M_{max}$  while the full catalog results are not as clear.

Estimates of  $M_{max}$  may be improved by substituting space for time (ergodic assumption), though there is always an issue whether different regions are “alike” (Vanneste et al., 2016). Estimates of the lower bound for  $M_{max}$  can be improved by paleoseismic investigation of active and apparently inactive faults to assess the size of past earthquakes over periods longer than the instrumental catalog (e.g., Camelbeeck et al., 2007). Geodetic studies can constrain the minimum magnitude of future earthquakes from the strain accumulation rate and time since the last large earthquake (Manaker et al., 2008). However, there is no reliable way to infer an upper bound, although various plausible assumptions can be made. Thus the only certainty about  $M_{max}$ , is that it is at least as large as that observed to date, as in the adage “anything that did happen, can happen.”

## 2.6. Conclusion

Synthetic earthquake catalog simulations for the eastern North America margin suggest that the historical record is probably not long enough to observe the region’s largest possible earthquakes. The inferred frequency of these large magnitude events is much more sensitive to a region’s  $b$  value than to its  $a$  value. For the eastern North America margin, these large magnitude events likely occur on time scales much longer than our historical record. Due to the infrequent nature of these events, the absence of a large earthquake along a segment of the margin need not reflect the long-term spatial distribution of seismicity. The different observed  $M_{max}$  values between the northern and southern margins may simply reflect the limited length of historical catalogs. Only historical and

paleoseismic catalogs that are longer than the mean recurrence time are likely to contain earthquakes approaching  $M_{max}$ . In the absence of additional evidence, it may be more appropriate to assume the hazard is uniformly distributed along the eastern North America continental margin.

## CHAPTER 3

## Large Uncertainties in Earthquake Stress Drop Estimates and Their Tectonic Consequences

### 3.1. Summary

Earthquake stress drop, the stress change on a fault due to an earthquake, is important for seismic hazard analysis because it controls the level of high frequency ground motions that damage structures. Numerous studies report that stress drops vary by tectonic environment, providing insight into a region's seismic hazard. Here we show that teleseismic stress drop estimates have large uncertainties that make it challenging to distinguish differences between the stress drops of different earthquakes. We compared stress drops for 900 earthquakes derived from two independent studies using teleseismic data and found practically zero correlation. Estimates for the same earthquake can differ by orders of magnitude. Therefore, reported stress drop differences between earthquakes may not reflect true differences. As a result of these larger uncertainties, some tectonic environment stress drop patterns that appear in one study do not appear in the other analysis of the same earthquakes. These large uncertainties in teleseismic estimates might lead to erroneous inferences about earthquake hazards. In many applications it may be more appropriate to assume that earthquakes in different regions have approximately the same average stress drop.

### 3.2. Introduction

Earthquake stress drop, the difference in stress on a fault before and after an earthquake, is an important earthquake parameter for seismic hazard analysis (Boore, 1983). An earthquake's stress drop controls the level of high frequency ground motions that damage structures (Hanks and McGuire, 1981; Cotton et al., 2013). Many studies (Kanamori and Anderson, 1975; Cocco and Rovelli, 1989; Allmann and Shearer, 2009; Courboulex et al., 2016) report that stress drops vary by tectonic environment. If a region's future earthquakes have similar stress drops to its past earthquakes, stress drop studies can provide insight into the expected levels of shaking and the region's seismic hazard. However, the uncertainties in stress drop estimates and thus these estimates' ability to resolve differences between earthquakes and different tectonic environments are poorly understood.

### 3.3. Estimating Earthquake Stress Drop

Although widely used, earthquake stress drops cannot be directly observed. Rather they are estimated via a combination of observed and assumed parameter values. An earthquake's stress drop ( $\Delta\sigma$ ) is proportional to the strain released and thus to the ratio of the seismic moment ( $M_o$ ) to the cube of the characteristic rupture dimension ( $L$ ) times a shape factor ( $c$ ) (Stein and Wysession, 2003). This assumes that there is a scaling law between  $M_o$  and  $L$ , and no saturation of  $L$  occurs at larger magnitudes.

$$(3.1) \quad \Delta\sigma = c \frac{M_o}{L^3}$$

Field studies can measure a characteristic rupture dimension (typically the fault length) for large earthquakes on land that rupture the earth's surface. However, for

most earthquakes  $L$  is estimated from seismograms. Although some seismogram techniques directly estimate rupture dimensions (Silver, 1983), most stress drop studies indirectly estimate the characteristic rupture dimension via an earthquake’s rupture duration ( $T$ ) (Courboulex et al., 2016) or its frequency-domain counterpart, corner frequency ( $f_c$ ) (Brune, 1970; Abercrombie, 1995; Allmann and Shearer, 2009; Baltay et al., 2011) (Figure 3.1). The measured rupture duration or corner frequency is combined with an assumed rupture velocity and rupture propagation model to estimate the characteristic rupture dimension (Brune, 1970; Sato and Hirasawa, 1973; Madariaga, 1976; Kaneko and Shearer, 2015). Rupture duration and corner frequency are inversely proportional ( $f_c = qT^{-1}$ ) with a constant of proportionality ( $q$ ) based on the assumed rupture model (Godano et al., 2015). For complex ruptures with multiple pulses of moment release (such as in Figure 3.1), this approach is a simplification, because the assumption of a single constant of proportionality between corner frequency and inverse rupture duration for all earthquakes may no longer be applicable. Similarly, a single stress drop estimate likely does not capture the true heterogeneity of stress release along the fault during a complex rupture.

To estimate stress drop, seismologists deconvolve seismograms to remove the effects of the seismometer and the seismic wave’s path and recover the original earthquake source signal in either the time or frequency domain (Figure 3.1). Time-domain estimates use the duration of an earthquake’s source time function (STF), a representation of the time history of the rupture (Courboulex et al., 2016). Frequency-domain studies estimate the corner frequency of the source time spectrum (Abercrombie, 1995; Baltay et al., 2011), which marks the transition between the flat, low frequency plateau and the high frequency falloff.

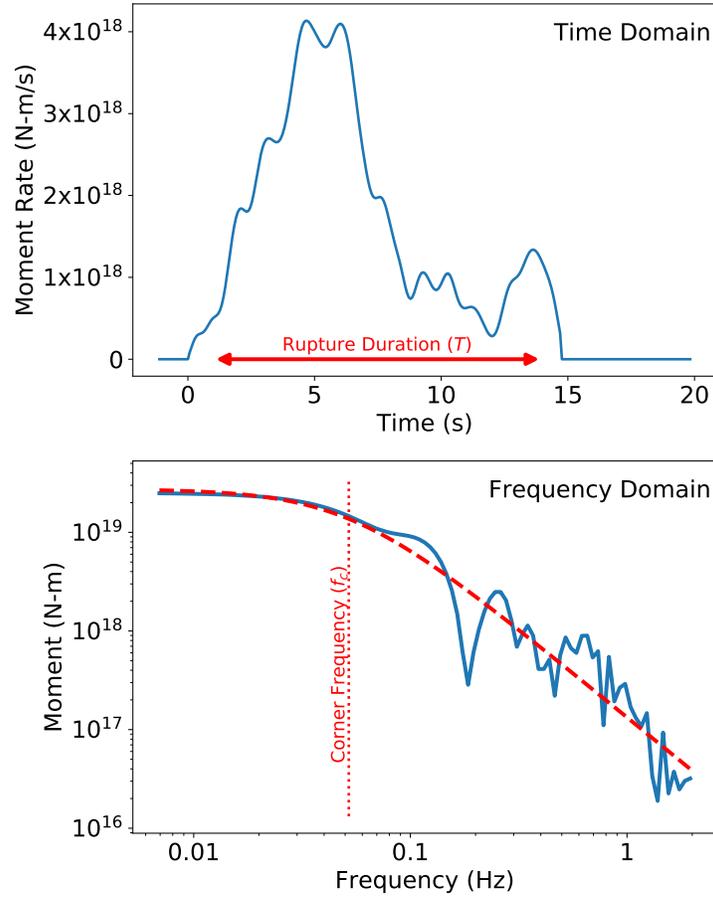


Figure 3.1. Earthquake stress drop estimates depend on estimates of characteristic rupture dimension from either rupture duration or corner frequency. Top: An earthquake’s source time function (STF) with a rupture duration estimate. Bottom: Amplitude spectrum of the above STF (solid line) with best fitting Brune spectral model (dashed line) and corner frequency (dotted line) indicated. The STF and spectrum are from an October 4th, 2000  $M_w$  6.9 earthquake near Vanuatu in the Pacific Ocean. We estimated the rupture duration and corner frequency using the methods described in this paper.

Stress drop estimates from time-domain ( $\Delta\sigma_T$ ) and frequency-domain ( $\Delta\sigma_f$ ) methods are theoretically equivalent. Both require simplifying assumptions about the earthquake’s geometry and rupture process. Most studies assume a circular fault model with a shape

factor ( $c$ ) of  $7/16$  (Eshelby, 1957), a shear wave velocity ( $\beta$ ), and a constant ( $k$ ) that incorporates an assumed rupture velocity and rupture propagation model (Brune, 1970; Sato and Hirasawa, 1973; Madariaga, 1976; Kaneko and Shearer, 2015) (Equation 3.2). Time-domain stress drop estimates include the same assumed parameters with the reciprocal of the rupture duration and a constant of proportionality ( $q$ ) replacing corner frequency.

$$(3.2) \quad \Delta\sigma_f = \frac{7M_o}{16} \left( \frac{f_c}{k\beta} \right)^3 = \Delta\sigma_T = \frac{7M_o}{16} \left( \frac{q}{k\beta T} \right)^3$$

For a given moment, higher corner frequencies or shorter rupture durations correspond to a smaller estimated characteristic rupture dimension and thus higher stress drop and greater high-frequency energy.

### 3.4. Previously Reported Differences in Earthquake Stress Drop

Stress drop estimates generally range from 0.1 to 100 MPa (1 to 1000 bars) (Allmann and Shearer, 2009) across a wide range of magnitudes (Kanamori and Anderson, 1975; Abercrombie, 1995). However, studies have suggested that stress drops may vary by tectonic environment. Previous studies have proposed that intraplate earthquakes have larger stress drops than interplate earthquakes (Kanamori and Anderson, 1975), and subduction zone earthquakes have a lower mean stress drop than non-subduction zone earthquakes (Couboulex et al., 2016). Stress drop may also vary with fault geometry. Some studies report higher stress drops for thrust fault earthquakes than normal fault earthquakes (Cocco and Rovelli, 1989). However, others find similar values for thrust and normal faulting events, with strike-slip faulting earthquakes having the largest median

stress drop (Allmann and Shearer, 2009). There is also ongoing debate whether human-induced earthquakes have lower stress drops than tectonic earthquakes (Hough, 2014; Huang et al., 2017). Previous studies have also suggested that average stress drops are similar across a broad moment magnitude range (Allmann and Shearer, 2009; Cocco et al., 2016) while other studies indicate that stress drop may vary systematically with moment magnitude (Mayeda and Walter, 1996; Malagnini et al., 2014).

Whether these reported tectonic stress drop patterns are real or artifacts of uncertainty is critical to understanding earthquake behavior. Because stress drop estimates depend on the cube of the characteristic rupture dimension—and therefore on the cube of rupture duration or corner frequency—small uncertainties in the estimated characteristic rupture dimension can significantly impact stress drop estimates (Stein and Wysession, 2003; Kane et al., 2011; Cotton et al., 2013; Kaneko and Shearer, 2015). Comparisons of multiple stress drop estimates for the same earthquakes provide insight into the impacts of these uncertainties. Although the magnitudes of the stress drop estimates may differ between studies, the relative values (i.e., which earthquakes have small stress drops and which have large ones) should be similar if the uncertainties are small. However, large discrepancies between stress drop estimates for the same earthquakes would suggest that the estimates have significant uncertainties. Previous comparative studies have been limited in scope and produced mixed results, with some indicating strong correlation between different estimates (Baltay et al., 2011, Baltay et al., 2013) and others showing only limited correlation between a few of the earthquakes (Allmann and Shearer, 2007; Abercrombie, 2014).

### 3.5. Large Uncertainty in Stress Drop Estimates

We expand on these previous comparative studies by using a much larger dataset and comparing two independent stress drop studies using teleseismic (distances greater than 1000 km) data for the same earthquakes. The earthquakes are mostly shallower than 50 km, with moment magnitudes from approximately 5.5 to 8.3. We estimate time-domain stress drop values using source time functions from the SCARDEC database (Vallée and Douet, 2016) and compare them to frequency-domain estimates from a previous independent study by Allmann and Shearer (2009) (hereinafter called AS09). We compared magnitudes, locations, and event times to identify 894 earthquakes common to both databases. For a full list of earthquakes used in this analysis, see Data S1 in the Supplemental Material.

The SCARDEC database contains two types of STF: average and optimal. The optimal STF is the individual STF that most closely matches the earthquake’s average STF. Because averaging STF reduces the high frequency content, the optimal STF is more appropriate for spectral analyses (Vallée and Douet, 2016). We use the optimal STF in this study.

In principle, the rupture duration can be estimated from the full duration of the source time function (STF). However, STF often have long tails that do not necessarily reflect the true rupture duration, which tends to be shorter. We use two methods proposed by Courboux et al. (2016) to estimate rupture duration. The first method uses the first and last points of the STF with values that exceed 10% of the maximum STF moment rate ( $STF_{max}$ ) to determine the rupture duration. The second method models the source time function as an isosceles triangle with height =  $STF_{max}$  and area equal to the total

seismic moment released (the integral of the STF). The width of this isosceles triangle is then an estimate of rupture duration. While the precise rupture duration values differ somewhat, both methods generally give similar rupture duration estimates (Courboulex et al., 2016). We average the two rupture duration estimates to produce a duration value ( $T$ ) for the earthquake based on the optimal STF.

To ensure consistency between our stress drop estimates and the AS09 frequency-domain estimates, we use the same assumed parameters ( $k = 0.32$ ;  $\beta = 3900$  m/s) as AS09. For our time-domain estimates, we use a value of 1 for the constant of proportionality ( $q$ ) between corner frequency and inverse rupture duration. Our time-domain stress drop estimates use the seismic moment values from the SCARDEC database.

Because stress drop estimates span several orders of magnitude, the logarithmic domain is a natural scale for stress drop analyses (e.g., Cotton et al., 2013). Therefore, we perform our statistical analyses using the  $\log_{10}$  values of the stress drop estimates.

Surprisingly, the time-domain estimates are essentially uncorrelated with the AS09 frequency-domain estimates (Figure 3.2a). If the two estimates produced similar results, we would expect a clear linear trend. Instead we find large scatter and very low correlation (0.07) between the stress drop estimates. These differences suggest significant uncertainty in stress drop estimates. However they do not indicate whether the uncertainty arises from the time-domain estimates, the AS09 frequency-domain estimates, or both. Using a different constant of proportionality between corner frequency and inverse rupture duration shifts the time-domain estimates, but does not change the correlation.

The distribution of the frequency-domain to time-domain stress drop ratios provides insight into the reproducibility of a stress drop estimate (Figure 3.2b, Figure 3.3). On

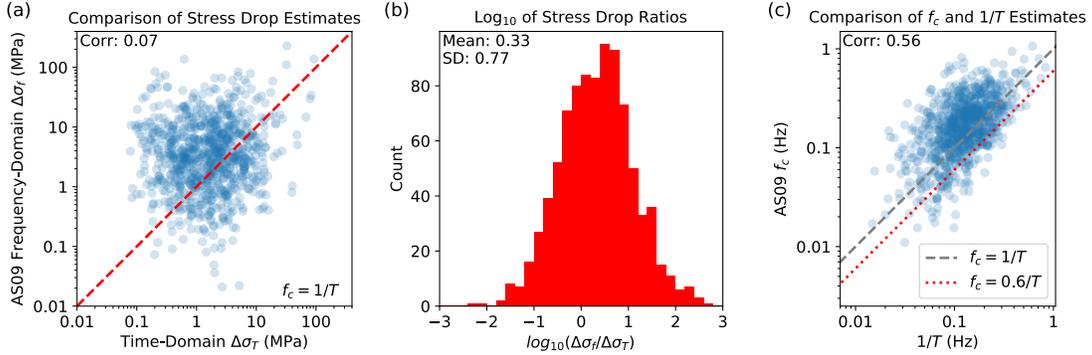


Figure 3.2. Correlation plots of AS09 frequency-domain and time-domain estimates of stress drop and rupture duration for individual earthquakes. (a) Time-domain stress drop estimates and corresponding AS09 frequency-domain estimates. Dashed line indicates ideal corresponding values. The two estimates are essentially uncorrelated. (b) Histogram of the logarithm of the ratios of AS09 frequency-domain to time-domain stress drop estimates with mean and standard deviation (SD) indicated. All units are  $\log_{10}$ . (c) AS09 corner frequencies ( $f_c$ ) versus inverse rupture duration ( $\frac{1}{T}$ ) estimates, which should be equivalent but instead show only moderate correlation.

average, the AS09 frequency-domain estimates are twice as large (0.33 in  $\log_{10}$  units) as the time-domain estimates, but the ratio's large standard deviation (0.77 in  $\log_{10}$  units) indicates that the relative values of the two estimates differ substantially between earthquakes. Changing the constant of proportionality ( $q$ ) between corner frequency and inverse rupture duration shifts the mean value of the stress drop ratios but does not impact the standard deviation. For a time-domain stress drop estimate of 1 MPa (0 in  $\log_{10}$  units), the corresponding AS09 frequency-domain estimate would be on average 2.1 MPa (0.33 in  $\log_{10}$  units). If we consider the one standard deviation plus/minus range, the corresponding value may fall between 0.4 to 12.6 MPa (-0.44 to 1.1 in  $\log_{10}$  units). Using plus/minus two standard deviations, the possible range spans 0.06 to 74.1 MPa (-1.21 to 1.87 in  $\log_{10}$  units). Therefore, attempts to reproduce an estimate using the

other method may produce a corresponding value that falls anywhere within a 3 orders of magnitude range. Our possible range of corresponding values is approximately the same size as previously reported stress drop variations among earthquakes (Allmann and Shearer, 2009). These similar ranges suggest that reported stress drop differences may not reflect the true differences in stress drop between earthquakes.

The large scatter in stress drop results from discrepancies in the characteristic rupture dimension inferred from our time-domain and the AS09 frequency-domain results. Figure 3.2c compares the AS09 corner frequency estimates to the inverse of our rupture duration estimates. Although the two estimates should tightly cluster about a straight line, they have only a moderate correlation (0.56) with considerable scatter. The higher correlation reflects the similar dependence of corner frequency and inverse rupture duration on moment because larger earthquakes have longer durations and lower corner frequencies. However, the large scatter suggests large uncertainties in estimates of characteristic rupture dimension.

### **3.6. Possible Sources of Earthquake Stress Drop Uncertainty**

Understanding the cause of the corner frequency and rupture duration discrepancies has significant implications for the estimation of stress drop. Although the time-domain and frequency-domain methods are theoretically equivalent, the observed discrepancies could simply reflect the uncertainties of analyzing real data using different methods. The corner frequency and rupture duration estimation methods assume relatively simple earthquake source models. It is possible that this simple parameterization of an earthquake rupture is inappropriate for complex earthquakes.

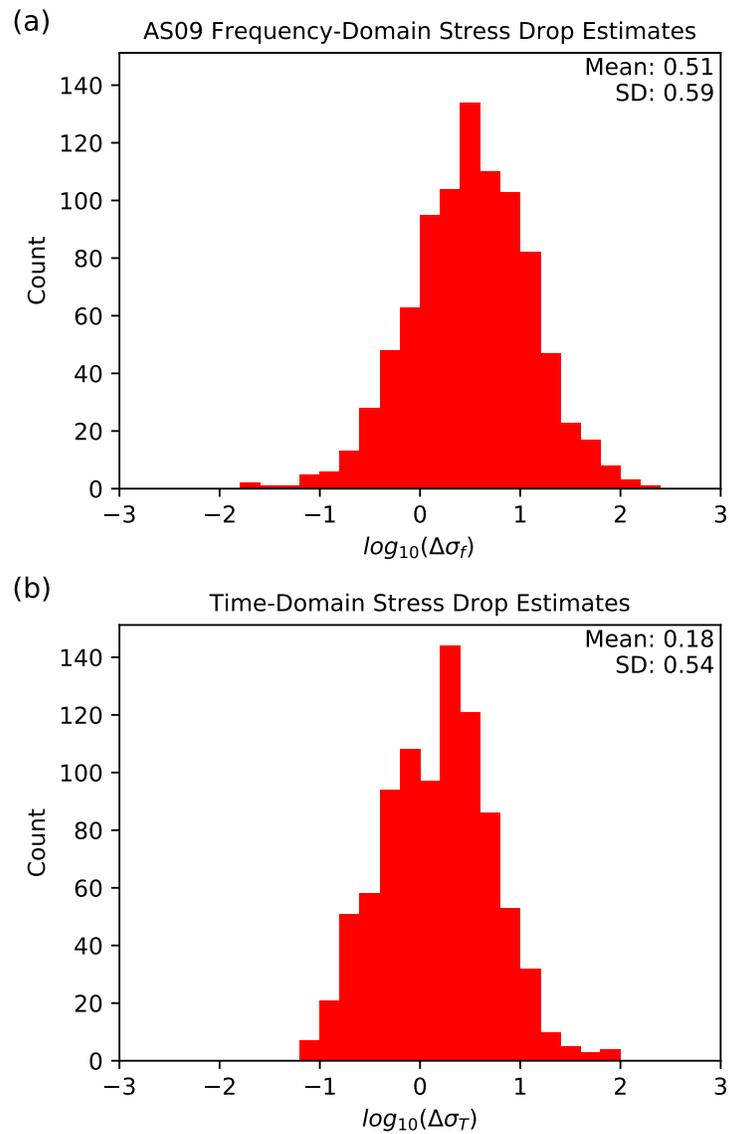


Figure 3.3. Distributions of stress drop estimates in  $\log_{10}$  space. Panel (a) shows AS09 frequency-domain stress drop estimates, and panel (b) shows our time-domain estimates. All units are  $\log_{10}$  MPa. Mean and standard deviation (SD) indicated. The ratio of these estimates is shown in Figure 3.2b.

We explore the possibility that simpler earthquake ruptures may produce better agreement between the two studies by binning the earthquakes by magnitude and recalculating the stress drop estimate correlations. Dividing the earthquakes into four magnitude bins ( $M_w < 6.0$ ;  $6.0 \leq M_w < 6.5$ ;  $6.5 \leq M_w < 7.0$ ;  $M_w \geq 7.0$ ) produces a slight decrease in correlation with increasing magnitude (Figure 3.4). The slightly better agreement in stress drop estimates for smaller magnitude earthquakes is consistent with these events generally having simpler ruptures. However, even the correlation for the smallest earthquakes ( $M_w < 6.0$ ) is still only 0.25 (Figure 3.4a), indicating significant disagreement between the two studies for even the simplest ruptures in the dataset.

We cannot attribute all of the stress drop discrepancies to the different time-domain and frequency-domain methods because there are several other major methodological differences between the two studies. The two studies considered here used different deconvolution procedures to remove the effects of seismic wave propagation between the earthquake and seismometers and recover the original earthquake signal. Additionally, the two studies did not necessarily use the same seismograms for the same earthquakes. To assess the contribution of the time-domain and frequency-domain methods to the observed stress drop discrepancies, we must apply these methods to the same exact data. We address this data consistency issue by transforming the SCARDEC STFs into the frequency domain and finding their corner frequencies. This analysis ensures that any observed stress drop discrepancy only reflects the different techniques used to infer rupture duration in the time and frequency domains.

We pad the SCARDEC STFs with zeros to increase the length of the signal to 5 times the original STF length and perform a Fourier Transform. We make two modifications to

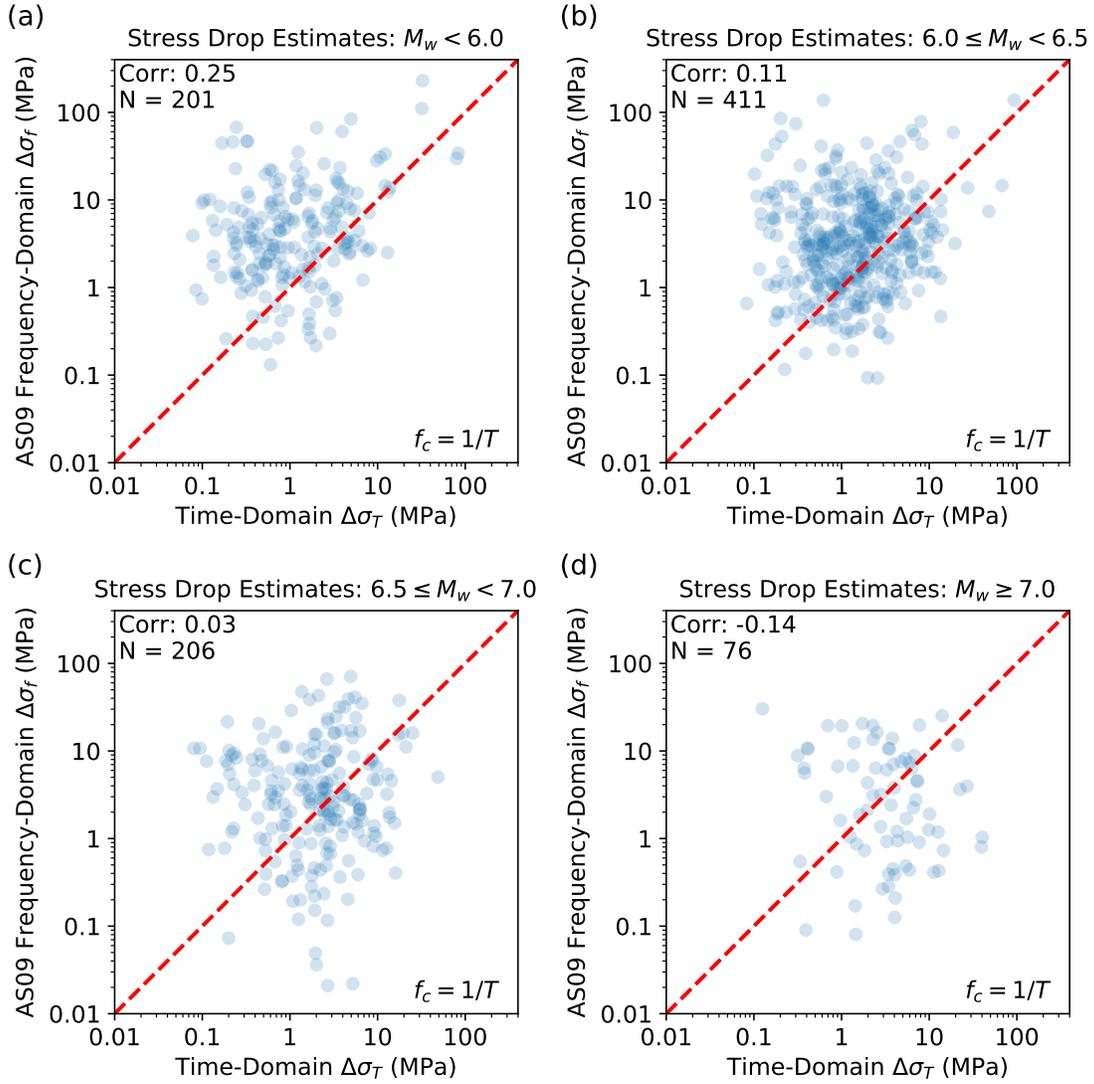


Figure 3.4. Correlation plots of AS09 frequency-domain and time-domain estimates of stress drops for earthquakes grouped by magnitude. Correlation is marginally better for smaller earthquakes (a) but significant scatter still exists for all magnitude ranges. Number of earthquakes ( $N$ ) indicated for each magnitude group. Dashed line indicates ideal corresponding values.

the resulting amplitude spectrum before fitting the Brune model (Equation 3.3). First, we low-pass filter the signal below 2 Hz. This ensures that high frequency noise does not

unduly influence the fitting of the Brune model. Second, we evenly resample the spectrum amplitude in  $\log_{10}$  space using a 0.025 spacing and remove the zero frequency amplitude. This creates additional points along the low frequency plateau of the spectrum and ensures that the fitting procedure is not dominated by the high frequency falloff above the corner frequency (Imanishi and Ellsworth, 2006). Using the Brune (1970) displacement spectral model,

$$(3.3) \quad A(f) = \frac{\Omega_0}{\left[1 + \left(\frac{f}{f_c}\right)^n\right]}$$

we solve for the best fitting corner frequency ( $f_c$ ), low frequency plateau height ( $\Omega_0$ ), and spectral falloff ( $n$ ) in  $\log_{10}$  space. At frequencies ( $f$ ) below the corner frequency, the spectral amplitude  $A(f)$  has a flat plateau. At frequencies above the corner frequency, the spectrum decreases at falloff rate  $n$ . We perform a grid search to find the optimal parameter combination that minimizes the least squares residuals. The corner frequency grid search spans 0.001 Hz to 1 Hz in increments of 0.001 Hz. The low frequency plateau search spans 0.5 to 3.5 times the maximum spectral amplitude in increments of 0.05. The spectral falloff range covers 0.5 to 4.5 in increments of 0.1. Our procedure differs from the AS09 corner frequency fitting procedures, which assumes the same fixed falloff rate of 1.6 for all earthquakes (Allmann and Shearer, 2009).

The resulting stress drop estimates (Figure 3.5a) are better correlated (0.66) than those for the two independent studies (0.07). The two estimates have a clear linear trend with some scatter. The histogram of the stress drop ratios (Figure 3.5b) also has a much smaller standard deviation (0.49) compared to the standard deviation of the ratios of the two independent studies (0.77). The increased agreement between the two stress

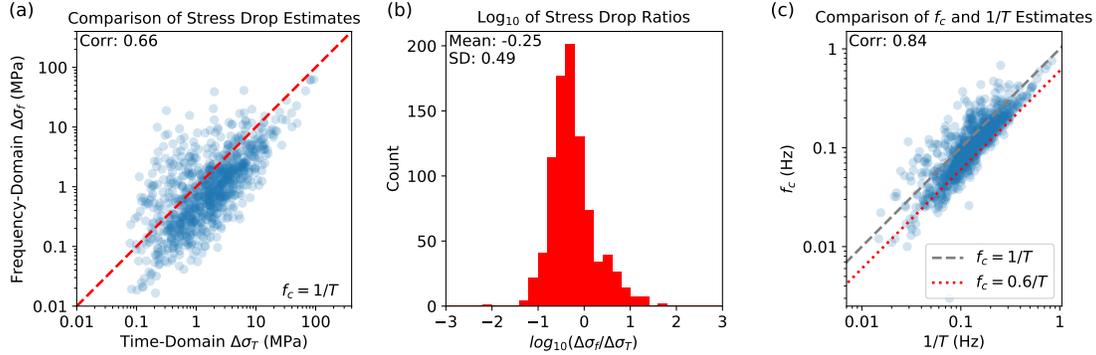


Figure 3.5. Correlation plots of frequency-domain and time-domain estimates of stress drop and rupture duration for individual earthquakes showing the effect of applying different methods to the same data. (a) Time-domain stress drop estimates and corresponding frequency-domain estimates for the same data. Dashed line indicates ideal corresponding values. These estimates show better correlation compared to Figure 3.2. (b) Histogram of the logarithm of the ratios of frequency-domain to time-domain stress drop estimates with mean and standard deviation (SD) indicated. All units are  $\log_{10}$ . (c) Corner frequency ( $f_c$ ) and inverse rupture duration ( $\frac{1}{T}$ ) estimates show a strong correlation.

drop estimates reflects the improved correlation (from 0.56 to 0.84) between the corner frequency and inverse rupture duration estimates (Figure 3.5c).

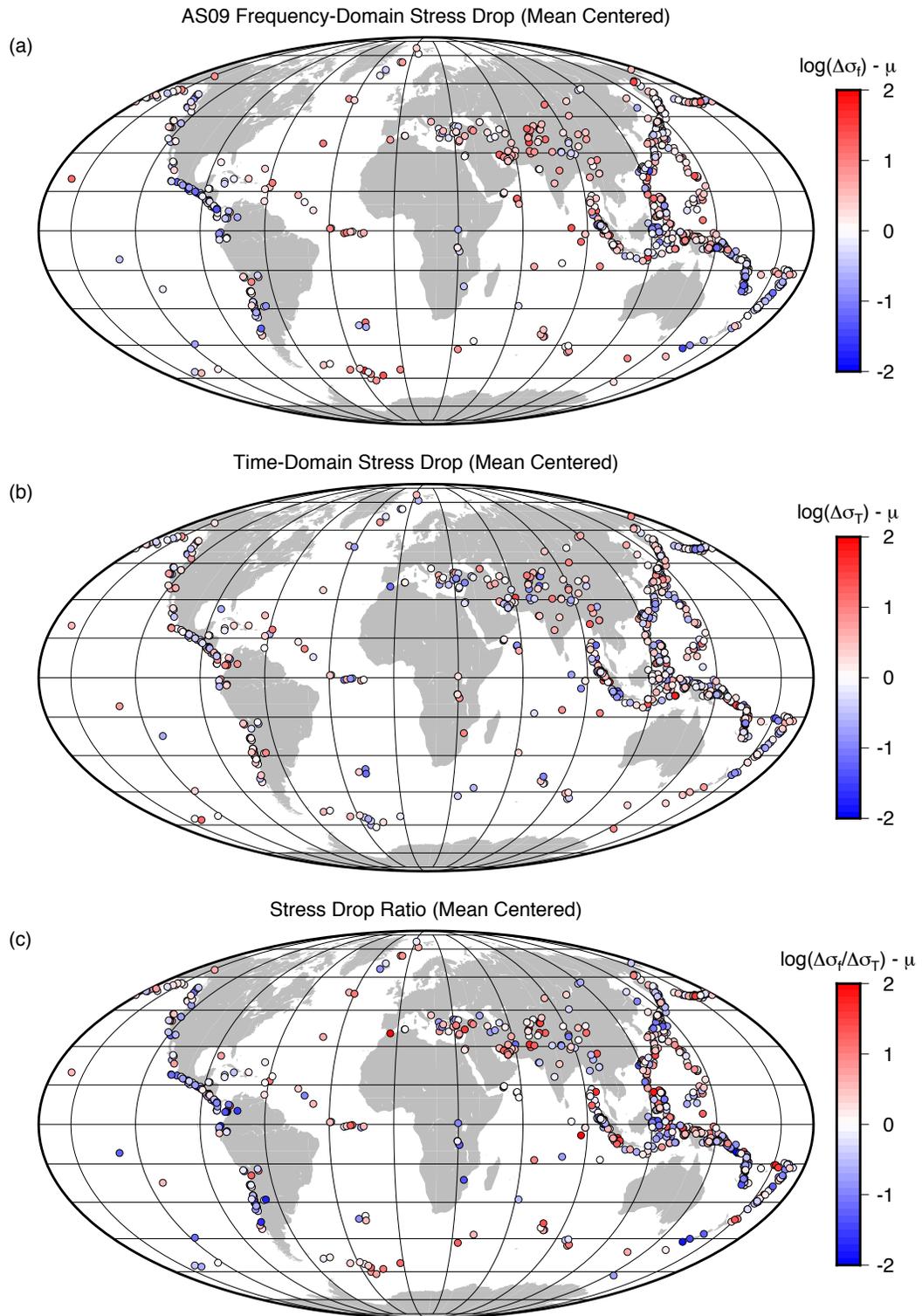
The remaining scatter between corner frequency and inverse rupture duration (Figure 3.5c) may reflect uncertainties in the analytical procedures or real differences in the constant of proportionality ( $q$ ) between corner frequency and inverse rupture duration between individual earthquakes (Atkinson and Beresnev, 1997). We conjecture that the different deconvolution procedures between the two studies may account for the low correlation between the AS09 corner frequencies and the inverse rupture duration estimates (Figure 3.2c). Additional sources of uncertainty may include the specifics of how rupture duration and corner frequency are estimated and the possibility that the two studies included different seismograms for the same earthquakes.

### 3.7. Observed Stress Drop Trends Vary Depending on Analysis

Spatial stress drop trends that appear in one dataset do not necessarily appear in the other. Figure 3.6 shows the spatial distribution of the stress drops for each dataset. The AS09 frequency-domain stress drop dataset shows a few spatially coherent stress drop patterns. Earthquakes along Central America have relatively lower stress drops, whereas earthquakes in southern Central Asia and Iran have higher stress drops (Figure 3.6a). In the time-domain results (Figure 3.6b), the negative anomaly stress drop trend along Central America is less pronounced, and the southern Central Asia and Iran earthquakes do not show a spatially coherent pattern.

Comparison of the stress drop ratios (AS09 frequency-domain stress drop divided by the time-domain stress drop) allows for the identification of spatially coherent patterns in the ratio of the two datasets' values. In Figure 3.6c, reds indicate earthquakes where the stress drop ratio is larger than the average ratio, and blues indicate earthquakes where the value is lower than average. A few regions exhibit some spatial coherence. Ratios lower than expected (blues) dominate along the west coasts of Central and South America. Conversely regions like the South Sandwich Trench in the South Atlantic, southern Central Asia, and southern Iran seem to have predominantly higher stress drop

Figure 3.6 (*following page*). Stress drop estimates for the 894 earthquakes in this study. Plots are mean ( $\mu$ ) centered so that the color scale corresponds to the mean  $\log_{10}$  value for that dataset. (a) Mean  $\log_{10}$  centered AS09 frequency-domain stress drop estimates. (b) Mean  $\log_{10}$  centered time-domain stress drop estimates. (c) Mean  $\log_{10}$  centered ratio of AS09 frequency-domain and time-domain stress drop estimates. Only limited regional stress drop coherence appears between the two studies.



ratios than expected. Other regions, especially along the trenches of the Indian and Pacific Oceans, show significant heterogeneity in stress drop ratios.

The large discrepancies between the time-domain and AS09 frequency-domain stress drop estimates also suggest that reported trends in stress drop between tectonic environments may not be robust. Dividing our dataset of 894 earthquakes by fault geometry and by lithospheric type produces strikingly different trends in stress drop distributions between the two studies. We classify earthquake fault geometry based on the plunge of the P and T axes of the moment tensor (Frohlich, 1992) from the GCMT catalog (Ekström et al., 2012). The Frohlich (1992) classification scheme is arbitrarily asymmetric with the thrust classification spanning a larger segment of the T axis space (Todes et al., 2021). As previously reported (Allmann and Shearer, 2009), the AS09 frequency-domain estimates indicate that strike-slip earthquakes have a larger median stress drop (11.0 MPa) than other fault geometries combined (2.7 MPa) (Figure 3.7, top). Our time-domain estimates, however, indicate that the median strike-slip earthquake (1.4 MPa) is indistinguishable from the other geometries combined (1.7 MPa) (Figure 3.7, top). Classifying the AS09 frequency-domain estimates by lithosphere type using the U.S. Geological Survey’s SeismoTectonic Regime Earthquake Calculator (STREC) software indicates that earthquakes in oceanic lithosphere have a higher median stress drop (4.8 MPa) than non-oceanic lithosphere earthquakes (3.0 MPa) (Figure 3.7, middle). However, in the time-domain data, the median estimate for oceanic lithosphere (1.8 MPa) is very similar to non-oceanic lithosphere (1.6 MPa) (Figure 3.7, middle). Although the Wilcoxon rank-sum test indicates that the higher strike-slip ( $p$ -value  $< 0.001$ ) and oceanic lithosphere ( $p$ -value  $< 0.001$ ) AS09 frequency-domain stress drops are statistically significant at the 0.01 level,

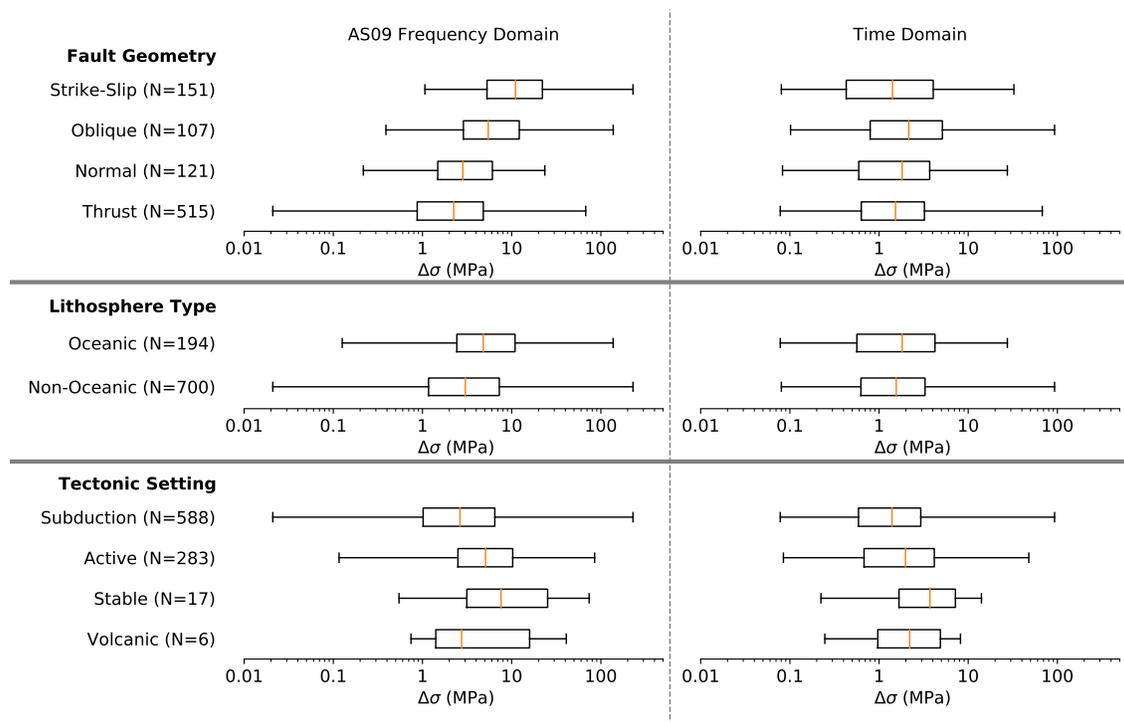


Figure 3.7. Stress drop distributions for different tectonic environments. Box and whisker plots show median value (vertical line), 1st and 3rd quartiles (ends of boxes), and min/max values. The three rows indicate different classification schemes with distinct sub-groups. Left and right columns are AS09 frequency-domain and our time-domain estimates. Number of events (N) in each subgroup indicated. Tectonic stress drop trends vary between the two estimates.

the time-domain estimates for the same earthquakes show no statistically significant differences (p-values of 0.26 and 0.27, respectively) (Table 3.1).

Classifying the earthquakes by tectonic setting using STREC indicates similar trends between the two studies (Figure 3.7, bottom). Both show that subduction zone earthquakes have a lower median stress drop than active region (non-subduction) earthquakes—2.6 vs. 5.1 MPa for AS09 frequency-domain estimates and 1.4 vs. 2.0 MPa for time-domain estimates. Regions of ongoing tectonic deformation that are not subduction zones are

Table 3.1. Wilcoxon rank-sum test (two-tailed) for statistical significance of tectonic environment stress drop patterns. We compared the distributions of strike-slip vs. non-strike-slip, oceanic vs. non-oceanic, and subduction vs. active zone (non-subduction) earthquakes. We used the Wilcoxon rank-sum statistic to test the null ( $H_0$ ) and alternative ( $H_a$ ) hypotheses.

Hypothesis Test	AS09 Frequency Domain		Time Domain	
	Statistic	p-value	Statistic	p-value
$H_0$ : Strike-slip = Non-Strike-slip	92652	< 0.001	52843	0.26
$H_a$ : Strike-slip $\neq$ Non-Strike-slip				
$H_0$ : Oceanic = Non-Oceanic	83252	< 0.001	71396	0.27
$H_a$ : Oceanic $\neq$ Non-Oceanic				
$H_0$ : Subduction = Active	58341	< 0.001	72922	0.003
$H_a$ : Subduction $\neq$ Active				

considered active tectonic settings. These differences are statistically significant at the 0.01 level (p-values 0.003 and < 0.001 for time and frequency, respectively) (Table 3.1). Although the stable (i.e. intraplate) earthquakes appear to have a higher stress drop distribution in both estimates, the low number (N=17) of stable earthquakes means that the statistical test has low power, which prevents us from making useful inferences. The same limitation applies to volcanic earthquakes.

Although the tectonic setting patterns are consistent between the two datasets, the fact that the fault geometry and lithosphere type trends differ so vastly is a cause for concern. Such differences suggest that reported tectonic stress drop patterns may depend on the details of the analytical procedures. Multiple analyses of the same earthquakes can produce different patterns. Therefore, stress drop studies should be cautious before treating apparent or reported differences as meaningful.

We observe similar discrepancies for stress drop relative to seismic moment. Figure 3.8 shows seismic moment plotted against the AS09 corner frequency and inverse rupture duration for the 894 earthquakes in our dataset. As expected, corner frequency and

inverse rupture duration, which are both inversely proportional to rupture dimension, decrease with increasing seismic moment. However, the different slopes of the regression lines for the AS09 frequency-domain and the time-domain data suggest that the relationship between seismic moment and stress drop varies between the two methods. For the AS09 frequency-domain data, the regression line has a slope similar to the lines of constant stress drop, indicating that the mean stress drop is relatively constant with seismic moment (Figure 3.8a). The slope of the regression line for the time-domain data (Figure 3.8b) differs from the lines of constant stress drop, indicating that mean stress drop increases somewhat with increasing seismic moment. Once again, two analyses of the same earthquakes produce different trends.

### 3.8. Discussion

Our results suggest that stress drop estimates are quite sensitive to the details of the analytical procedure. Deconvolution methods used to remove propagation effects appear likely to be a primary cause of the large discrepancies between the two studies. One key difference is how each method accounts for seismic wave attenuation. The application of an attenuation model seeks to correct for anelastic effects and thus more accurately represent the source signal. For the SCARDEC STFs, Vallée and Douet (2016) apply the same global, frequency-dependent attenuation model to all earthquakes. In contrast, the AS09 deconvolution implicitly corrects for both path and seismic station site attenuation. It also contains additional source-specific attenuation corrections based on a previous upper mantle attenuation study (Warren and Shearer, 2002). Therefore, in the AS09 analysis, the applied attenuation model varies between earthquakes.

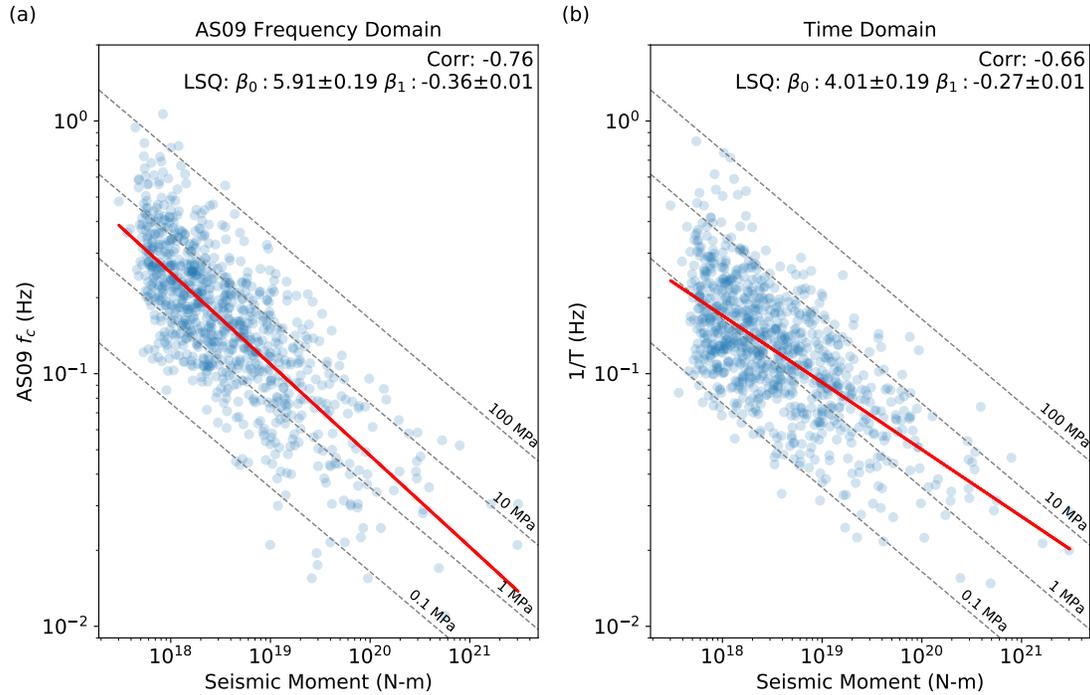


Figure 3.8. Corner frequency and inverse rupture duration as a function of seismic moment. (a) AS09 frequency-domain data. (b) Time-domain data. Circles indicate earthquakes. Dashed lines are lines of constant stress drop. Solid line indicates best-fitting regression line with intercept ( $\beta_0$ ) and slope ( $\beta_1$ ) with 1 standard deviation and correlation indicated. Differing regression line slopes suggest that the inferred relationship between seismic moment and stress drop varies depending on details of analytical procedure.

Small variations in an attenuation model can significantly impact stress drop estimates (Allmann and Shearer, 2009). Deconvolution models that assume higher attenuation yield higher frequency content and thus higher stress drops. If the AS09 source-specific adjustments accounted for most of the discrepancy between the time-domain and the AS09 frequency-domain stress drop estimates, we would expect a clear spatial correlation between the relative stress drop ratios (Figure 3.6c) and the source specific attenuation adjustments made in AS09 (see Figure 3 in Allmann and Shearer, 2009). Earthquakes in

the AS09 dataset that have larger attenuation adjustments would produce higher stress drop values relative to the time-domain results and thus larger stress drop ratios. However, we do not observe any clear correlations between the stress drop ratios (Figure 3.6c) and Allmann and Shearer’s (2009) attenuation adjustments. This suggests that the earthquake-specific attenuation adjustments are not the sole cause of the stress drop estimate discrepancies. The sensitivity of stress drop to attenuation corrections, however, suggests that the attenuation model differences may still play a role in the observed stress drop discrepancies.

The different deconvolution methods may have disparate impacts on different frequencies. Because rupture duration estimates depend largely on the long period (low frequency) components of the source time function and corner frequency estimates depend largely on the short period (high frequency) components, the uncertainty introduced by deconvolution may impact one method more than the other. These frequency-specific effects will likely increase with the distance the waves travel, so stress drops inferred from studies using seismograms at teleseismic distances would have greater uncertainties than ones derived from seismometers closer to the source.

Our results have implications for ongoing debates about the meaning and utility of stress drop in earthquake studies (Atkinson and Beresnev, 1997). Some studies attribute geologic significance to this parameter itself (Kanamori and Anderson, 1975; Abercrombie, 2014), whereas others (Hanks and McGuire, 1981) use it as a proxy for the relative level of high-frequency energy. At least for the datasets we compared, the lack of correlation between stress drop estimates suggests that they might not be useful for either application.

Our analysis shows that large uncertainties in stress drop estimates might lead to erroneous inferences about earthquakes and their hazards. The theoretical link between earthquake stress drop and ground motion is well established (Boore, 1983). Ideally, ground motion prediction equations (GMPE) developed for one region could be applied to another region if adjustments for regional differences in earthquake stress drop (and other source and site processes) are made (Baltay and Hanks, 2014). This would be especially useful when assessing earthquake hazards in regions where existing ground motion data are relatively sparse. However, incorrect assumptions about a region’s median earthquake stress drop level and between-event variability can lead to significant discrepancies between a region’s expected hazard and its true hazard (Baltay et al., 2017).

Other studies have found that assumptions about the earthquake rupture (Kaneko and Shearer, 2015) and variability between seismograms (Kane et al., 2011) can introduce considerable uncertainty into stress drop estimation. Because stress drop estimates contain many sources of uncertainty that are poorly understood, the teleseismic evidence for differences in average stress drop between earthquakes in different regions is weak. Therefore, it may be appropriate in many applications to assume that earthquakes in different regions have approximately the same average stress drop.

## CHAPTER 4

## Why Do Continental Normal Fault Earthquakes Have Smaller Maximum Magnitudes?

### 4.1. Summary

Continental normal fault earthquakes have been reported to have smaller maximum magnitudes ( $M_{max}$ ) than continental earthquakes with other fault geometries. This difference has significant implications for understanding seismic hazards in extensional regions. Using the Global Centroid Moment Tensor (GCMT) catalog, we examine how  $M_{max}$  varies with fault geometry in continental regions, whether these trends are robust, and potential physical reasons for the smaller magnitudes of continental normal fault earthquakes.

We find that the largest continental normal fault earthquakes are in the low  $M_w$  7 range whereas other fault geometries can reach  $M_w \approx 8$ . The continental normal fault earthquake magnitude-frequency distribution has a lower corner magnitude (one possible parameterization of  $M_{max}$ ) than other fault geometries. The observed smaller continental normal fault  $M_{max}$  is not an artifact of classification criteria or catalog length. Probability calculations indicate that the GCMT catalog is long enough to capture differences in  $M_{max}$  due to fault geometry. Additionally, our analysis indicates that neither fault length nor width is limiting the size of continental normal fault earthquakes. Fault complexity can limit rupture extent, but it is likely not the primary reason for the smaller continental normal fault  $M_{max}$ .

Rather, lithosphere yield stress (strength) appears to be the main factor controlling  $M_{max}$ . In extension, lithosphere is weaker, failing at lower yield stresses than in compression. Although this yield stress difference is consistent with smaller continental normal fault earthquakes, it appears inconsistent with the occurrence of large oceanic normal fault earthquakes. However, the largest oceanic normal fault earthquakes occur near subduction zones where the lithosphere is bending. Laboratory studies indicate that bending lithosphere likely has a higher yield stress than lithosphere in pure extension, which may allow for larger oceanic normal fault earthquakes. Therefore, yield stress—rather than fault geometry alone—appears to be the key factor limiting an earthquake’s maximum magnitude.

## 4.2. Introduction

How fault geometry influences an earthquake’s maximum magnitude ( $M_{max}$ ) is important for understanding seismic hazards. In continental regions, it is commonly assumed that the largest normal fault earthquakes are smaller than those of other fault geometries (Jackson and White, 1989). Some of the largest historical continental strike-slip and thrust earthquakes include the 1906  $M_w = 7.9$  San Francisco (Biasi et al., 2013), 1911  $M_w = 8.0$  Chon-Kemin, Kazakhstan (Kulikova and Krüger, 2015), 1920  $M_w = 8.0$  Haiyuan, China (Deng et al., 1984), 1957  $M_w = 8.1$  Gobi-Altai, Mongolia (Okal, 1976), 1990  $M_w = 7.7$  Luzon, Philippines (Velasco et al., 1996), 2002  $M_w = 7.8$  Denali, Alaska (Ekström et al., 2012), and 2008  $M_w = 7.9$  Wenchuan, China earthquakes (Yu et al., 2010). These are much larger than the largest historical normal fault earthquakes, which include the 1887  $M = 7.5$  Sonora, Mexico (Suter, 2015), 1915  $M_w = 7.3$  Pleasant Valley,

Nevada (Wesnousky, 2008), 1954  $M_w = 7.1$  Fairview Peak, Nevada (Doser, 1986), and 1959  $M_w = 7.3$  Hebgen Lake, Montana (Doser, 1985) earthquakes.

The past 100 years of earthquake observations suggest that continental normal fault earthquakes have a smaller  $M_{max}$ . Whether this observation reflects a fundamental limitation on their size, however, remains unresolved. This question has serious ramifications for seismic hazard in extensional regions. The expected hazard of large normal fault systems, like the 370-km-long Wasatch Fault, changes depending on whether the fault ruptures in single or multiple segments (DuRoss et al., 2016). Likewise, the expected hazard posed by low angle normal faults, which are widespread in extensional regions (Collettini, 2011) but rarely host large earthquakes (Wernicke, 1995; Axen, 1999), strongly depends on whether very large continental normal fault earthquakes will occur.

Before continuing, we should clarify the meaning of  $M_{max}$ .  $M_{max}$  can either be a “hard” or “soft” cutoff value (Kagan, 2002). Under a “hard”  $M_{max}$  framework, it is assumed that no earthquakes can exceed  $M_{max}$ . However with a “soft”  $M_{max}$ , earthquakes can exceed  $M_{max}$  but they are far less likely to occur than we would expect under the linear Gutenberg-Richter magnitude-frequency relationship. It is helpful to think of these “hard” and “soft”  $M_{max}$  differences in terms of the Gutenberg-Richter earthquake magnitude-frequency relationship. For a “hard”  $M_{max}$ , the Gutenberg-Richter curve abruptly terminates at  $M_{max}$ , above which no earthquakes are predicted. For a “soft”  $M_{max}$ , larger magnitude earthquakes are possible, but their frequency is significantly lower than predicted by the unrestricted Gutenberg-Richter curve. As Kagan (2002) notes, a “soft”  $M_{max}$  better reflects the behavior of dissipative physical systems,

and even if a “hard”  $M_{max}$  exists, the uncertainties in seismic moment estimation mean a “soft”  $M_{max}$  is more appropriate. In this paper,  $M_{max}$  refers to a “soft”  $M_{max}$ .

Here, we examine the observation that continental normal fault earthquakes have a smaller  $M_{max}$  than other continental fault geometries. We explore whether these lower magnitudes are an artifact of how we classify earthquakes or relatively short catalog lengths, and (if the observation proves true) potential physical reasons for the smaller magnitudes.

### 4.3. Data Set and Earthquake Classifications

We use the Global Centroid Moment Tensor (GCMT) catalog (Dziewonski et al., 1981; Ekström et al., 2012) from its inception in 1976 through the end of 2019 to examine how earthquake magnitudes vary with fault geometry. Although the catalog includes earthquakes smaller than moment magnitude ( $M_w$ ) 5, it is only complete down to  $M_w$  5.8 (Kagan, 2003). We group earthquakes into fault geometry classifications - normal, strike-slip, thrust, and oblique – using Frohlich’s (1992) classification based on the plunge of the P, T, and B axes of the earthquake’s moment tensor. In this classification, normal earthquakes have a P-axis plunge greater than  $60^\circ$  and thrust earthquake have a T-axis plunge greater than  $50^\circ$ .

We further divide the earthquakes into six categories by depth and tectonic environment. We classify earthquakes as shallow ( $\leq 40$  km) or deep (40 - 200 km) and either continental, oceanic, or convergent. These tectonic zone classifications are based on lithosphere type and proximity to a convergent plate boundary. For our purposes, convergent

Table 4.1. Numbers of earthquakes  $M_w \geq 5.8$  in the GCMT database classified by tectonic environment and fault geometry.

Tectonic Environment	Normal	Thrust	Strike-Slip	Oblique	Total
Shallow Continental	135	267	321	162	885
Shallow Oceanic	366	235	1251	158	2010
Shallow Subduction	212	2329	443	441	3425

plate boundary earthquakes occur in island arcs, close to ( $< 20$  km) or below the surface of the subducting slab, or in continental lithosphere within 100 km of a convergent plate boundary. Continental earthquakes are those within continental lithosphere that do not meet the convergent criteria. Oceanic earthquakes occur within oceanic lithosphere and do not meet either the convergent or continental criteria. In addition to continental interiors, our definition of continental lithosphere includes continental shelves as well as complex tectonic regions like the Aegean Sea and maritime Southeast Asia.

We classify the earthquakes using a slightly modified version of Matthews et al.’s (2016) island arc and continental polygons. Unlike in the Matthews et al. (2016) dataset, we classify the full Aleutian Islands chain, Japan, and the Okinawa Trough as island arcs. Additionally, we reclassify the Sea of Japan and portions of the Bering Sea as oceanic lithosphere. For convergent boundary earthquakes, we use the USGS’s SeismoTectonic Regime Earthquake Calculator (STREC) to calculate the depth to the subducting slab and Coffin et al.’s (1998) dataset to determine distance to a convergent boundary. Table 4.1 shows the number of shallow earthquakes (the earthquakes of interest here) by fault geometry and tectonic environment with  $M_w \geq 5.8$ .

#### 4.4. Global Distribution of Large Normal Fault Earthquakes

Large normal fault earthquakes occur in various tectonic environments. Figure 4.1 shows all  $M_w \geq 6.5$  normal fault earthquakes shallower than 200 km in the GCMT catalog. Most great ( $M_w \geq 8$ ) normal fault earthquakes occur near subduction zones, some within the subducted plate (Okuwaki, and Yagi, 2017). These larger normal fault earthquakes occur due to bending (flexural) stresses within the subducting plate between the trench and outer rise (Craig et al., 2014). Although spreading ridges have numerous small normal fault earthquakes,  $M_w \geq 6.5$  earthquakes are rare there. The 1983 Chagos Archipelago earthquake is the largest normal fault oceanic earthquake not located near a trench. However, this earthquake may not be as large as it appears to be. Although the GCMT catalog indicates  $M_w = 7.7$ , the U.S Geological Survey's (USGS) National Earthquake Information Center (NEIC) lists a lower value of  $M_w = 7.3$ .

Shallow continental normal fault earthquakes in the GCMT catalog, on the other hand, rarely reach  $M_w > 7$ . The largest normal fault earthquakes occur in extensional plate boundary zones like the Basin and Range province in the western U.S., Italy, eastern Mediterranean, and the East Africa Rift extension zones. Since 1976 (the start of the GCMT catalog), the largest normal fault earthquakes in these regions range from  $M_w$  6.5 to  $M_w$  7.1. Surprisingly, the Baikal Rift in Siberia has no large normal fault earthquakes during this time period. Another conspicuously quiet region is the Basin and Range province in the western U.S. Only one earthquake with  $M_w \geq 6.5$  appears in the GCMT catalog (1983  $M_w$  6.9 Borah Peak, Idaho), although this region hosted large earthquakes with  $M_w \geq 7$  earlier in the 20th century. Figure 4.2, showing the depth versus magnitude of normal fault earthquakes, highlights the lack of large shallow continental normal fault

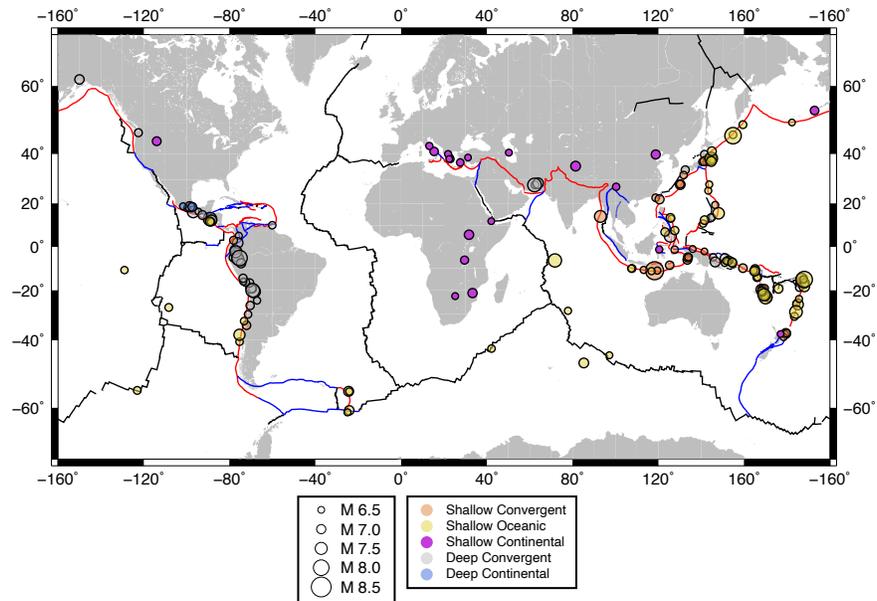


Figure 4.1. Distribution of  $M_w \geq 6.5$  normal fault earthquakes in the GCMT catalog. Circle color corresponds to tectonic environment and depth. Circle size indicates magnitude. Color lines show convergent boundaries (red), spreading centers (black), and transform boundaries (blue) (Coffin et al., 1998). No deep oceanic earthquakes exceed  $M_w$  6.5.

earthquakes despite their prevalence near convergent plate boundaries and in oceanic environments.

#### 4.5. Variations in Earthquake Magnitude Distribution with Fault Geometry

The above observations show that over the past 44 years shallow continental normal fault earthquakes rarely exceed  $M_w$  7, although both shallow oceanic and convergent normal fault earthquakes can be much larger. To better understand how fault geometry impacts magnitude, we compare the magnitude distributions of shallow continental, oceanic,

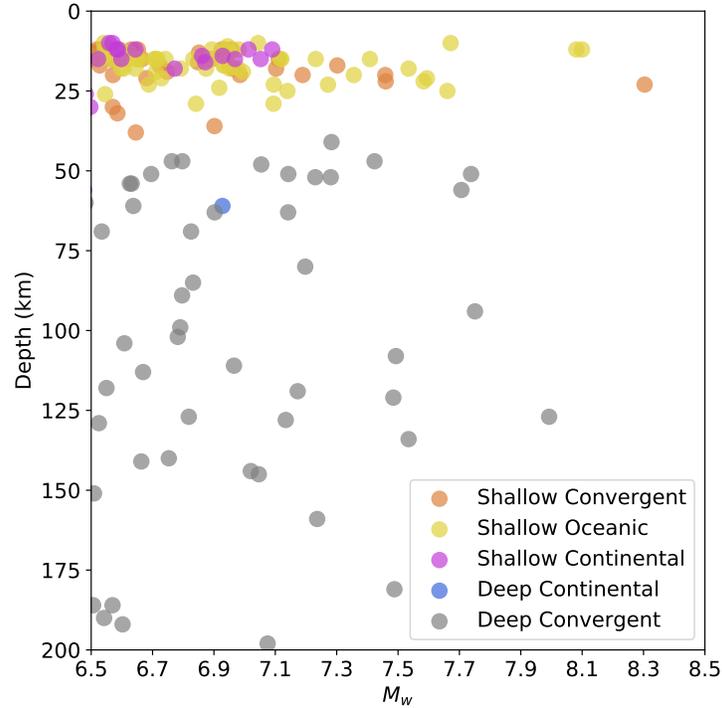


Figure 4.2. Depth versus magnitude plot for normal fault earthquakes, showing depth and tectonic environment classification. No deep oceanic earthquakes have  $M_w > 6.5$ .

and convergent earthquakes (Figure 4.3). For each fault geometry (normal, thrust, strike-slip, oblique), we fit a tapered Gutenberg-Richter (TGR) distribution using the maximum likelihood method (Kagan, 2002). The TGR distribution includes two key parameters:  $b$ -value and corner magnitude ( $M_c$ ). The  $b$ -value corresponds to the slope of the cumulative magnitude distribution line.  $M_c$  is one specific parameterization of the “soft”  $M_{max}$  concept (Kagan, 2002). The inclusion of  $M_c$  causes the distribution for larger magnitudes to deviate below the straight line assumed in a traditional unrestricted Gutenberg-Richter distribution. Although various hypotheses have been proposed for what  $b$ -value (Rundle,

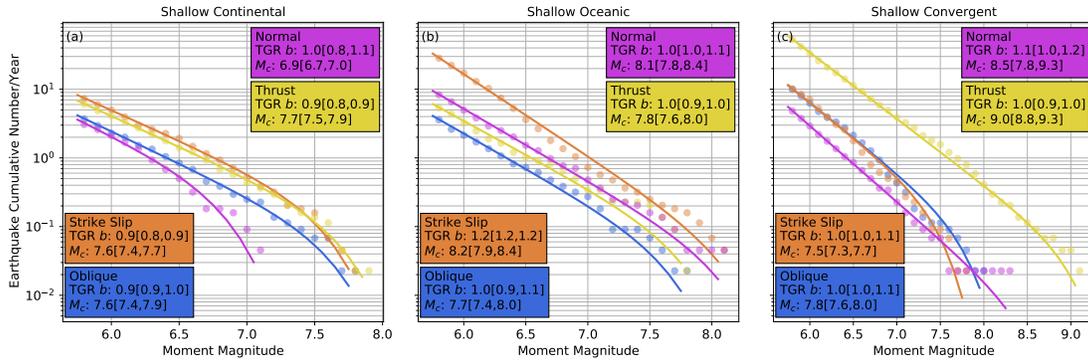


Figure 4.3. Tapered Gutenberg-Richter distributions for shallow continental (a), oceanic (b), and convergent (c) earthquakes. Color corresponds to earthquake fault geometry with  $b$ -values and  $M_c$  listed for each fault geometry. 1-sigma range indicated for each parameter.

1989) and corner magnitude (Okal and Romanowicz, 1994) physically represent, we simply use these values to parameterize the magnitude distribution curve. Previous studies noted that normal fault earthquakes have a higher  $b$ -value (Schorlemmer et al., 2005) and that  $M_c$  trends can differ between tectonic environments (Kagan, 2002).

The greatest differences between normal fault earthquakes and other geometries occur for shallow continental earthquakes. For this environment, normal fault earthquakes have the lowest  $M_c$  (6.9) compared to the other geometries (Figure 4.3a) but a similar  $b$ -value. Thrust, strike-slip, and oblique geometries have an  $M_c$  between 7.6 and 7.7. The more earthquakes in the dataset that exceed  $M_c$ , the better constrained the estimate of  $M_c$  (Kagan, 2002). In the shallow continental environment, the normal fault  $M_c$  is best constrained with seven earthquakes exceeding it (Figure 4.3a). The other fault geometries have at most one or two earthquakes exceeding their  $M_c$ .

Normal fault earthquake distributions do not stand out in shallow oceanic (Figure 4.3b) or shallow convergent (Figure 4.3c) environments. In shallow oceanic environments,

normal fault earthquakes have an  $M_c$  (8.1) indistinguishable from strike-slip (8.2) and thrust (7.8) geometries, however these are all poorly constrained. In the shallow convergent environment, the  $b$ -values are all similar, but thrust earthquakes have the highest  $M_c$  (9.0), reflecting the large megathrust earthquakes along convergent interfaces. Although normal fault earthquakes have the next highest  $M_c$  (8.5), this estimate is poorly constrained. However, the normal fault  $M_c$  for convergent earthquakes appears to be higher than the corresponding  $M_c$  values for strike-slip and oblique earthquakes, although these are also poorly constrained. Thus, normal fault earthquakes have significantly smaller maximum magnitudes compared to other fault geometries only in shallow continental regions.

#### 4.6. Smaller Normal Fault $M_{max}$ Not Due to Classification

Our analysis indicates that in the GCMT catalog, shallow continental normal fault earthquakes have smaller maximum magnitudes. To assess the robustness of this result, we examine how fault geometry classification impacts our assessment of  $M_{max}$ . The hard cutoffs in Frohlich's (1992) classification cause similar earthquakes just above and below the cutoff to be grouped separately. In this classification, normal-fault-like earthquakes with P-axis plunges just below  $60^\circ$  are classified as oblique. If the smaller continental normal fault  $M_{max}$  observation is robust, oblique earthquakes just below the cutoff should have a similar  $M_{max}$  to those above the cutoff.

A plot of magnitude versus P-axis plunge (Figure 4.4a) shows that oblique earthquakes just below the  $60^\circ$  cutoff are no bigger than those above it. Shallow continental earthquakes (Figure 4.4a) show a clear pattern in maximum magnitude as P-axis plunge

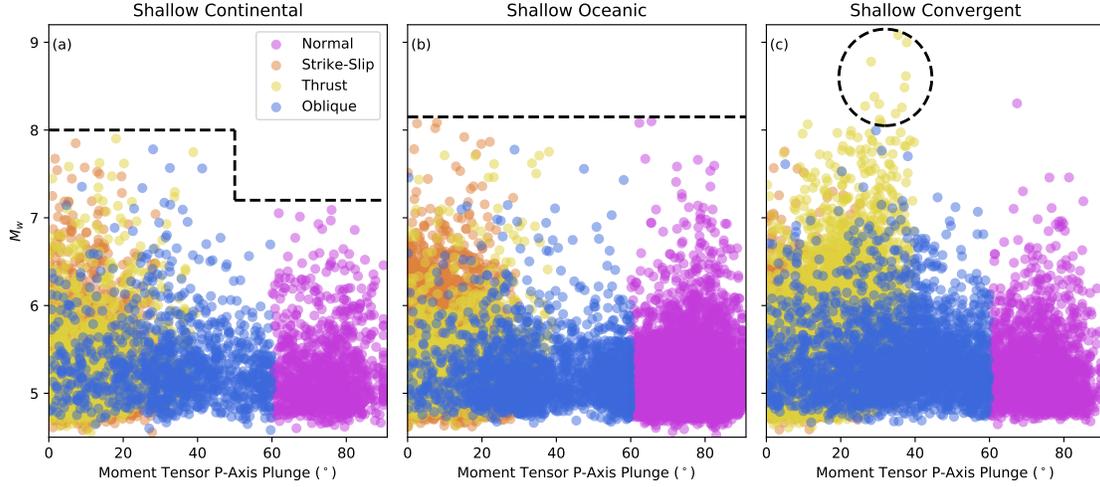


Figure 4.4. Plot of magnitude versus P-axis plunge for shallow continental (a), oceanic (b), and convergent (c) earthquakes. Circle color corresponds to fault geometry classification. In panel a, the dashed line indicates the drop in  $M_w$  with increasing P-axis plunge. In panel b, the dashed line shows constant  $M_w$  with increasing P-axis plunge. In panel c, the dashed circle indicates megathrust earthquakes.

increases from  $0^\circ$  to  $90^\circ$ . Maximum magnitudes approach  $M_w$  8 for strike-slip, thrust, and oblique earthquakes with P-axis plunges less than  $50^\circ$ . Above  $50^\circ$  the largest earthquakes drop to approximately  $M_w$  7 for both normal and oblique geometries. Thus, maximum magnitude drops as the fault geometry becomes more normal-fault like

We do not observe similar sharp drops in maximum magnitude as the P-axis plunge increases for shallow oceanic (Figure 4.4b) or shallow convergent (Figure 4.4c) environments. For shallow oceanic earthquakes, the largest  $M_w$  stays relatively constant as P-axis plunge increases. A different  $M_{max}$  trend occurs for shallow convergent earthquakes. Instead of a constant  $M_{max}$  across the P-axis plunge range, a peak occurs between  $30^\circ$  and  $40^\circ$ , corresponding to megathrust earthquakes at subduction interfaces. Aside from these megathrust earthquakes, the overall trend between  $M_{max}$  and P-axis plunge appears flat.

#### 4.7. Smaller Normal Fault $M_{max}$ Not Due to catalog Length

In the GCMT catalog, shallow continental normal fault earthquakes have a smaller corner magnitude (due to the lack of large earthquakes), but is 44 years long enough to be confident in these differences? We explore this question using probability density functions (PDF) to estimate the likelihood of observing these trends. The tapered Gutenberg-Richter distribution (Figure 4.3) assumes that the magnitude of each shallow continental normal fault earthquake is independent and drawn from the same PDF (Kagan, 2002). Hence, we can calculate the probability that an earthquake falls within a given magnitude range. By assuming that the earthquake magnitudes are independent and identically distributed, we can also calculate the probability that a number of earthquakes all fall within the same magnitude range by taking the probability for one earthquake and raising it to a power equal to the number of earthquakes. We can use the same procedure to estimate the probability that no earthquakes exceed a specified magnitude.

The GCMT catalog contains 135 shallow continental normal fault earthquakes with  $M_w \geq 5.8$ . We calculate the probability that none of 135 such earthquakes would have  $M_w > 7.1$ , the largest in the catalog, for a range of  $b$ -values and  $M_c$ . A high probability indicates the lack of earthquakes with  $M_w > 7.1$  is a likely outcome. A low probability indicates that it is an unlikely outcome. This lets us assess whether our  $M_c$  estimate of 6.9 for shallow continental normal fault earthquakes is reasonable or is artificially low because the GCMT catalog is too short to capture rarer larger events.

The probability estimates show that the catalog is long enough to provide a reasonable estimate of  $M_c$ . In Figure 4.5, the purple lines correspond to the probability that  $M_w > 7.1$  is the largest earthquake for three different  $b$ -values and a range of  $M_c$  values. As

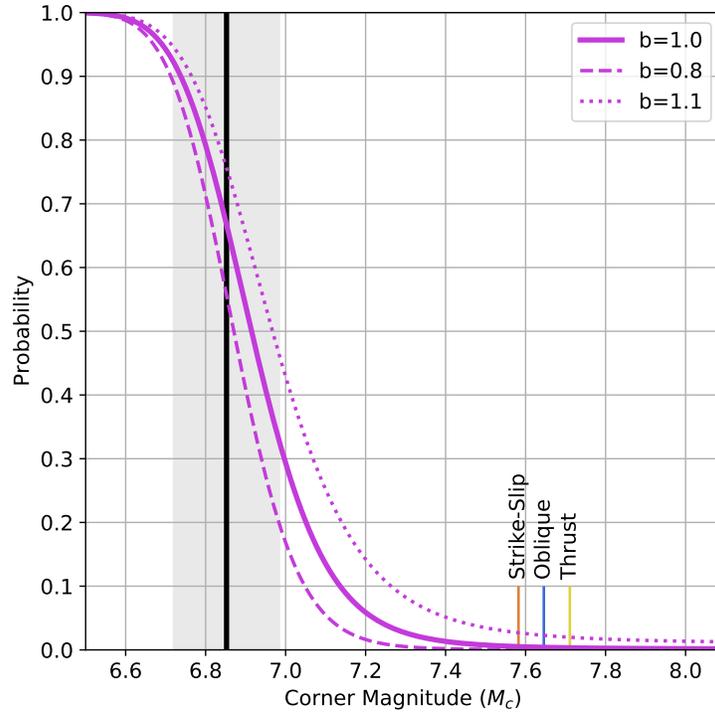


Figure 4.5. Probability that for 135 earthquakes,  $M_w$  7.1 is the largest observed earthquake for a range of  $M_c$  using a tapered Gutenberg-Richter distribution. Purple lines indicate probabilities for different  $b$ -values. Best estimate of shallow, normal continental  $M_c$  (black line) with 1 sigma indicated (grey shading). Corner magnitudes for other fault geometries indicated.

$M_c$  increases, the probability that an  $M_w = 7.1$  earthquake is the largest in the catalog decreases. For our best estimate of  $b$ -value (1.0) and  $M_c$  (6.9) there is an approximately 70% chance that an  $M_w = 7.1$  would be the largest earthquake. If we assume a lower  $b$ -value (0.8) and higher  $M_c$  (7.0) based on the parameters' 1-sigma uncertainties, then the probability decreases to 20%. Conversely, for a higher  $b$ -value (1.1) and lower  $M_c$  (6.7), the probability is closer to 95%.

It is highly unlikely that the  $M_c$  for shallow continental normal fault earthquakes is as large as the  $M_c$  for thrust, strike-slip, and oblique earthquakes, which we estimated to be between 7.6 to 7.7. If normal fault earthquakes had an  $M_c$  value to similar to the other fault geometries, there is less than an approximately 5% chance (and possibly even lower for smaller  $b$ -values) that the observed  $M_w = 7.1$  event would be the largest shallow continental normal fault earthquake. These low probabilities indicate that shallow continental normal fault earthquakes likely have a smaller  $M_c$  (and therefore  $M_{max}$ ) than the other fault geometries, and the GCMT catalog is long enough to observe these differences.

#### 4.8. Reconciling GCMT $M_{max}$ with Early Instrumental Earthquakes

Larger magnitude historical earthquakes that pre-date the GCMT catalog have been reported in extensional environments. However, the lack of a global seismographic network makes it difficult to compare historical instrumental earthquake magnitudes to modern instrumental estimates. Accurate  $M_w$  estimates for pre-instrumental earthquakes require well-constrained fault length, width, and slip measurements, but we only directly observe surface rupture length and surface displacement. For instance, the  $M$  7.5 estimate for the 1887 Sonora Earthquake is based on extrapolating Wells and Coppersmith's (1994) empirical scaling relations between surface rupture length and  $M_w$  (Suter, 2015). For early instrumental-era earthquakes with few available seismograms like the 1915 Pleasant Valley, Nevada earthquake, discrepancies exist between the  $M_w = 7.3$  geologic  $M_w$  estimate (Wesnousky, 2008) and instrumental  $M_w = 6.9$  to 7.0 estimates (Doser, 1988).

These historical large extensional environment earthquakes may also contain significant oblique motion (Doser and Yarwood, 1990) resulting in an oblique rather than

normal fault classification using the Frohlich (1992) criteria. For instance, the 1910  $M_w = 7.4$  Rukwa earthquake, the largest instrumentally recorded earthquake in East Africa (Ambraseys, 1991), appears to have a significant strike-slip component (Ayele and Kulhánek, 2000). The 1956  $M_w = 7.7$  Amorgos, Greece earthquake—the largest Aegean Sea earthquake over the past 100 years—may also contain significant oblique motion because some studies find predominantly normal faulting (Okal et al., 2009) whereas others indicate it was primarily strike-slip (Papazachos & Delibassis, 1969; Ritsema, 1974). In the Baikal rift region, the largest instrumentally recorded earthquake, the complex 1957  $M_w = 7.8$  Musik earthquake, involved primarily strike-slip faulting (Doser, 1991). The 1959  $M_w = 7.3$  Hebgen Lake, Montana earthquake is another large historical event in an extensional environment with a somewhat complicated geometry. Doser’s (1985) body-wave inversion suggests that this earthquake is best described by two subevents, the larger of which contains oblique motion, however the surface rupture indicates predominantly normal faulting displacement (Johnson et al., 2018).

How do we reconcile the absence of any shallow, normal fault earthquakes with an  $M_w > 7.1$  in the GCMT catalog with the presence of larger continental normal fault earthquakes in the pre-GCMT instrumental record? Although there is some uncertainty in the moment magnitude and geometry of these early instrumental earthquakes, the 1959 Hebgen Lake, Montana and possibly the 1915 Pleasant Valley, Nevada are  $M_w = 7.3$  earthquakes that exceed the largest GCMT shallow continental normal fault earthquakes. However, the occurrence of these larger magnitude earthquakes is expected given the longer observation window.

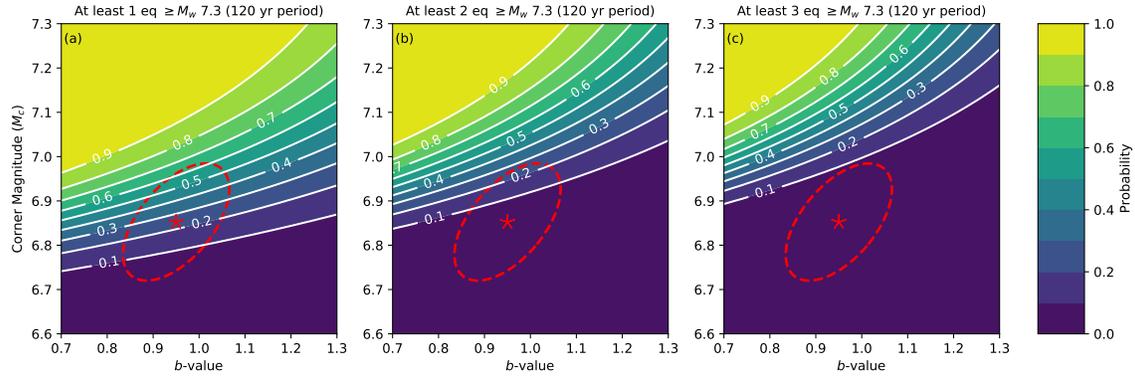


Figure 4.6. Probability of observing at least 1 (a), 2 (b), or 3 (c) shallow continental normal fault earthquakes  $\geq M_w$  7.3 for a given combination of  $b$ -value and  $M_c$  over a 120-year period. Colors and contours indicate the probability. The red star and dashed red ellipse indicate the best fitting  $b$ -value and  $M_c$  and 2-sigma uncertainty ellipse estimated from the GCMT shallow continental normal fault earthquakes.

If we assume the annual rate of shallow, continental, normal fault earthquakes with  $M_w \geq 5.8$  is constant (approximately 3.07 per year based on the GCMT catalog), then we can calculate the probability of observing these larger earthquakes over the length of the historical instrumental record. Figure 4.6 shows the probability of at least one (Figure 4.6a), two (Figure 4.6b), and three (Figure 4.6c) shallow continental normal fault earthquakes with  $M_w \geq 7.3$  occurring in a 120-year span—the approximate length of the historical instrumental catalog (Di Giacomo et al., 2015)—for a range of  $M_c$  and  $b$ -values. The red star and dashed red ellipse indicate the best fitting  $b$ -value and  $M_c$  and 2-sigma uncertainty ellipse estimated from the GCMT shallow continental normal fault earthquakes.

The probability estimates based on GCMT  $b$ -value and  $M_c$  parameters seem to be roughly in line with what we observe in the historical instrumental record, but the GCMT-based  $M_c$  estimate might be slightly low. The best fitting  $b$ -value and  $M_c$  parameters

indicate that there is a 20% to 30% chance of at least one shallow continental normal fault earthquake with  $M_w \geq 7.3$  (1959 Hebgen Lake) in the historical instrumental record (Figure 4.6a). If we consider the upper end of the 2-sigma uncertainty ellipse, there is an approximately 60% chance. The probability of at least two earthquakes with  $M_w \geq 7.3$  (1915 Pleasant Valley and 1959 Hebgen Lake) occurring is significantly lower (Figure 4.6b). The best fitting  $b$ -value and  $M_c$  parameters indicate less than a 10% chance of at least two earthquakes with  $M_w \geq 7.3$  occurring, but this rises to at least 20% if we consider the upper end of the 2-sigma uncertainty. Although 20% may seem low, it is still a higher probability than rolling a specific number on a 6-sided die. More importantly, due to the tight clustering of the probability contours, a small increase in  $M_c$  to 7.0 or 7.1 increases the probability to between 50 and 90%. In the scenario of at least three earthquakes with  $M_w \geq 7.3$  (Figure 4.6c), only a small increase in  $M_c$  above the GCMT-based estimate is needed to increase the probability to 50%. Thus an  $M_c$  near 7.0 could account for the observed shallow continental normal fault earthquake magnitude distribution over the entire historical instrumental record. A normal fault  $M_c$  of 7.0 would still be significantly lower than observed  $M_c$ 's (7.6-7.7) for thrust, strike-slip, and oblique earthquakes, indicating significant magnitude distribution differences due to fault geometry.

#### **4.9. Fault Dimensions Not a Limitation on Normal Fault Earthquake Magnitude**

Understanding why shallow continental normal fault earthquakes have a smaller  $M_{max}$  has important implications for seismic hazard analysis. Because larger earthquakes require

longer faults, fault length may limit the size of normal fault earthquakes. We examine this possibility by comparing the global distribution of extensional faults from the Global Earthquake Model (GEM) Global Active Fault Database (Styron and Pagani, 2020) to the expected surface rupture lengths for normal fault earthquakes (Wells and Coppersmith, 1994).

The fault length histogram (Figure 4.7) shows that there are at least 890 normal faults long enough to host earthquakes with  $M_w \geq 7.0$ , 547 faults long enough for earthquakes with  $M_w \geq 7.2$ , 223 faults long enough for earthquakes with  $M_w \geq 7.5$ , and 22 faults long enough for earthquakes  $M_w \geq 8.0$ , especially considering the variability in earthquake rupture length for a given magnitude. Because Wells and Coppersmith's (1994) regression only includes earthquakes with magnitudes from 5.2 to 7.3, it is possible that larger earthquakes do not follow the same relation between fault length and magnitude. However, larger earthquakes generally follow similar trends as smaller earthquakes (Figure 4.8), but earthquakes with similar magnitudes can have very different surface rupture lengths. For  $M_w < 7.5$ , normal fault rupture lengths do not appear to differ significantly from the other two fault geometries, so there is little reason to suggest that larger earthquakes would behave differently.

Although the fault length data suggest that some continental normal faults are long enough for larger magnitude earthquakes, the fault database does not indicate how continuous these large faults are. A comprehensive study of normal faults in the Afar region of the East Africa Rift showed that nearly all the faults showed some segmentation regardless of fault length (Manighetti et al., 2015). Fault segmentation is also clearly visible along the Wasatch Fault Zone in Utah and the Fucino Fault Zone in Italy (DuRoss, 2016).

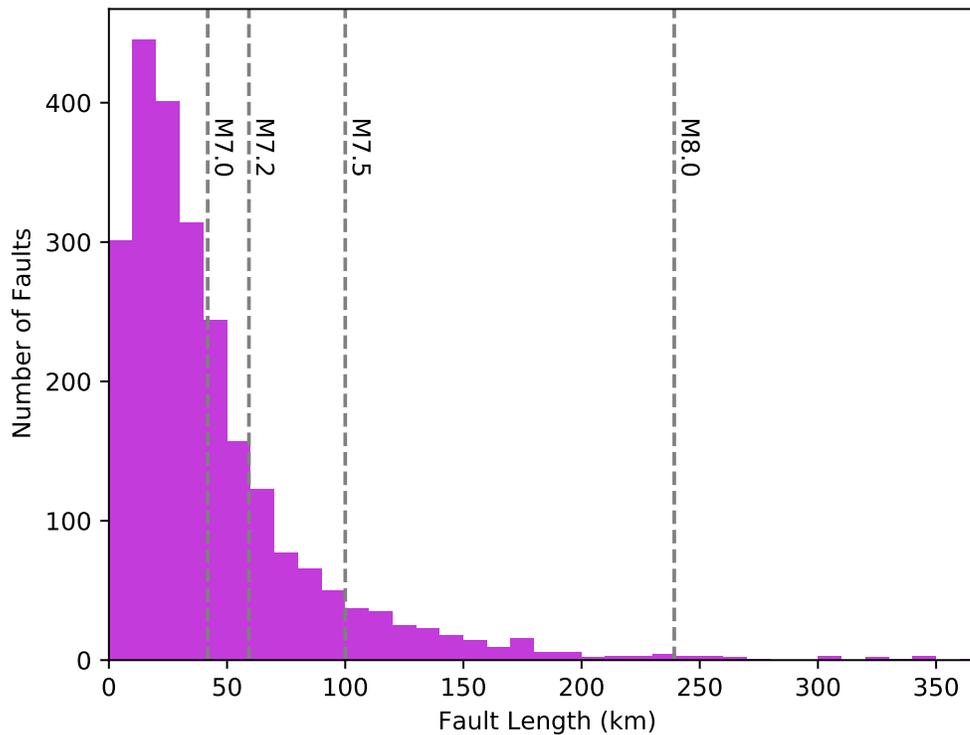


Figure 4.7. Histogram of active continental normal fault lengths from the GEM-Active Fault database. Vertical lines indicate expected magnitude for the surface rupture length shown (Wells and Coppersmith, 1994), and extrapolated at higher magnitudes.

Fault segments are delineated by fault gaps, fault branches, fault steps, or changes in fault strike, and it is thought that these complexities can limit rupture extent and hence earthquake magnitude (Wesnousky, 1988) but not always (Ando, 1975).

Detailed paleoseismic and historical rupture studies of normal fault earthquakes suggest that fault complexities sometimes, but not always, control rupture extent. A paleoseismic study (DuRoss et al., 2016) found that, along the central portion of the Wasatch Fault Zone, the most recent earthquakes ( $< 3$  ka) appear to be confined to individual

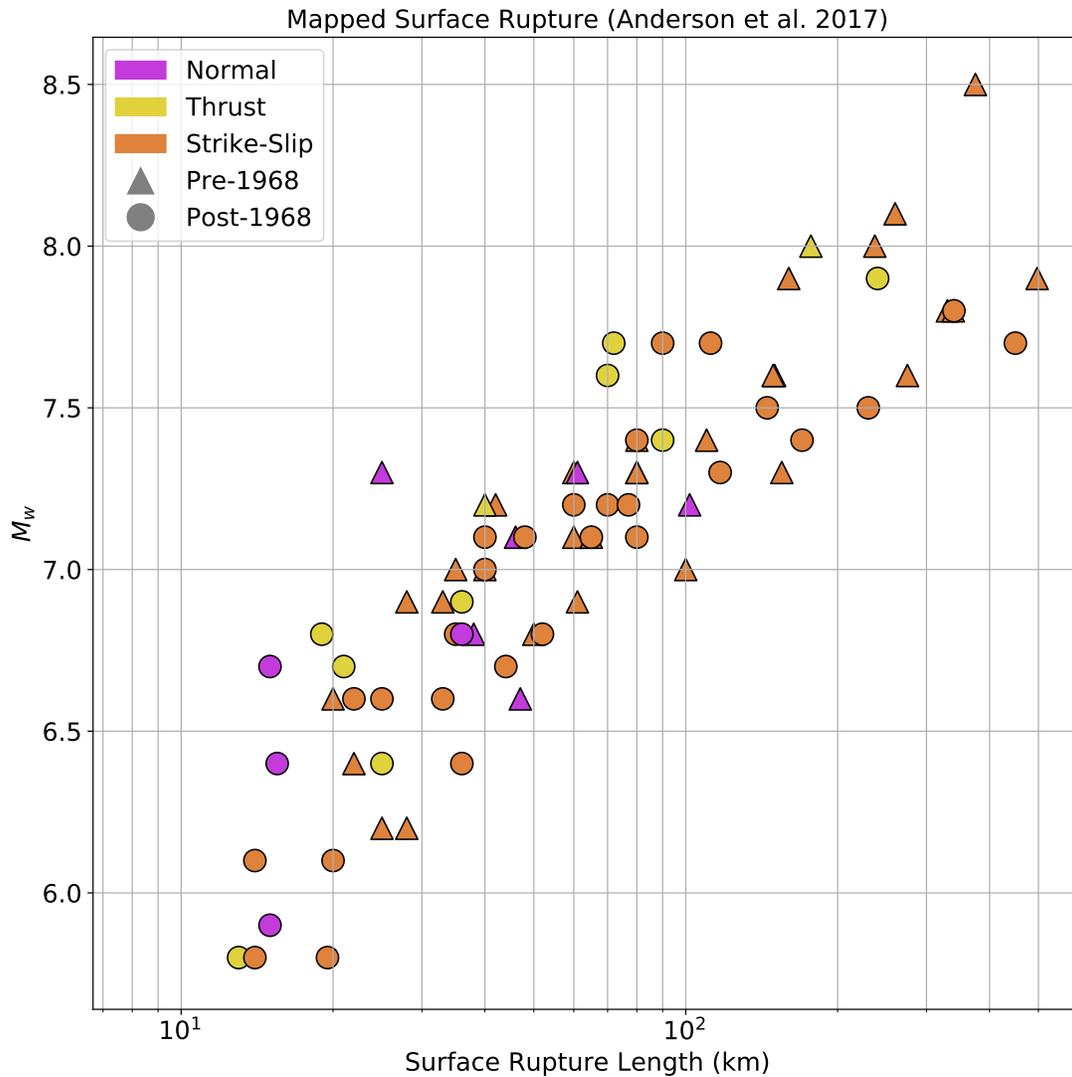


Figure 4.8. Mapped surface rupture length versus magnitude for large historical earthquakes. Data compiled by Anderson et al. (2017). Color corresponds to geometry. Shape corresponds to earthquake date.

fault segments. However, older earthquakes may have been multi-segment ruptures. Historical earthquake fault rupture studies also suggest a range of behavior. Jackson and

White (1989) observed that the largest normal fault ruptures consist of multiple, disjointed segments. Similarly, DuRoss et al. (2016) noted that while some of the largest historical Basin and Range province normal fault earthquakes appear to be limited to a single fault segment, others overcame fault complexity and ruptured as least parts of multiple segments.

If fault complexity were the primary reason for smaller normal fault earthquakes, we would expect to see a stronger spatial correlation between fault rupture extent and fault complexity for normal fault earthquakes compared to other fault geometries. However, this does not appear to be the case. In a study of historical earthquake fault step size—the perpendicular distance between two distinct fault traces—Wesnousky (2008) noted that both strike-slip and normal fault earthquake rupture end points correspond to fault steps approximately 70% of the time. However, normal fault earthquakes can jump larger fault steps (5 to 7 km) than strike-slip earthquakes (3 to 4 km). In a larger study, Biasi and Wesnousky (2016) also observed that normal and thrust fault earthquakes can propagate across larger fault steps than strike-slip earthquakes. They also observed that for similar length ruptures, dip-slip earthquakes (normal and thrust) include more gaps—the absence of surface rupture along an assumed continuous fault trace—than strike-slip earthquakes. Between 60% to 70% of the studied earthquake ruptures end at either a fault step or fault end, with strike-slip earthquakes more likely to end at a fault step while dip-slip earthquakes end at the fault end. For both strike-slip and dip-slip ruptures, in 30% to 40% of earthquakes the rupture ends but the fault trace continues (Biasi and Wesnousky, 2016).

Fault bends—changes in fault strike—are also thought to limit rupture extent. However, Biasi and Wesnousky (2017) found that fault bends at dip-slip rupture ends are no larger than bends within the rupture. In contrast, the ends of strike-slip ruptures corresponded to larger bends than those within the ruptures. These results suggest that strike-slip earthquake rupture extent is more sensitive to fault bends than dip-slip earthquakes. These multiple, detailed fault rupture studies show that fault complexity limits earthquake rupture, but not that normal fault earthquakes are more sensitive to this complexity. In many cases, normal fault ruptures seem to overcome more fault complexity than strike-slip earthquakes.

Fault width, the down-dip fault extent, also impacts earthquake magnitude. As magnitude increases, so does fault width (Wells and Coppersmith, 1994). However, width is limited by the depth of the seismogenic zone (Sibson, 1986). For a given seismogenic zone depth, steeply dipping faults will have smaller widths than shallowly dipping faults. Along the San Andreas Fault, which is capable of hosting  $M_w \approx 8.0$  earthquakes, estimates of the seismogenic zone thickness are approximately 15 km (Nazareth and Hauksson, 2004), so large earthquakes can occur even for relatively thin seismogenic zones. In extensional environments, seismogenic zones are on average 10 to 15 km thick (Jackson and Blenkinsop, 1993), although some zones like the Baikal Rift have seismogenic zones more than 30 km thick (Déverchère et al., 2001). Extensional seismogenic zones are thick enough to host larger magnitude earthquakes, especially considering that most normal fault earthquakes have dips between  $30^\circ$  to  $60^\circ$  (Jackson and White, 1989; Collettini and Sibson, 2001), increasing the potential fault area within the seismogenic zone. Therefore, neither

fault length nor fault width appears to limit the size of shallow continental normal fault earthquakes.

#### 4.10. Lithosphere Yield Stress Controls $M_{max}$

If fault length and width are not limiting the size of continental normal fault earthquakes, then what is? Some argue that the primary energy source driving faulting differs between normal fault earthquakes and other fault geometries and may impact  $M_{max}$  (Doglioni et al., 2015; Bignami et al., 2020). Others argue that continental lithosphere may be too weak to host large normal fault earthquakes (Jackson and White, 1989). Lithosphere is weaker in extension than in compression (Sibson, 1977). Lithosphere yield-stress envelopes (also termed strength envelopes) (Figure 4.9), showing the difference between the most compressive and least compressive principal stress axes required to induce failure, illustrate that the lithosphere fails at lower stress differentials in extension than in compression. Previous research also indicates a link between earthquake magnitude and the lithosphere stress differential (Scholz, 2015). In laboratory experiments, Scholz (1968) observed that lower stress differentials produce larger  $b$ -values (relatively few large events) and proposed that large magnitude events occur when multiple high stress asperities are linked in a rupture. Schorlemmer (2015) and Petrucci et al. (2019) also observed larger  $b$ -values for normal fault earthquakes and attributed them to smaller stress differentials in extensional environments.

For the yield stress envelopes in Figure 4.9, we use a modified equation 22a from Burov (2011)

$$\Delta\sigma_{max}(z) = \text{sign}(\epsilon) \min(|\sigma^b|, |\sigma^d|)$$

where  $\sigma^b$  and  $\sigma^d$  are the maximum brittle and ductile yielding stresses,  $sign(\epsilon)$  is a sign function equal to -1 for compression and 1 for extensions, and  $z$  is depth.

We use equation 12 from Burov (2011) for brittle deformation

$$\sigma^b = \alpha \rho g z (1 - \lambda)$$

where  $\alpha$  is a fault geometry dependent parameter,  $\rho$  is the rock density,  $g$  is the acceleration due to gravity, and  $\lambda$  is the ratio of water to rock density. We use dislocation creep (modified from equation 15 in Burov, 2011) for ductile deformation.

$$\sigma^d = \left( \frac{\dot{\epsilon}}{A f_w \Delta \exp(-Q/(RT))} \right)^{(1/n)}$$

where  $A$  and  $n$  are material constants,  $Q$  is activation energy,  $\dot{\epsilon}$  is strain rate,  $f_w$  is the water fugacity factor,  $R$  is Boltzmann gas constant, and  $T$  is temperature. For oceanic lithosphere we assume a dry olivine rheology. For continental lithosphere, we assume a two-layer rheology with a wet quartz crust and wet olivine mantle (Kohlstedt et al., 1995). We use a half-space cooling model to calculate the oceanic lithosphere and a steady-state model with shallow radioactivity for continental lithosphere—see equation 4.31 in Turcotte and Schubert (2014). Parameters used in these calculations are listed in Tables 4.2 and 4.3.

Because lithosphere is weaker in extension for both oceanic (Figure 4.9a) and continental (Figure 4.9b) lithosphere, we would expect normal fault earthquakes to have a smaller  $M_{max}$  in both shallow continental and oceanic environments. The lithosphere yield-stress argument is compelling except that oceanic earthquakes do not appear to have the same  $M_{max}$  pattern as continental earthquakes. However, a closer examination

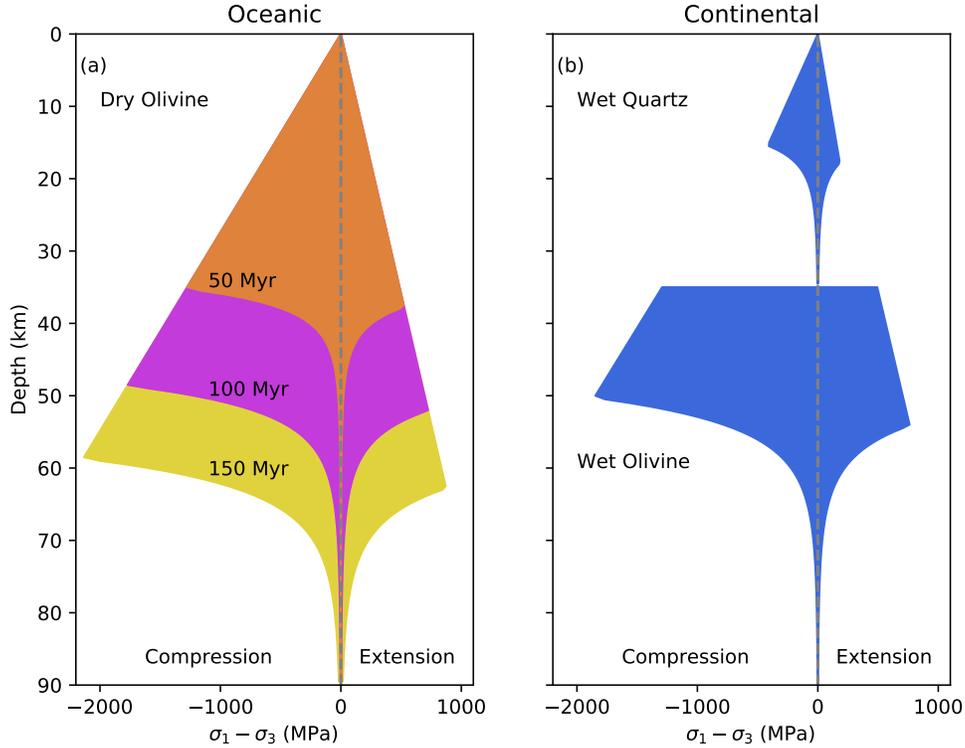


Figure 4.9. Yield-stress envelope (YSE) for oceanic (a) and continental (b) lithosphere in compression and extension. The oceanic YSE (a) assumes a dry olivine rheology and half-space lithosphere cooling model for 50, 100, and 150 million year-old (Myr) lithosphere. The continental YSE (b) assumes a steady-state geotherm with near-surface radioactivity. An upper wet quartz and lower wet olivine rheology are assumed. Yield stresses are higher in compression than in extension.

of shallow oceanic normal fault earthquakes suggests that their  $M_{max}$  may actually be similar to their continental brethren.

As shown in Figure 4.1, the largest oceanic normal fault earthquakes occur between the trench and outer rise due to bending stresses in the subducting plate (Craig et al., 2014). In the GCMT catalog, these outer rise events reach  $M_w \approx 8$ , and larger ones, including the

Table 4.2. Lithology parameters used in yield stress envelope calculations. Values from Burov (2011).

Mineral/Rock	$A$ (MPa $^{-n}$ s $^{-1}$ )	$n$	$Q$ (kJ mol $^{-1}$ )	Density (kg/m $^3$ )
Wet Quartz	1e-4	2.4	160	2700
Wet Olivine	4.876e6	3.5	515	3300
Dry Olivine	1e4	3	520	3300

Table 4.3. Earth structure and geotherm calculation parameters.

Coefficient of Friction	0.5
Oceanic thermal diffusivity	0.804e-6 m $^2$ s $^{-1}$
Oceanic base temperature	1330 °C
Oceanic surface temperature	0 °C
Continental surface temperature	10 °C
Continental mantle heat flow	30 mWm $^{-2}$
Length scale of radioactivity	10 km
Continental surface heat flow	56.5 mWm $^{-2}$
Continental thermal conductivity	3.35 Wm $^{-1}$ °C $^{-1}$
Continental crustal thickness	35 km

1933  $M_w = 8.6$  Sanriku earthquake off the coast of Japan (Kanamori, 1971), have been observed. If we remove outer rise earthquakes from the dataset,  $M_{max}$  drops as the P-axis plunge increases, as observed for continental earthquakes (Figure 4.10). The outlier is the 1983 Chagos earthquake, with a significant discrepancy between GCMT ( $M_w = 7.7$ )

and USGS NEIC ( $M_w = 7.3$ ) magnitude estimates. If the USGS NEIC  $M_w$  better reflects the true value, then shallow oceanic earthquakes away from trenches have  $M_{max}$  in the low 7 range. This event occurred in a tectonically complex region and may not represent genuine oceanic lithosphere away from plate boundaries (Wiens & Stein, 1984).

Why might normal faulting earthquakes in flexural regions between the trench and outer rise have a higher  $M_{max}$  than those in extensional regions far from the trench? Perhaps this occurs because the stress fields due to plate bending differ from those due to pure extension. For a homogenous material, pure extension produces uniform extensional stress within the material (Figure 4.11). Bending, however, produces extensional stress that is highest at the top of the material and decreases until the neutral plane where no extensional stresses occur (Turcotte and Schubert, 2014). Below the neutral plane, the material is in compression that increases with depth, producing thrust fault earthquakes (Craig et al., 2014).

The different stress fields may cause different failure behavior. Homogenous material fails at the same yield stress in both pure extension and bending. However, experiments (Campo, 2008; Whitney and Knight, 1980) show that materials are generally stronger in bending than in pure extension. Small defects in materials cause this strength discrepancy (Leguillon et al., 2015). In pure extension with a uniform stress field, when stress exceeds the yield stress of the weakest point, the material fails. In bending, however, stress in parts of the material may exceed the yield stress of the weakest point. Unless the weak point is near the outer surface, it will not experience the highest bending stresses. Instead, other regions will continue to experience increasing stress levels until the weak zone's yield stress is reached. When it fails, other parts of the material are actually at higher stresses.

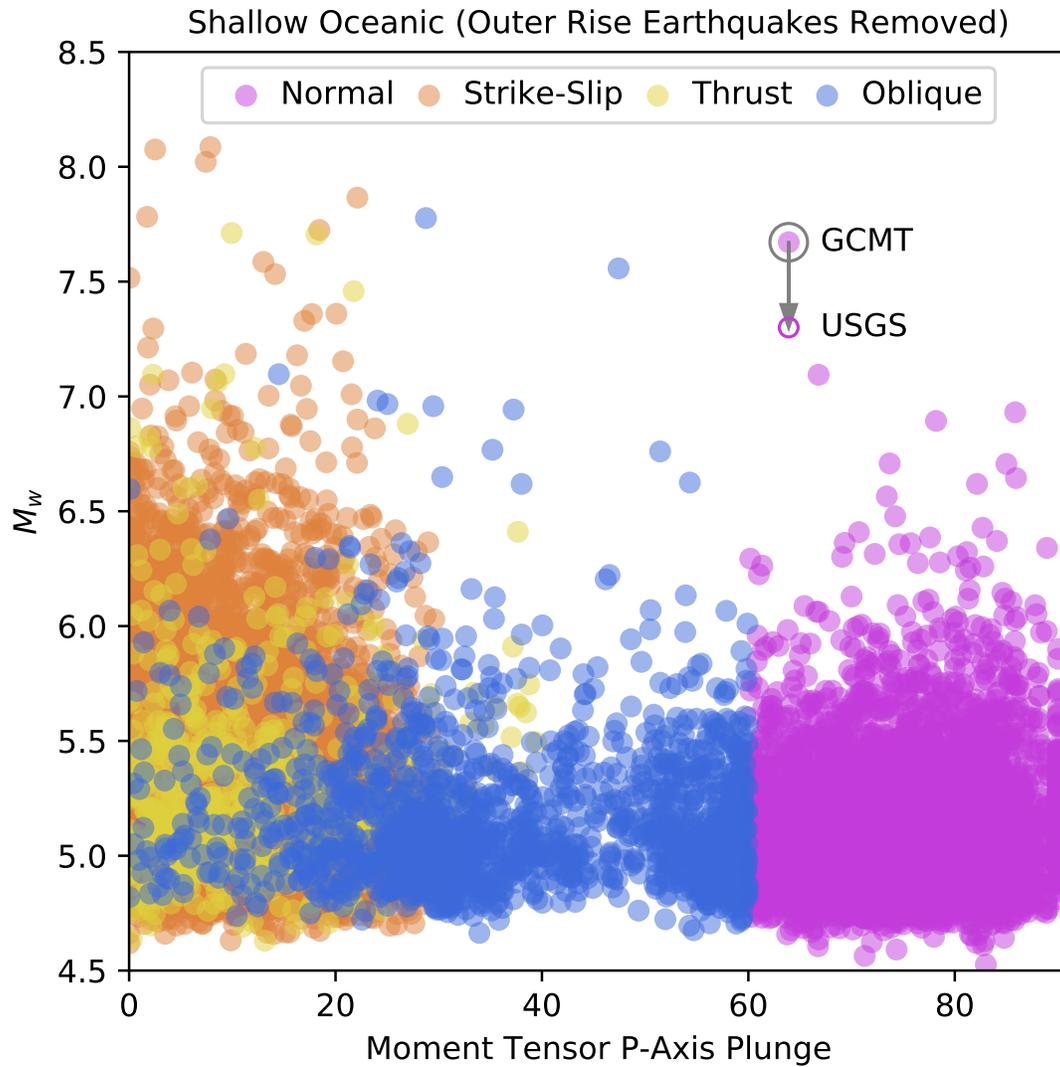


Figure 4.10. P-axis plunge versus moment magnitude for oceanic earthquakes with outer rise earthquakes (those earthquakes within 100 km of the trench) removed. GCMT and USGS NEIC moment magnitudes for the 1983 Chagos earthquake are indicated. With outer rise earthquakes removed,  $M_{max}$  drops as P-axis plunge increases.

Although the real stress state within lithosphere is complicated (Buck, 1991; Craig et al., 2014), these simple models suggest that varying yield stress may explain normal fault

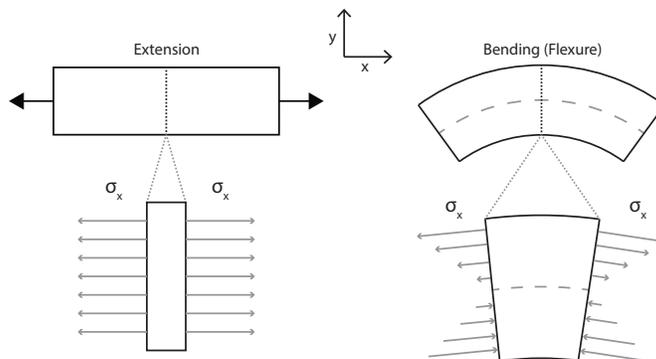


Figure 4.11. Schematic plot of stress within homogenous material for extension and bending. In extension, the internal stress is uniform whereas in bending the largest extensional and compressional stresses are at the outermost points.

earthquake magnitude. In bending oceanic lithosphere, when failure occurs, the additional areas of high stress may allow earthquakes to grow large. However, in continental lithosphere under pure extension, the yield stresses may be too low to allow large magnitude earthquakes. This reasoning follows Scholz's (1968) hypothesis that large earthquakes occur from linking several high stress asperities.

Therefore, the mode of deformation's impact on yield stress appears to be important for maximum earthquake magnitude. The weakness of lithosphere in extension appears to prevent shallow continental normal fault earthquakes from growing as large as those for other fault geometries. However, bending oceanic lithosphere's ability to produce large normal fault earthquakes indicates that fault geometry alone is an insufficient predictor of  $M_{max}$ . Understanding the lithosphere's stress state and deformation mode is thus critical in assessing a region's seismic hazard.

#### 4.11. Conclusion

Our analysis shows that shallow continental normal fault earthquakes have a smaller maximum magnitude (in the low  $M_w$  7 range) than other fault geometries ( $M_w \approx 8$ ). This maximum magnitude difference appears to be real and not an artifact of catalog length or earthquake classification. Although fault length, width, and complexity can impact the extent of an earthquake's rupture, these do not appear to be the primary reason for the smaller maximum magnitudes of shallow continental normal fault earthquakes. Instead, we propose that the weakness of lithosphere in extension is what primarily limits the size of normal fault earthquakes. The smaller maximum magnitudes of shallow continental normal fault earthquakes have important implications for seismic hazard assessment in extensional tectonic environments. In such environments, normal fault earthquakes are unlikely to exceed a low  $M_w$  7 earthquake, even if the fault system is long enough to host much larger earthquakes.

## CHAPTER 5

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

### 5.1. Summary

Forecasts of the probability of a large earthquake occurring on a fault during a specific time interval assume that a probability distribution describes the inter-event times between large earthquakes. However, current models do not incorporate fundamental aspects of the strain accumulation and release processes that cause earthquakes. 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 of partial strain release in earthquake histories showing clusters and gaps. Here we calculate earthquake probabilities using the Long-Term Fault Memory (LTFM) model, which better reflects the strain accumulation and release processes. Using the southern San Andreas fault as an example, we show that LTFM yields a more realistic earthquake forecast. Whereas current models estimate the earthquake probability will be essentially unchanged in the next 80 years, LTFM predicts that the probability will continue to grow, resulting in a 30-year earthquake probability that is 38% higher than the other models. By allowing partial strain release, LTFM incorporates the specific timing of past earthquakes, which commonly used probability models cannot

do. Thus LTFM better forecasts the exceptionally short inter-event time before the 1857 Fort Tejon earthquake. Although LTFM is more complex than existing models, it is also more powerful.

## 5.2. 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 due to motion between the two sides of a locked fault is released by slip on the fault when an earthquake occurs (Reid, 1910). However, current 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 better reflects 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 that it produces more realistic forecasts.

Table 5.1. Dates of the most recent 10 earthquakes along the Pallett Creek section of the San Andreas fault as listed in Table 4 in Scharer et al. (2011). Note that the earthquake in the year 645 is event C from Biasi et al. (2002)

Date									
645	728	805	957	1102	1181	1339	1508	1812	1857

### 5.3. Limitations of Current Earthquake Probability Methods

The Mojave section of the San Andreas is of concern because large earthquakes occur on average every 135 years, most recently the 1857  $M = 7.9$  Fort Tejon earthquake. Prior to 1857, a  $M7.5$  earthquake occurred in 1812 (an inter-event time of 45 years), which was preceded by a long 304-year quiescent period (Scharer et al., 2011) (Table 5.1). Studies suggest that the probability of a large earthquake here in next 30 years is 20 - 25% (Biasi et al. 2002; Field et al., 2015). How such probabilities are calculated varies between studies but follows a general methodology. Studies start with a paleoseismic record giving dates of large past earthquakes (Figure 5.1A) and fit some probability density function (PDF) to the distribution of inter-event times (Figure 5.1B). Here we show four 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 assume the probability changes with time) lognormal, Brownian Passage Time (BPT), and Long-Term Fault Memory (LTFM) models. From these PDFs, one can calculate the probability of a large earthquake during a time period, typically the next 30 years, given that one has not occurred since 1857 (Figure 5.1C).

The estimated 30-year probabilities depend dramatically on each PDF’s assumption about how the probability of an earthquake changes with time. For the time-independent

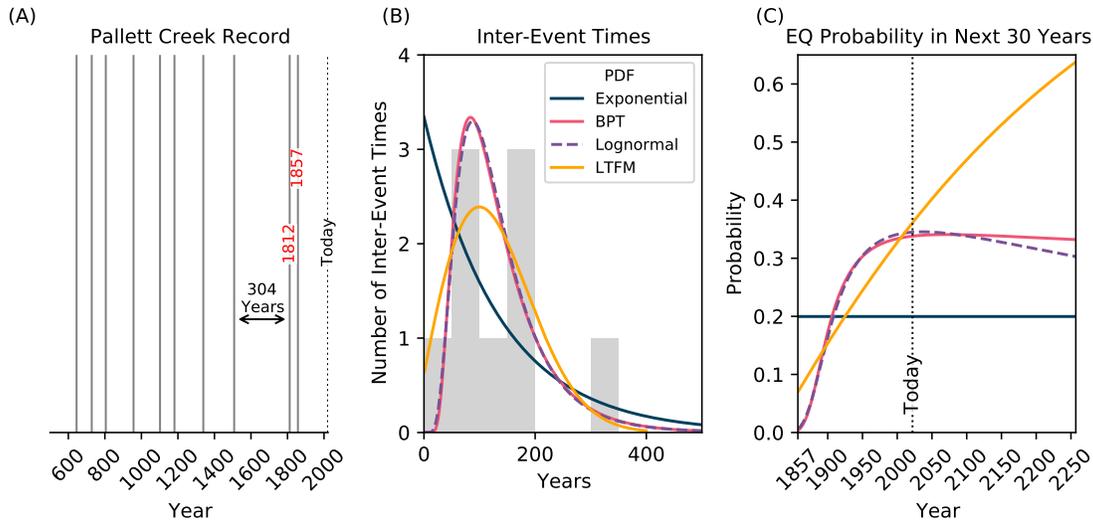


Figure 5.1. Estimating earthquake probabilities. (A) Pallett Creek paleoseismic record of past earthquakes (Scharer et al., 2011, listed in Table 5.1). (B) Histogram of inter-event times (grey bars) and estimated probability distributions for inter-event times for exponential, lognormal, BPT, and LTFM models. The lognormal and BPT models have a coefficient of variation of 0.54. (C) Conditional probability of an earthquake (EQ) in the next 30 years given the date of the last earthquake. Only LTFM yields steadily rising and thus higher earthquake probability, consistent with continuing strain accumulation on the San Andreas fault.

exponential model, the 30-year probability stays constant at 20% (Fig 1C). The 30-year probabilities change with time for the time-dependent lognormal, BPT, and LTFM models. In 2022, the lognormal (34%), BPT (34%), and LTFM (36%) produce similar 30-year estimates but diverge significantly in the future. The lognormal and BPT probabilities have been relatively flat since the expected mean recurrence time (approximately 1990s), whereas LTFM's has steadily increased. Moving forward, if no earthquake occurs by 2050, the lognormal and BPT models give probabilities of 35% and 34%, whereas LTFM predicts 40%. By 2100, the lognormal and BPT forecasts remain stable at 34%, but LTFM's

forecast increases to 47%—a 38% higher chance. By 2150, the LTFM forecast is nearly 60% higher (53% chance) than the lognormal and BPT (33 to 34% chance) forecasts. By 2250, the BPT forecast has leveled out, but the lognormal model continues to decrease.

These time-probability relationships reflect assumptions about accumulated strain and earthquake probability. Intuitively we expect that earthquake probability should increase with time as strain accumulates on a fault, so the exponential model’s flat 30-year probability seems unrealistic. Similarly, the lognormal and BPT’s decreasing and flattening 30-year probabilities after the expected mean recurrence interval seem implausible. The behavior of these two models change if an earthquake has not happened by the mean recurrence time, whereas in LTFM the probability keeps increasing with time.

The difference arises because LTFM is designed to reflect paleoseismic records suggesting that complex patterns of strain accumulation and release influence the timing of earthquakes, often giving rise to clusters of earthquakes and gaps between them (Wallace, 1970; Rockwell et al., 2000; Friedrich et al., 2003; Weldon et al., 2004; Sieh et al., 2008; Goldfinger et al., 2013; Salditch et al., 2020; Hecker et al., 2021). However, unlike LTFM, the common earthquake probability models do not incorporate this effect and its implications. To better understand the accumulated strain/probability relationship, we examine each PDF’s corresponding hazard-rate function. The hazard rate is 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  $f(t)/[1 - F(t)]$ , where  $t$  is the time since the most recent earthquake,  $f(t)$  is the PDF, and  $F(t)$  is the cumulative distribution function (CDF), the integral of the PDF. Fig 5.2A shows four PDFs with the same mean and standard deviation (except for the exponential

PDF whose mean is also the standard deviation), and Fig 5.2B 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 & 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 (Figure 5.2C) 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. Although past studies acknowledged this behavior (Davis et al., 1989, Matthews et al., 2002), no physical mechanism has been proposed for how strain accumulation could lead to variable probability increments. Unlike the other models, in LTFM each additional strain unit corresponds to a constant increase in earthquake probability.

Another limitation of the common time-dependent models is their assumption that the probability resets to zero after each earthquake. They assume that the inter-event times between earthquakes are independent, treating earthquakes as a renewal process

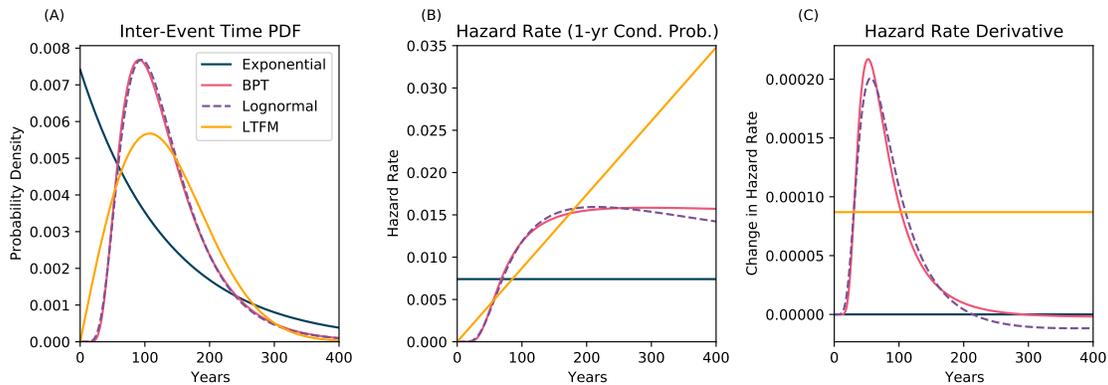


Figure 5.2. Pallett Creek earthquake probabilities. (A) Probability density functions for the inter-event time with same mean (135 years) and standard deviation (70 years)—except exponential where mean equals standard deviation. (B) Corresponding hazard-rate functions. The hazard rate at time  $t$  is the probability 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/decrease/staying constant with time as each new strain increment accumulates.

(Cornell & Winterstein, 1988). Hence the calculated probability of the next earthquake depends only on the time since the most recent earthquake and the known distribution of inter-event times, but not the specific sequence of 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 temporal clusters of earthquakes followed by long gaps, indicating that earthquakes often only partially release the accumulated strain (Wallace, 1970; Rockwell et al., 2000; Friedrich et al., 2003; Weldon et al., 2004; Sieh et al., 2008; Borrero et al., 2009; Goldfinger et al., 2013; Hecker et al., 2021). LTFM overcomes this limitation by allowing partial probability drops, reflecting partial strain releases in

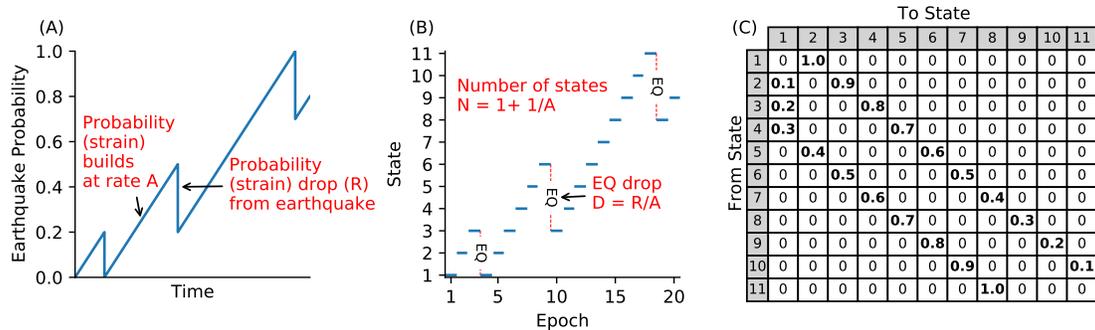


Figure 5.3. Long-term fault memory (LTFM) as a hidden Markov process. (A) Example with 3 earthquakes. (B) Representation of (A) as a Markov model with the probability space discretized into 11 states. (C) Transition probability matrix indicating probability of transition from state  $j$  to state  $k$  for this set of model parameters.

earthquakes. This behavior allows LTFM to incorporate more available information—the specific sequencing of earthquakes—in its probability forecasts.

#### 5.4. Calculating Earthquake Probabilities with Long-Term Fault Memory

The Long-Term Fault Memory (LTFM) model (Figure 5.3A) builds upon previous earthquake probability models, but with modifications to model the temporal patterns of strain accumulation and release. LTFM, like some earlier models (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 inter-event times are not independent.

LTFM has two basic parameters:  $A$  – the rate at which strain, described by earthquake probability, increases per year, 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 unless multiple paleoseismic sites can be correlated to bound the length of the rupture. Here, we assume a fixed value for  $R$ . 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 the probability drop is much larger than the amount of probability accumulated in a year ( $R \gg A$ ), the probability usually resets to zero after an earthquake, as in renewal models. Otherwise, some residual probability (aka memory) is generally retained after the earthquake. Thus, if an inter-event time much longer than the average inter-event 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 where the probabilities of future outcomes only depend on the system’s most recent state (Girardin & 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, so that the probability of an earthquake in state 1 is 0 and the probability of an earthquake in state  $N$  is 1 (Figure 5.3B). In this formulation,  $N = 1 + \frac{1}{A}$ , where  $A$  is the probability increment in the original formulation. The earthquake drop  $D$  corresponds to  $\frac{R}{A}$  in the original formulation.

At each epoch (time step), an earthquake either occurs or does not. If an earthquake occurs, the system moves from state  $n$  to a lower state  $n - D$  or 1 (whichever is larger). 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 5.3C 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 3 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. Using this matrix, we can calculate earthquake probabilities following the procedures in the Appendix. Like 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 which can have a range of magnitudes, but treat each earthquake as having the same strain drop. 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 unless multiple sites are correlated.

Markov models have been used in seismic hazard analysis (Anagnos & Kiremidjian, 1988; Ebel et al., 2007; Votsi et al., 2013), but generally did not tie the states to strain increments as LTFM does. Anagnos & 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, the resulting probabilities should better reflect the earthquake process and the history of earthquakes on the fault.

With just two 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). For given  $D$  and  $N$  values, we can calculate the long-run distribution of inter-event times. This long-run PDF shows the expected distribution of inter-event times if we observed the fault for a sufficiently long time, analogous to the exponential, lognormal, and BPT PDFs in Figure 5.1B. Figure 5.4 shows a set of simulations with a fixed  $N$  and increasing  $D$  and the resulting long-run distribution of inter-event times and hazard rates. The hazard-rate curves represent the probability of an earthquake in the next time epoch, which we assume to be one year.

Simulations (i)-(iii) in Figure 5.4A 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 include the state space history showing the probability/strain history, and the corresponding earthquake record. In simulation (i)  $D \ll N$ , so after an earthquake the system rarely drops to the lowest state (where the probability of an earthquake equals 0), leaving residual probability for another earthquake. Hence the

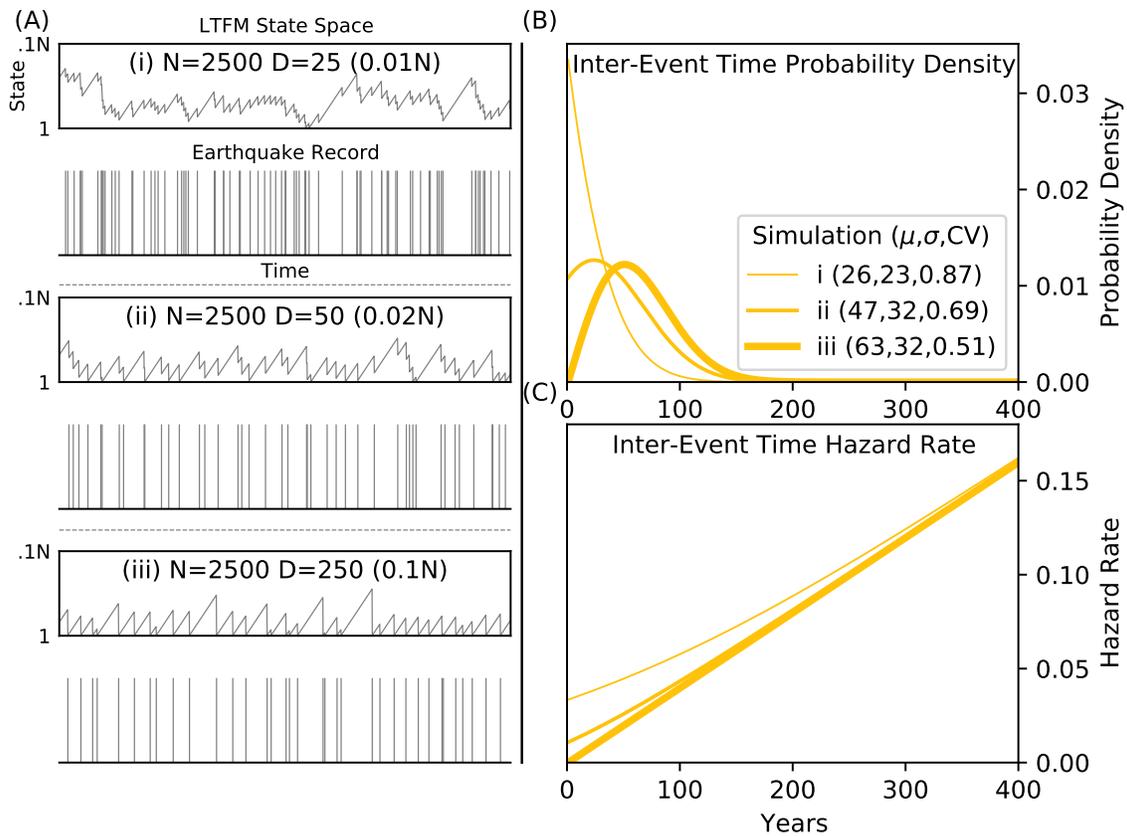


Figure 5.4. LTFM simulations with different input parameters. (A) Earthquake record simulations with fixed number of states  $N$  and increasing earthquake drop  $D$  (*i* – *iii*). (B) Long-run inter-event time probability distributions for the simulations in panel A. Thicker lines indicate larger values of  $D$ . Mean ( $\mu$ ), standard deviation ( $\sigma$ ), and coefficient of variation (CV) indicated. (C) Hazard-rate curves for probability density functions in panel B.

corresponding earthquake record contains clusters and long gaps. However as  $D$  increases relative to  $N$  (simulations *ii* & *iii*), 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 inter-event time and decreases the coefficient of variation (standard deviation/mean) (Figure 5.4B). With

a lower coefficient of variation, the earthquake record looks more regular (and periodic) with fewer outlier (short or long) inter-event times.

The shapes of the long-run inter-event time distributions (Figure 5.4B) 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 inter-event times possible, as in simulation (*i*) for  $D \ll N$ . Conversely, simulation (*iii*) 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 inter-event time distribution. A simulation with  $D=500$  and  $N=2500$  would be indistinguishable from simulation (*iii*) with  $D=250$  and  $N=2500$ , and the coefficient of variation does not continue to decrease below approximately 0.5 in the long run without the addition of a 3rd parameter (the threshold parameter) in the model.

The hazard curves also reflect LTFM's incorporation of fault memory. When a model is unlikely to have residual earthquake probability (*iii* in Figure 5.4C), the hazard function is a straight line, reflecting the linear strain/probability accumulation in LTFM. Simulations (*i*) and (*ii*)—which have the same number of states  $N$  as (*iii*) but different  $D$  values—initially have a higher hazard rate because they have more residual probability on average. However, as the quiescent period continues, the hazard curves with the same  $N$  (*i* and *ii*) asymptotically converge towards the curve in which residual strain is unlikely (*iii*). The number of states  $N$ , which reflects the strain accumulation rate, controls the slope of the hazard-rate curve. Increasing  $N$  decreases the slope, so the hazard rate grows

more slowly. These plots assume that one epoch equals one year. Assuming that one epoch equals another time increment would simply scale the probability/strain increment.

### 5.5. Inter-Event 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 inter-event times but in a different order (Figure 5.5A & 5.5B). We use a maximum likelihood estimation (MLE) approach to find the distribution parameters that best fit each record. Because the exponential, lognormal, and BPT models assume inter-event time independence, the order does not matter so the best fitting parameters (and therefore probability distributions) for each model are the same for both records (Figure 5.5C & 5.5D). 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. (For details of the MLE procedure and how the probability density and hazard curves are calculated, see the Appendix). Because LTFM does not assume inter-event time independence,  $N=11300$  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 inter-event times (solid LTFM-LR lines in Figure 5.5C & 5.5D). Although both records contain the same inter-event times, LTFM fitted to the original record indicates that the fault has more memory (retained strain) in the long run compared to the LTFM fitted to the reordered record. This is shown by the LTFM-LR distribution fitted to the original record having a larger positive  $y$ -intercept (Figure 5.5C) than the LTFM-LR distribution fitted to the reordered record (Figure 5.5D). The coefficients of variation for the LTFM-LR models differ from the

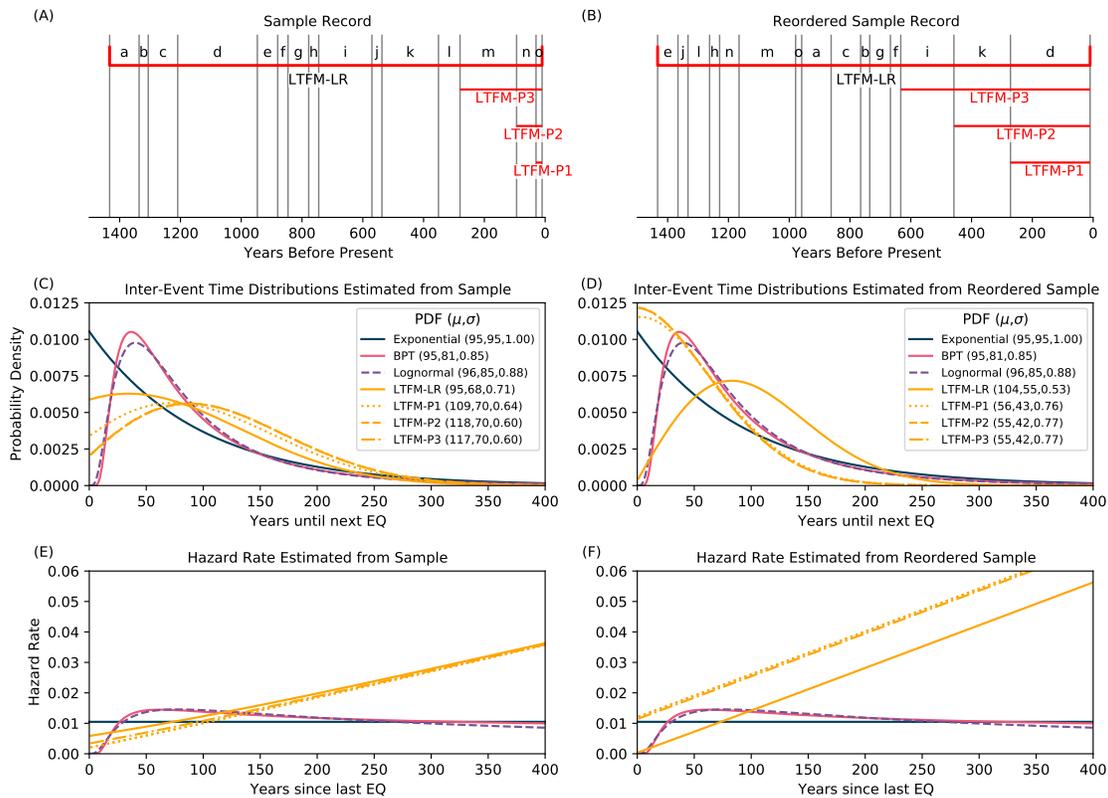


Figure 5.5. LTFM applied to sample records. (A) Earthquake record with inter-event times labeled. The best fitting LTFM model parameters for this sequence of inter-event times yield long-run probability distributions for the time until the next earthquake (LTFM-LR) and conditional probability distributions for the next earthquake given the most recent  $P$  inter-event times (LTFM-P1, LTFM-P2, LTFM-P3). (B) Reordered version of the record in panel A. (C) Estimated inter-event time probability distributions for the record in panel A with mean ( $\mu$ ), standard deviation ( $\sigma$ ), and coefficient of variation (CV) indicated. (D) Estimated inter-event 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 inter-event times. (E, F) Hazard-rate curves for inter-event time distributions in panels C and D.

time-dependent renewal models. For the lognormal and BPT models, the MLE produces parameters that match the paleoseismic record's coefficient of variation. The LTFM-LR,

however, assumes that a paleoseismic record is described by both the coefficient of variation and the long-term fault memory, so the MLE parameter estimates reflect both pieces of information. As Salditch et al. (2020) demonstrated, records with the same coefficient of variation can look very different, so the coefficient of variation only tells part of the story about the earthquake record.

Furthermore, we can use the LTFM parameters to calculate short-term earthquake probabilities conditional on the most recent sequence of inter-event times. These short-term probabilities incorporate extra information—the timing of the most recent prior  $P$  inter-event times (LTFM-P1, LTFM-P2, 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 inter-event 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 inter-event time) or a gap (longer average inter-event 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 making continuous adjustments in a data-driven manner.

The LTFM long-run and short-term forecasts estimate the probability that the fault is in a given state to calculate earthquake probabilities (see Appendix). 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  $P$  inter-event 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 inter-event times, the short-term forecast stabilizes and does not change noticeably with the inclusion of additional inter-event times farther back in time.

For the original record, conditioning on the most recent 1, 2, or 3 inter-event times decreases the probability of an earthquake in the near future by shifting the probability distributions (LTFM-P1, LTFM-P2, and LTFM-P3 in Figure 5.5C) to the right. This increases the expected mean inter-event time for the next earthquake relative to the long-run distribution (LTFM-LR). The distribution shifts to the right because the most recent inter-event times are short compared to 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 once information about the two most recent inter-event 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 inter-event times shifts the distribution to the left (LTFM-P1, LTFM-P2, and LTFM-P3 in Figure 5.5D) with a lower expected mean inter-event time until the next earthquake. The distribution shifts to the left because the recent inter-event times have been exceptionally long, leaving significant residual strain/probability after the most recent earthquake.

Comparison of the hazard rates highlights key differences between the LTFM and other probability models. Long-run LTFM-LR models for both the original (Figure 5.5E) and reordered records (Figure 5.5F) behave similarly. Initially, their hazard-rate curves are below the exponential distribution but higher than the lognormal and BPT distributions.

As the quiescent time increases, the lognormal and BPT distributions surpass the long-run LTFM-LR, indicating that an earthquake is more likely if it has not yet occurred. However, because in LTFM probability accumulates linearly, its hazard-rate function continues to climb, soon surpassing both the flat exponential distribution and the flattening/dropping lognormal and BPT distributions. Once LTFM-LR has crossed above these curves, it indicates that an earthquake is even more likely if it has not yet occurred. Conditioning on the most recent inter-event times (LTFM-P1, LTFM-P2, LTFM-P3) shifts when these crossovers occur (Figure 5.5E) or eliminates them (Figure 5.5F).

### **5.6. Earthquake Forecast Immediately after the 1812 Pallett Creek Earthquake**

To illustrate the effects in Figure 5.5, we use the Pallett Creek record to develop forecasts shortly after the 1812 earthquake. In hindsight, we know that the next earthquake occurs in 1857—a remarkably short 45-year inter-event period—but how well would any of the probability models have forecast such a short interval? Prior to 1812, there was an exceptionally long 304-year inter-event time. We re-analyze the Pallett Creek record up to and including the 1812 earthquake to get a set of parameters for each model and compare their forecasts. Figure 5.6A shows the estimated PDFs for the different models. Integrating the PDFs from 0 to 45 years yields estimated probabilities of another earthquake within 45 years of the 1812 earthquake (Figure 5.6B). Here we include an additional time-dependent renewal model—the Weibull distribution, discussed shortly—for comparison. The probability of such a short inter-event time is extremely low in the

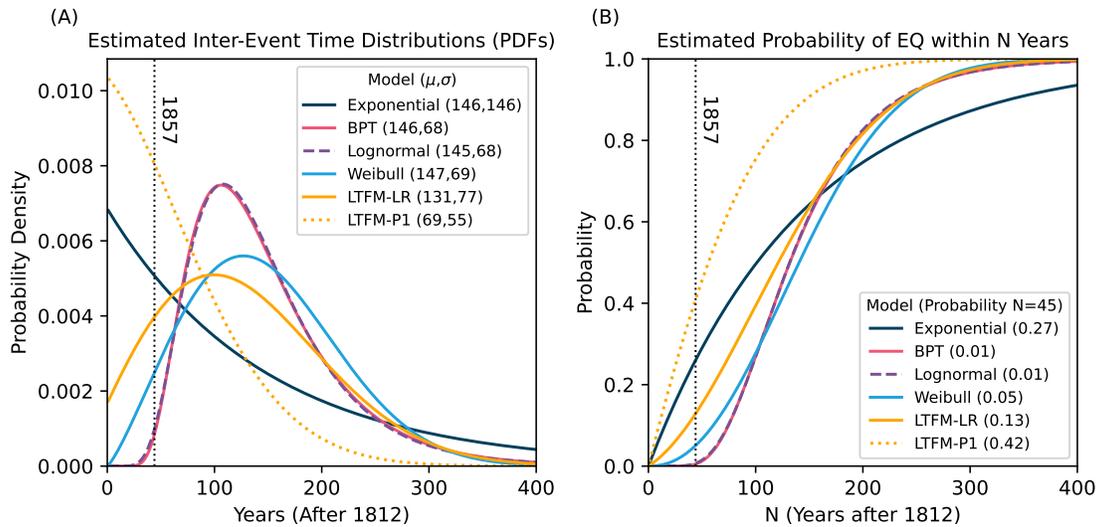


Figure 5.6. Probability estimates immediately after 1812 Pallett Creek earthquake. (A) Estimated probability density functions (PDFs) for inter-event time until the next earthquake. Empirical mean ( $\mu$ ) and standard deviation ( $\sigma$ ) indicated. BPT and lognormal curves overlap. (B) Probability of an earthquake within 45 years of 1812 earthquake for models in panel A. The next earthquake occurred in 1857 (dashed line). Values listed are probabilities of an earthquake within 45 years of 1812.

lognormal (0.01), BPT (0.01), and Weibull (0.05) models. LTFM-LR (0.13)—which allows for residual strain after an earthquake—and the exponential (0.27) model—which assumes constant earthquake probability—forecast a higher probability of a short inter-event time compared to the renewal models. Strikingly, LTFM-P1, which conditions on the date of the previous earthquake (1508) and so explicitly incorporates the recent long prior inter-event time, forecasts the highest probability (0.42) of observing an earthquake by 1857, as actually occurred. This example suggests that allowing residual earthquake strain and conditioning on the prior inter-event times provides a more accurate assessment

of future earthquake hazards, especially when the most recent earthquake is preceded by a relatively long inter-event time.

### 5.7. Assumptions About Earthquake Probability and Accumulated strain

LTFM's accurate forecast for the 1857 Fort Tejon earthquake is gratifying. However, it is just one earthquake, and because of the stochastic nature of the forecasts, there are likely some instances where the LTFM model performs worse. Unfortunately, determining which models perform best on average is challenging because paleoseismic records contain relatively few earthquakes. Time-dependent models fit the observed inter-event times better than a time-independent model (Scharer et al., 2010), but statistical tests struggle to identify the most appropriate time-dependent model due to the relatively few earthquakes in a paleoseismic record (Matthews et al., 2002). Hence in our view, model choices should reflect what is known about the relationship between accumulated strain and earthquake probability. Although some schematic earthquake recurrence models assume that tectonic strain and thus earthquake probability accumulate linearly in the interseismic period (Shimazaki and Nakata, 1980), the commonly used time-dependent models do not.

The choice of probability density function determines the hazard function, and hence the relationship between accumulating strain and earthquake probability. Thus when selecting a probability density function, attention should be paid to both the shape of the PDF and the resulting behavior of the hazard-rate function. In particular, the decline and flattening of the lognormal and BPT hazard-rate curves (Fig 5.2B) are intrinsic to these PDFs and do not reflect any updated knowledge of processes on the fault. These are hard

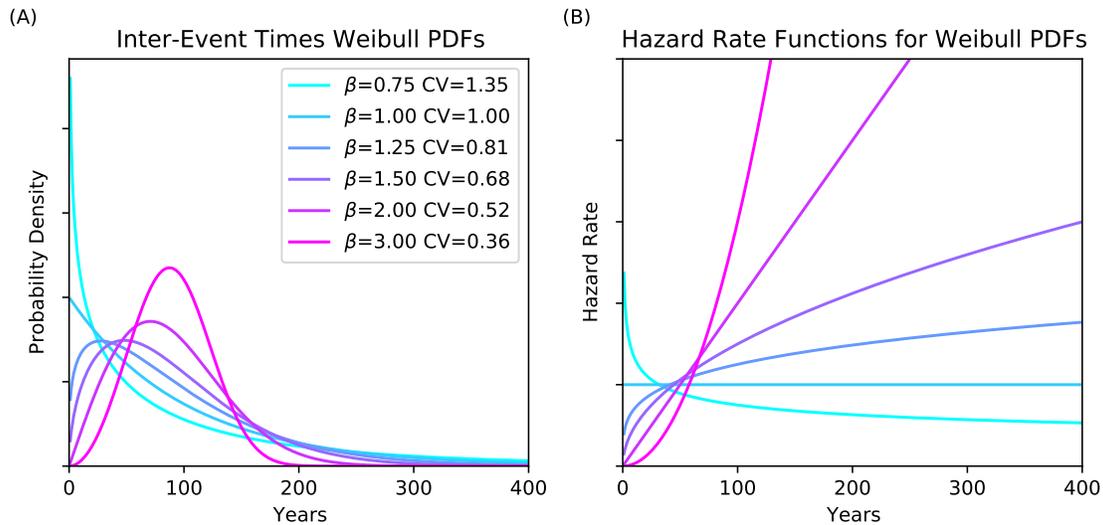


Figure 5.7. Weibull distribution with different shape parameters ( $\beta$ ). (A) Weibull probability density functions plotted with same scale parameter  $\eta$  but different  $\beta$  shape parameters. Coefficient of variation (CV) indicated. (B) Hazard-rate functions for the probability density functions shown in panel A.

to justify physically because tectonic plate motion loading the fault continues unabated. Although processes besides long-term tectonic loading like aseismic slip and Coulomb stress transfer from earthquakes on nearby faults may impact earthquake probability (Toda et al., 1998), these do not explain the shape of the lognormal and BPT probability curves.

Interestingly, some features of LTFM are shared by the Weibull distribution. Although the Weibull has fallen out of fashion in the seismological community, this distribution used in reliability engineering has been used to describe the distribution of earthquake inter-event times (Hagiwara, 1974; Chou and Fischer, 1975; Brillinger, 1982). The Weibull distribution has two parameters: a scale parameter  $\eta$  and shape parameter  $\beta$ .  $\beta$  controls the PDF's coefficient of variation (CV, Figure 5.7A) and the shape of the hazard-rate

function (Figure 5.7B). 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 inter-event times against theoretical cumulative hazard probabilities shows that the Weibull distribution is often appropriate (Chou & Fischer, 1975). If the 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  $\beta$ . Brillinger (1982) and Sieh et al. (1989) applied this method to an earlier Pallett Creek dataset, finding shape parameters of approximately 2 and  $1.5 \pm 0.8$ . A similar analysis (Figure 5.8A) shows that the Pallett Creek dataset we used is well-fit by a line whose corresponding  $\beta$  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  $\beta$  or with a fixed  $\beta = 2$ ) have quite similar PDFs (Figure 5.8B). The hazard-rate functions (Figure 5.8C) 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 (Figure 5.8B). The Weibull model, which is a renewal model, lacks long-term memory and so does not predict these. Hence Weibull, like the common

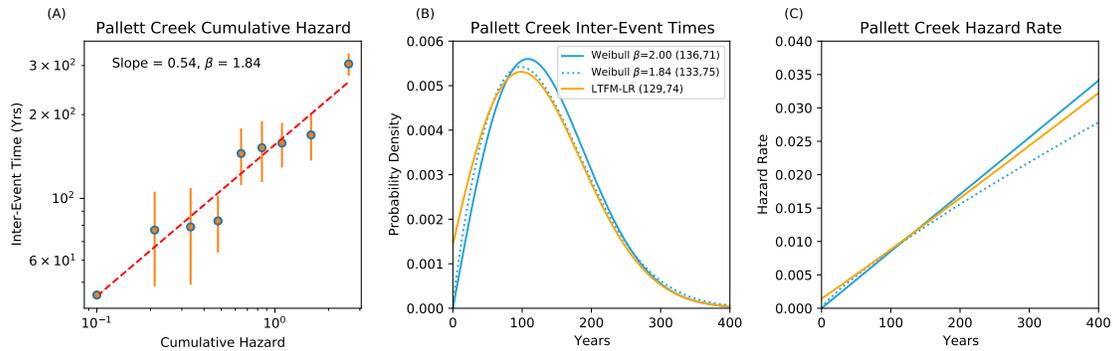


Figure 5.8. Comparisons of Weibull and LTFM. (A) Pallett Creek inter-event time cumulative hazard plot for Weibull distribution with  $1\sigma$  uncertainties indicated. (B) Comparison of best fitting Weibull distribution and  $\beta = 2$  Weibull distribution to the long-run LTFM-LR inter-event time distribution. Inter-event time mean ( $\mu$ ) and standard deviation ( $\sigma$ ) indicated. (C) Corresponding hazard-rate functions for PDFs in panel B.

renewal models, does not describe clusters and gaps and so does much worse than LTFM in forecasting the 1857 Fort Tejon earthquake (Figure 5.6).

## 5.8. Conclusion

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 avoids the implausible result in current models that probabilities flatten or even decrease with time after the expected mean recurrence interval. It also 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 inter-event time. Although LTFM is computationally more complex than the existing earthquake probability models, it is far more powerful. It allows forecasts with more information than the current models can include and should lead to more realistic earthquake probability estimates.

## CHAPTER 6

## The Generalized Long-Term Fault Memory Model and Applications to Paleoseismic Records

### 6.1. Introduction

What is the probability of an earthquake occurring in the near future? The answer to this question has important implications for both hazard planning and public awareness of a region's hazards. Estimates of the probability can vary significantly depending on methodology. For instance, along the Nankai Trough off the coast of Japan, estimates of the probability of an earthquake in the next 30 years range from 6% to 80% depending on model assumptions (Hashimoto, 2022). Because the forecasts depend heavily on model choice (Neely et al., in review), it is critical that the models reflect our understanding of earthquake processes.

However, commonly used earthquake probability models make assumptions that conflict with our understanding of earthquake processes informed by geologic evidence. Here, I present a new earthquake probability model that builds on prior work in Chapter 5 (Salditch et al., 2020; Neely et al., in review) to construct an earthquake recurrence model that reflects our understanding of earthquake processes. This new method is much easier to implement and less computationally expensive than its prior iterations.

The earthquake process is fundamentally about strain. Strain accumulates on faults during the inter-seismic period and is released suddenly during an earthquake. Simple

models of this strain accumulation and release, like the elastic rebound model (Reid, 1910) date back to the early 20th century. In principle, if we knew the current level of strain on a fault and the threshold level at which the accumulated strain overcomes the fault's strength, we could forecast when the next earthquake would occur. However, neither of these values are easy to measure.

Instead, most earthquake probability forecasts rely on an alternative approach. Seismologists choose a probability model that they think best characterizes the observed inter-event times (e.g., Field et al., 2015; Schorlemmer et al., 2018) to produce forecasts for the next earthquake. However some of the assumptions of the most commonly used models ignore key aspects of the strain accumulation and release process (Salditch, et al., 2020; Neely et al., in review).

Earthquake recurrence models can be broadly grouped into three categories based on how time impacts earthquake probability (Salditch, et al., 2020). *Memoryless models* assume earthquake probability does not change with time. Even when an earthquake occurs, the probability of an earthquake at the next moment remains unchanged. *Short-term memory models* assume that earthquake probability depends on the elapsed time since the most recent earthquake but not on the order of the prior inter-event times. When an earthquake occurs, the probability resets to zero, so these models only “remember” when the most recent earthquake occurred. *Long-term memory models*, however, depend both on the time since the most recent earthquake and the specific timing of prior earthquakes. These probability models do not necessarily reset after an earthquake. Thus they allow for the possibility of some residual probability that will impact the timing of the next earthquake in the sequence.

The memoryless and short-term models are purely statistical and assume that future earthquakes will satisfy a probability distribution that describes the times between past large earthquakes. These models are fit to the mean and standard deviation of the inter-event times. These models describe average earthquake behavior well, but not deviations from it, because they do not incorporate fundamental aspects of the strain accumulation and release processes that cause earthquakes. Long-term memory models, on the other hand, seek to model the specific sequence of observed events by mimicking the underlying strain accumulation and release process that produced that sequence.

Because tectonic loading leads to constant strain accumulation along faults (Gordon & Stein, 1992), these time/earthquake probability assumptions correspond to strain/probability assumptions (Neely et al., in review). For memoryless models, increasing strain accumulation has no impact on earthquake probability. For short-term memory models, increasing accumulated strain impacts earthquake probability, but because the probability resets after an earthquake, it implies that there is never any residual strain after an earthquake. For the long-term memory models, increasing accumulated strain impacts earthquake probability, but because the probability does not necessarily reset after an earthquake, the model allows the possibility of residual strain.

Paleoseismic records of past large earthquakes suggest that long-term memory models most accurately reflect the underlying strain processes. These records contain inter-event time trends that suggest a complex pattern of strain accumulation and release (Wallace, 1970; Rockwell et al., 2000; Friedrich et al., 2003; Weldon et al., 2004; Sieh et al., 2008; Goldfinger et al., 2013; Salditch et al., 2020; Hecker et al., 2021). Clusters of earthquakes and long gaps between them in the seismic record suggest that earthquakes may only

release some of the accumulated strain, so the level of residual strain impacts the timing until the next earthquake. Long-term memory probability models, which model the probability impact of residual strain over multiple earthquakes, attempt to recreate the behavior observed in the paleoseismic record.

## 6.2. The Mechanics of Probability Models

To construct an earthquake forecast using a memoryless or short-term memory model, seismologists follow a series of steps. First, they compile a record of when past large earthquakes occurred—usually by combining instrumental, historical, and geologic records (e.g. Sieh et al., 1989). These earthquake histories (referred to as paleoseismic records when they contain primarily geologic evidence of past earthquakes) indicate when earthquakes occurred but usually do not have much information about the size of the earthquakes except that they were large enough to leave observable geologic evidence (unless they could be correlated along a specific fault segment). In Figure 6.1A, I show a sample paleoseismic record consisting of  $n + 1$  earthquakes occurring at successive times  $t_0, t_1, \dots, t_n$  with positive inter-event times  $u_i = t_i - t_{i-1}, i = 1, \dots, n$ . With the paleoseismic record, seismologists fit a probability density function (PDF)  $f(u)$  to the distribution of the inter-event times  $u_i$  (Figure 6.1B). I have included fits (using a maximum likelihood method) for the memoryless exponential model as well as the short-term memory Brownian Passage Time (BPT) and Weibull models. Both the memoryless and short-term memory models assume that the inter-event times are independent, so reordering the  $u_i$ 's would not impact the parameter estimates. After finding the appropriate PDF parameters, it's straightforward to calculate the cumulative distribution function (CDF)  $F(u)$  by integrating the PDF from

0 to  $u$  (Figure 6.1C). The CDF indicates the estimated probability that an earthquake will occur within  $u$  years.

An important related function is the hazard rate  $h(u) = f(u)/[1 - F(u)]$ , which is calculated from the PDF and CDF. The hazard rate function indicates the rate at which an earthquake will occur at time  $u$  since the prior earthquake (Figure 6.1D). More specifically, it is the probability density of an earthquake occurring at time  $u$  given the prior earthquake occurred  $u$  years ago. Although the hazard rate is not a probability itself (it can exceed 1), the shape of the function shows how the probability of an earthquake changes with time (and therefore accumulated strain). The memoryless exponential model has a flat hazard rate function indicating that the probability does not change with time, whereas the short-term memory Brownian Passage Time and Weibull models change with time but in different ways.

Long-term memory models do not assume that the inter-event times are independent and instead allow the possibility of residual probability after an earthquake. Thus the specific order matters (Salditch et al., 2020). A long inter-event time may make a short inter-event time for the next earthquake more likely and vice versa (Neely et al., in review). Salditch et al.'s (2020) Long-Term Fault Memory (LTFM) model (described in detail in Chapter 5) assumes that earthquake probability increases linearly at rate  $A$  and decreases by some value  $R$  when an earthquake occurs. If the accumulated probability prior to the earthquake was greater than  $R$ , the remaining residual probability will impact the timing of the subsequent earthquakes. LTFM was designed to mimic the linear strain accumulation and partial strain release patterns suggested by paleoseismic records. Salditch et al. (2020) used this discrete-time numerical model to show how a wide range of earthquake

recurrence behaviors could be reproduced by modifying the input parameters. However, its numerical formulation meant that 1000s of simulations were needed to find  $f(u)$ ,  $F(u)$ , and  $h(u)$ . In subsequent work, I reformulated the LTFM as a hidden Markov model (see Chapter 5; Neely et al., in review) which allowed me to derive the necessary equations for  $f(u)$ ,  $F(u)$ , and  $h(u)$  to calculate probabilities of interest (see Appendix). However, the Markov reformulation required significant computational resources to construct the necessary matrices to make these calculations.

In this chapter, I present a revised version of the LTFM model, which I call the Generalized LTFM. Unlike the prior iterations of LTFM, the Generalized LTFM is in continuous, not discrete, time and the earthquake probability accumulation function can take any form (it does not need to accumulate linearly). The Generalized LTFM's equations are simpler to code than the hidden Markov formulation, which allows faster calculation of earthquake probabilities. Below, I present the equations necessary to implement the Generalized LTFM and apply this model to various paleoseismic records.

### 6.3. The Generalized LTFM

The Generalized LTFM (GLTFM) builds upon the key assumption in prior LTFM models: if there is residual strain after an earthquake it will likely impact the inter-event time to the next earthquake. The inter-event time PDF, CDF, and hazard rate function, denoted by  $f(u|z)$ ,  $F(u|z)$ , and  $h(u|z)$ , depend on the residual strain  $z$  at the start of the inter-event period, where  $z > 0$ . If the residual strain  $z = 0$ , then  $f(u|z = 0) = f(u)$  and likewise for  $F(u|z)$  and  $h(u|z)$ . Because I only have measurements of the time between

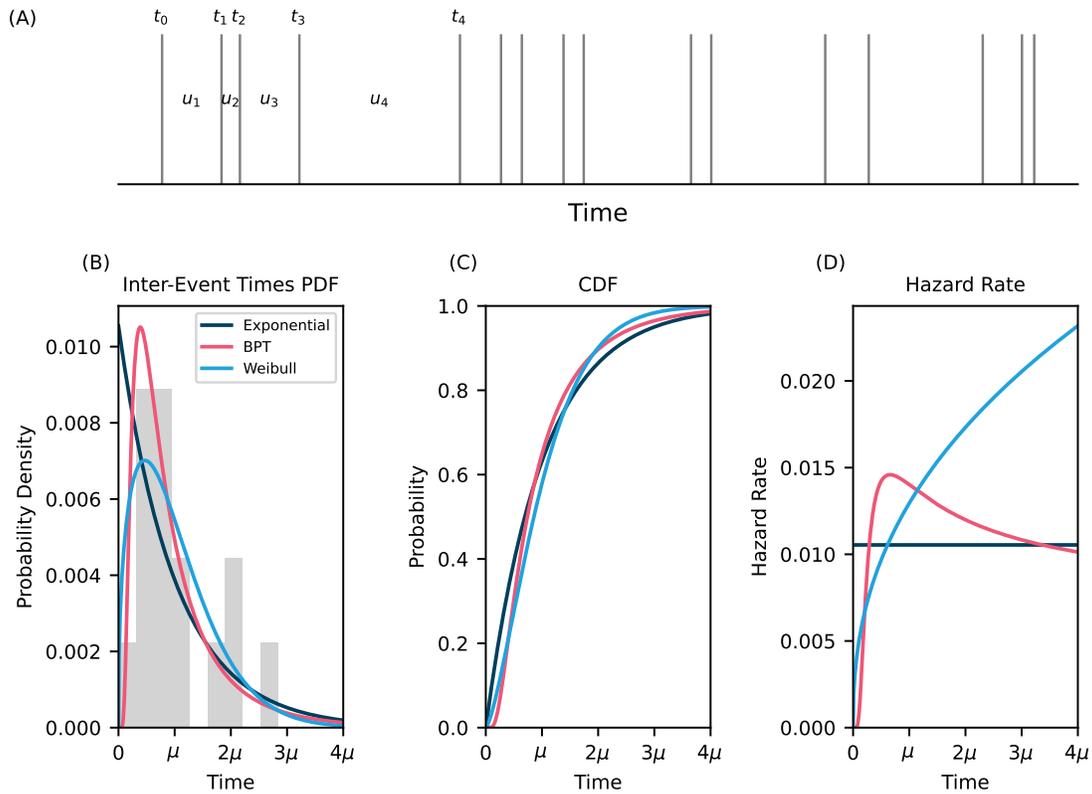


Figure 6.1. Application of memoryless and short-term memory models. (A) Sample paleoseismic record with earthquakes occurring at  $t_i$  and separated by inter-event times  $u_i$ . The first 5 earthquakes and four inter-event times are labeled. (B) Probability density functions (PDF) fit to the distribution of inter-event times  $u_i$  indicated by the histogram. Times are given in terms of  $\mu$ , the mean inter-event time from the paleoseismic record. (C) Cumulative distribution functions (CDF) for corresponding PDFs. (D) Hazard rate functions for PDFs.

earthquakes and not the residual strain, I define  $z$  in terms of time (aka  $z$  years of residual strain).

To calculate  $f(u|z)$  consider the following example. Find the probability of waiting exactly  $u_0$  years until the next earthquake given that we already have  $z_0$  years of residual strain at the start of the inter-event period  $f(u_0|z_0)$ . This probability is the same as

the probability of waiting  $u_0$  more years until the earthquake given that we have already waited  $z_0$  years without an earthquake. Therefore  $f(u|z)$  can be expressed as

$$(6.1) \quad f(u|z) = f(u+z)/[1-F(z)]$$

Figure 6.2A shows a simple example of how to calculate  $f(u|z)$  (red curve) using  $f(u)$  (black curve). To calculate  $f(u=75|z=50)$ —the probability (indicated by the red X) of waiting 75 years until the next earthquake given that there is 50 years of residual strain at the start of the inter-event period, first retrieve the value  $f(u+z=125)$  (indicated by the black X) then divide that by 1 minus the area between 0 and 50 under the black curve  $F(50)$  (shaded grey area). The function  $f(u|z)$  is simply  $f(u)$  shifted to the left by  $z$  time units and renormalized to keep the total probability as 1 by dividing by the area of the curve that is now to the left of time 0,  $1-F(z)$ . The concept of time shifting PDFs has been used by Stein et al. (1997) to model the probability impact of stress transfer between earthquakes on different fault sections, but this is the first time it has been used to model residual strain after an earthquake. With  $f(u|z)$ , I apply probability relationships to derive  $F(u|z)$  and  $h(u|z)$ . Integrating  $f(u|z)$  from 0 to  $u$  gives the CDF

$$(6.2) \quad F(u|z) = [F(u+z) - F(z)]/[1-F(z)]$$

Plugging  $f(u|z)$  and  $F(u|z)$  into the hazard rate function equation gives

$$(6.3) \quad h(u|z) = f(u+z)/[1-F(u+z)]$$

Note that Equation (6.3) indicates that  $h(u|z) = h(u+z)$  based on the definition of  $h(u)$ .

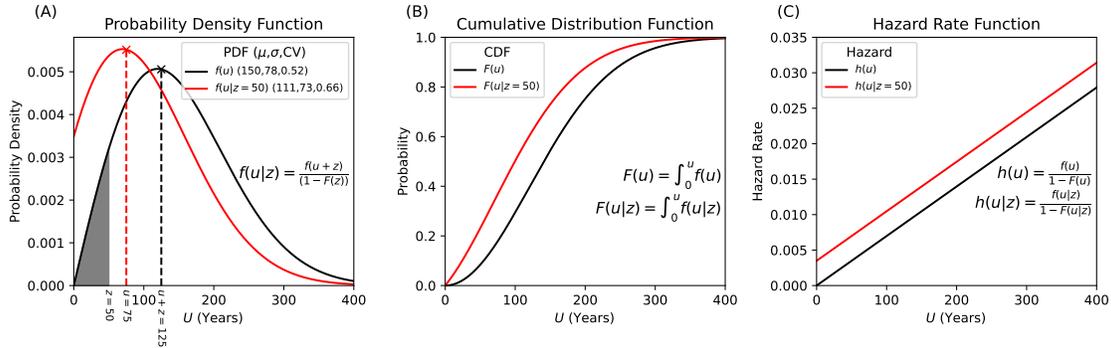


Figure 6.2. GLTFM calculation schematic. (A) PDF plot showing how to calculate  $f(u|z = 50)$  (red curve) from  $f(u)$  (black curve). The black X [ $f(75 + 50)$ ] and the shaded gray area [ $F(50)$ ] are the values needed to calculate  $f(u = 75|z = 50)$  (red X). Mean ( $\mu$ ), standard deviations ( $\sigma$ ), and coefficient of variation ( $\sigma/\mu$ ) indicated for the curves. (B) Cumulative distribution functions for the PDFs in (A). (C) Corresponding hazard rate functions.

Figure 6.2B and 6.2C further show how residual strain impacts our expectation of earthquake probability. For a given probability value,  $F(u|z)$  always corresponds to a smaller  $u$  value than  $F(u)$  indicating that residual strain makes shorter inter-event times more likely (Figure 6.2B). Likewise, for a given inter-event time  $u$ , the hazard rate (Figure 6.2C) is higher when residual strain is present if the hazard rate function is monotonically increasing.

To calculate the residual strain, I assume that the first observed earthquake occurred at time  $t_0$  and left residual strain of  $z_0$ . Although I do not observe  $z_0$ , I specify a value as if it were known (the impact of this assumption is discussed more in detail below). Suppose the second earthquake occurred at time  $t_1$ . The residual strain following that earthquake is  $z_1 = \max[0, (u_1 + z_0 - R)]$  where  $R$ , defined in terms of time, is the amount of strain released during the earthquake.  $z_1$  depends on  $u_1 = t_1 - t_0$ , which is observed,  $z_0$ , which

is specified, and  $R$  the amount of accumulated strain the earthquake releases. I choose  $R$  and the parameters of the assumed probability function  $f$  to maximize the likelihood of observing the earthquake sequence. Suppose the third earthquake occurred at time  $t_2$ . The residual strain following the earthquake at time  $t_2$  is  $z_2 = \max[0, (u_2 + z_1 - R)]$ . The value  $z_2$  depends on  $u_2 = t_2 - t_1$ , which is observed,  $z_0$ , which is specified, and  $R$ . Generally, the  $k^{\text{th}}$  earthquake occurs at time  $t_k$  with  $z_{k-1} = \max[0, (u_{k-1} + z_{k-2} - R)]$ . The values of  $z_1, \dots, z_k$  depend on  $u_1, \dots, u_k$ , which are observed,  $z_0$ , which is specified, and  $R$ .

Figure 6.3 shows the GLTFM applied to the sample paleoseismic record from Figure 6.1. Figure 6.3B illustrates how the hazard rate changes with time and whether there is residual strain after an earthquake. The color solid lines correspond to earthquakes in panel A, and the dashed color lines indicate the hazard rate in the inter-event period before the corresponding earthquake. The hazard rate builds during the inter-event period until an earthquake occurs. The hazard rate drops to zero after the first four earthquakes, indicating that there is no residual strain. However, after the long inter-event time  $u_4$ , the earthquake at  $t_4$  does not reset the hazard rate to zero, indicating residual strain. The hazard rate curve stays elevated over the next four earthquakes until the earthquake at  $t_8$  drops the hazard rate back to zero and wipes out any long-term memory in the system. Figure 6.3C shows the impact of the residual strain on the expected distribution of inter-event times. Each inter-event time  $u_k$  has an expected distribution based on the initial residual strain at the start of the inter-event time. In Figure 6.3C,  $\mu$  is the mean of the observed inter-event times, not the mean of the PDFs. The expected PDFs for  $u_1$  through  $u_4$  and  $u_9$  are identical because each of those inter-event times starts with

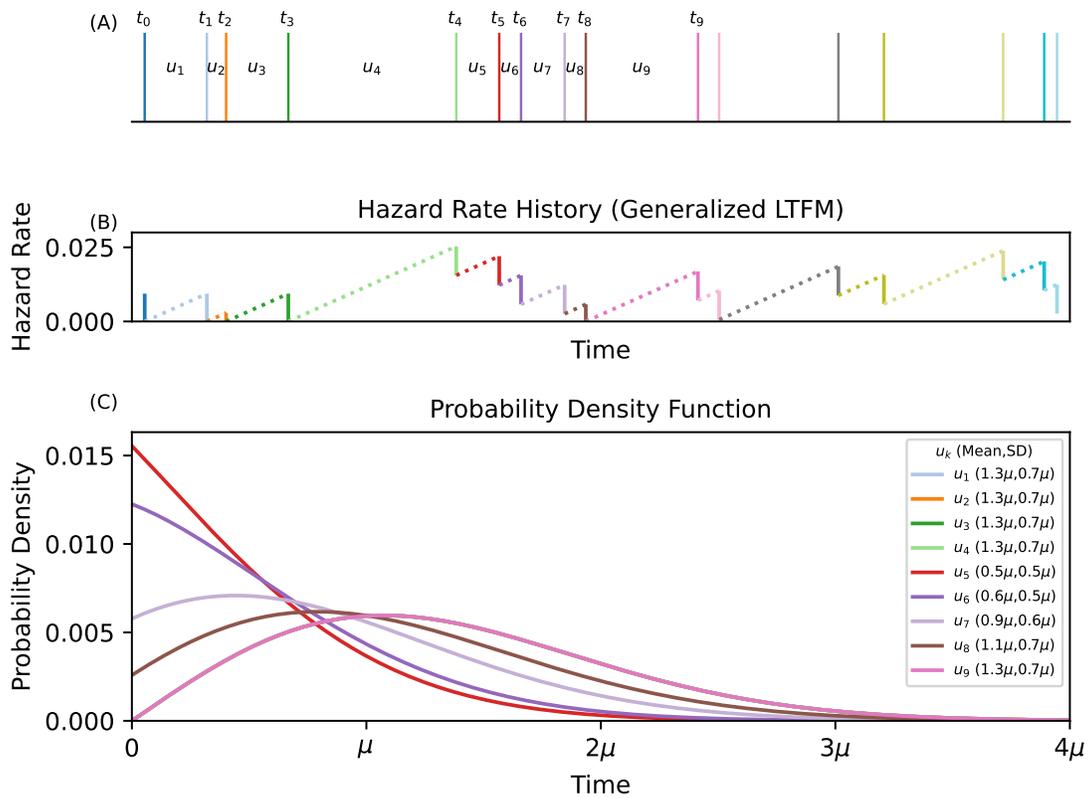


Figure 6.3. GLTFM applied to sample paleoseismic record. (A) Sample paleoseismic record with first 10 earthquakes and 9 inter-event times labeled. (B) Hazard rate history plot which shows the hazard rate changes with time and whether residual strain is present after an earthquake. Solid lines correspond to earthquakes in panel (A). Dashed lines show how hazard rate builds during the inter-event period. If the hazard rate history curve drops to zero then there is no residual strain, otherwise some residual strain remains after an earthquake. (C) Corresponding estimated probability density functions for first 9 inter-event times. Time and the mean and standard deviations for each PDF are given in terms of  $\mu$ , the mean inter-event time from the paleoseismic record.

zero residual strain. Inter-event times  $u_5$  through  $u_8$  have different PDFs with shorter expected intervals because these sequences each start with some residual strain.  $u_5$  has the shortest expected inter-event time mean with the values increasing with each subsequent

earthquake until the earthquake at  $t_8$ , after which no residual strain remains. In contrast, in the conventional short-term memory models, every time an earthquake occurs, the hazard rate drops to zero because of the assumption of no residual strain. Additionally, the PDFs for each inter-event time would be identical.

#### 6.4. Fitting Parameters and Choosing the Probability Function $f$

In addition to estimating  $R$ , the earthquake drop size in terms of time, I also need to estimate the parameters of the chosen probability function  $f$ . Although  $f$  can contain many parameters, the most commonly used models have just two parameters, which I will refer to generically as  $\alpha$  and  $\omega$ . The actual names of these parameters will vary depending on the probability function. To find these, I apply a maximum likelihood approach to the likelihood function  $L(\alpha, \omega, R)$  in Equation (6.4). The value of the likelihood function is the probability of observing the actual specific sequence of inter-event times

$$(6.4) \quad L(\alpha, \omega, R) = f(u_1|z_0) \times f(u_2|z_1) \times \dots \times f(u_n|z_{n-1})$$

The  $u_i$ 's indicate the inter-event times and  $z_i$ 's the residual strain at the start of the  $i^{th}$  inter-event time. I can find the values of  $\alpha$ ,  $\omega$ , and  $R$  that maximize the likelihood  $L(\alpha, \omega, R)$  using any number of optimization functions.

The GLTFM is designed so that any probability function  $f$  can be used. However some probability functions are more realistic for modeling earthquake recurrence than others. The lognormal and BPT distributions, two of the most commonly used probability models, behave in counterintuitive ways when an earthquake has not happened by the average inter-event time. The hazard rates for these functions decrease steadily to zero (lognormal)

or decrease and asymptotically approach a lower value (BPT) after the quiescent period extends beyond approximately the average inter-event time (Matthews et al., 2002). This behavior indicates that an earthquake becomes less likely as the quiescent time continues and implies that adding more stored strain on the fault decreases the probability of an earthquake, which seems unlikely (see Figure 5.2 in Chapter 5).

A model with a monotonically increasing hazard rate function makes more sense conceptually. Such functions imply that as the quiescent period lengthens (and more strain accumulates) the estimated earthquake probability always increases. The 2-parameter Weibull distribution, with shape parameter ( $\beta$ ) and scale parameter ( $\eta$ ) fits this criterion. A Weibull distribution's shape parameter ( $\beta$ ) controls the shape of the hazard function and implies different relationships between accumulated strain and earthquake probabilities. For  $\beta < 1$ , the hazard rate decreases with time (additional strain decreases the earthquake probability), for  $\beta = 1$  the hazard rate is constant (additional strain does not impact earthquake probability), for  $\beta > 1$  the hazard rate increases with time (additional strain increases earthquake probability). For  $1 < \beta < 2$  the hazard rate increases with time but at decreasing rate. For  $\beta = 2$ , the hazard rate increases linearly with time. For  $\beta > 2$ , the hazard rate increases with time at an increasing rate. As discussed in Chapter 5,  $\beta = 2$  corresponds to the Markov LTFM formulation which assumed a constant rate of probability accumulation.

I explore the effect of using different  $\beta$  values [2.5, 2, 1.5] on earthquake probability in Figure 6.4. I fix  $\beta$  and then solve for the best fitting  $\eta$  and  $R$ . Despite the different  $\beta$  values, the GLTFM hazard rate plots (Figure 6.4A-C) show many similarities. Regardless of  $\beta$ , the earthquakes from  $t_4$  through  $t_7$  and  $t_9$  through  $t_{15}$  show similar patterns of

partial strain release although the amount of residual strain varies among the models. The forecasts for the next earthquake, however, vary. The  $\beta = 1.5$  model predicts the likely time until the next earthquake to be shorter whereas the  $\beta = 2.5$  model predicts a longer inter-event time (Figure 6.4D). The 30-year earthquake forecasts, which show the probability of an earthquake occurring within the next 30 years, further highlight these differences (Figure 6.4E). Initially, the  $\beta = 1.5$  model forecasts a higher probability of an earthquake, but it is surpassed by the other models near the expected mean recurrence time. The  $\beta = 2.5$  model ultimately forecasts the highest probabilities as the quiescent period continues because earthquake probability accumulates faster with time.

### 6.5. GLTFM Applied to Real Paleoseismic Records

How does GLTFM compare to conventional models when applied to real earthquake records? Figures 6.5 – 6.9 show GLTFM applied to paleoseismic records of the Pallett Creek section of the southern San Andreas fault (Scharer et al., 2011), the Hayward fault along the east side of San Francisco Bay (Lienkaemper & Williams, 2007), the Nankai Trough subduction zone (Ando, 1975) off the coast of Japan, the Cascadia subduction zone (Goldfinger et al., 2012, 2017), and the Alpine fault in New Zealand (Clark et al., 2013; Cochran et al., 2017). For each location, I assume a 2-parameter Weibull distribution and apply the maximum likelihood method to find  $\beta$ ,  $\eta$ , and  $R$ . I assumed an initial residual strain of  $z_0 = 0$  for all locations. For each fault, the figure includes the hazard rate history plot and the corresponding forecast of the probability of an earthquake in the next 50 years. Above each earthquake in the hazard rate history plot, I show the remaining residual strain (defined in terms of years of residual strain) on the fault after

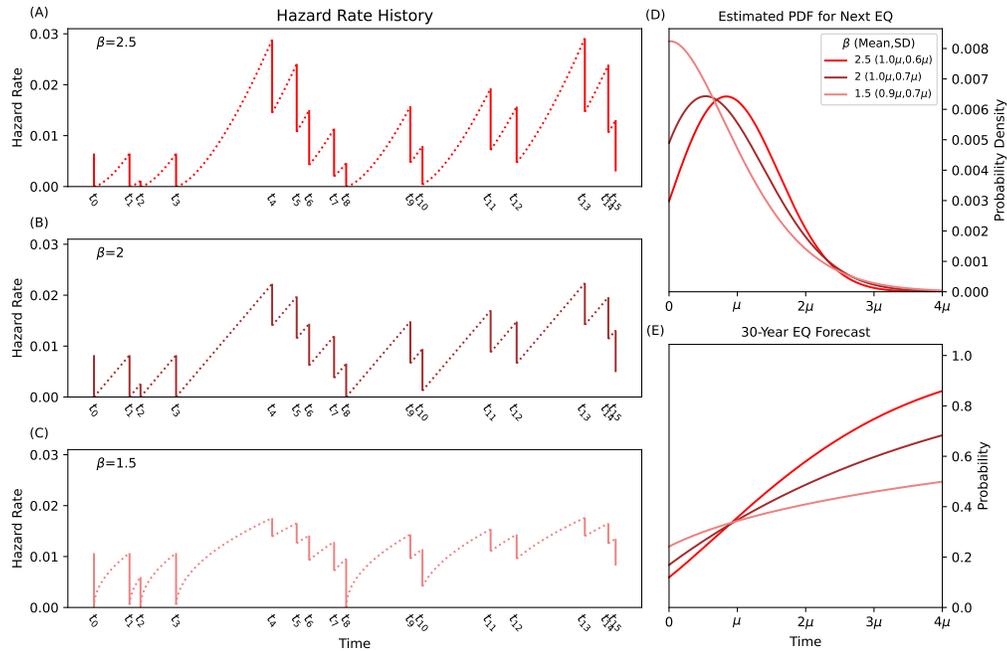


Figure 6.4. Impact of probability accumulation rate on earthquake probability. (A) – (C) GLTFM hazard rate history plots for sample paleoseismic record with 2-parameter Weibull distribution with different fixed  $\beta$ 's. (D) Estimated PDFs for the next earthquake in the sequence for models in (A) – (C). (E) Corresponding 30-year earthquake forecasts after the most recent earthquake. Time and the mean and standard deviations for each PDF are given in terms of  $\mu$ , the mean inter-event time from the paleoseismic record.

that earthquake. For comparison, I also show the 50-year forecasts for the memoryless exponential and short-term memory lognormal, Weibull, and BPT models.

GLTFM estimates that there was residual strain after the most recent earthquakes in the Pallett Creek (Figure 6.5A) and Hayward fault records (Figure 6.6A), which increases the likelihood of an earthquake in the near term. Along the Pallett Creek section of the San Andreas fault, GLTFM estimates that the most recent earthquakes in 1812 and 1857 left 135 years and 11 years of residual strain, respectively. Along the Hayward Fault, the

most recent earthquake likely left 70 years of residual strain. These residual strains arise because of long inter-event times in the sequences. At Pallett Creek, GLTFM estimates a 60% probability of an earthquake in next 50 years, higher than the exponential (31%), BPT (50%), lognormal (51%), and Weibull (55%) (Figure 6.5B). The impact of residual strain on earthquake probability is even more apparent for the Hayward fault (Figure 6.6B). GLTFM estimates that there is a 92% chance of an earthquake in the next 50 years, but the exponential (25%), BPT (66%), lognormal (67%), and Weibull (63%) are all much lower. For a 30-year forecast (not shown), the GLTFM (73%) is 60% higher than the lognormal and BPT (45%) models. The residual strain in both systems reflects the relatively recent and exceptionally long (for the respective faults) inter-event times.

Along the Nankai Trough (the paleoseismic record is for the A-B section of the subduction zone), the 1361 and 1605 earthquakes may also have left residual strain, but GLTFM estimates all residual strain has been released by the 1707 earthquake (Figure 6.7A), and none persists after the two more recent events. With no estimated residual strain at the start of the current quiescent period, the GLTFM 50-year forecast is lower (12%) than the exponential (24%), BPT (23%), lognormal (22%), and Weibull (16%) models (Figure 6.7B). These different estimates reflect the differences in how each model assumes earthquake probability accumulates with time. If no earthquake occurs in the next 100 years, the GLTFM forecast will overtake the exponential, lognormal, and BPT forecasts because the GLTFM assumes that additional accumulated strain always increases earthquake probability but these other models do not.

For the Cascadia subduction zone (Figure 6.8A) and Alpine fault (Figure 6.9A), the GLTFM estimates that each earthquake released all accumulated strain. In these cases,

the GLTFM is identical to the short-term memory Weibull model, and the 50-year forecasts (Figure 6.8B and 6.9B) are indistinguishable. Although the GTLFM estimates that there is no residual strain after any earthquake in either of these records, it does not mean that future earthquakes will also always release all accumulated strain. If an exceptional long inter-event time (relative to the rest of the paleoseismic record) occurs, some residual strain would likely remain after the subsequent earthquake, which would impact future earthquake probability.

Although I assumed the starting residual strain ( $z_0$ ) equaled 0, it could be a non-zero value. I performed a sensitivity analysis to determine how varying the starting strain threshold after the first earthquake impacts the 50-year forecast. For each paleoseismic record, I selected a new starting residual strain ( $z_0$ ) that corresponds to the probability that the accumulated strain right before the first earthquake is in the 99<sup>th</sup> percentile based on the estimated GLTFM parameters. This corresponds to a new  $z_0 = \max[0, (F^{-1}(.99) - R)]$  where  $F^{-1}(.99)$  is the inverse CDF evaluated at 0.99. Both  $F^{-1}(.99)$  and  $R$  are calculated from the initial  $z_0 = 0$  maximum likelihood parameters. With this new starting  $z_0$ , I refit the GLTFM model. I have included the resulting 50-year forecasts as grey dashed lines. For Cascadia and the Alpine fault, the starting residual strain ( $z_0$ ) remains 0, indicating that residual strain would be unlikely along those faults based on the observed paleoseismic record. For Pallett Creek and the Hayward fault, starting with the 99<sup>th</sup> percentile residual strain increases the likelihood of an earthquake in the next 50 years slightly, and for the Nankai Trough it decreases the probability.

The GLTFM suggests that relatively few earthquakes leave residual strain along the fault. This may be true or may reflect the simplifying assumptions in the model. The

minimization function only identifies residual strain when an abnormally long inter-event time is in the record. It is possible that residual strain remains at other times, but the likelihood function simply does not identify it. The lack of identified residual strain may also be due to the assumption that all paleoseismic earthquakes along a specific fault have the same magnitude. For many locations, this is the best that we can do. However, along faults where there are multiple paleoseismic study sites, earthquake magnitudes can be estimated. Along the Cascadia subduction zone, Goldfinger et al. (2012) has assigned different magnitudes to various earthquakes. Incorporating this magnitude information into GLTFM by allowing for a variable earthquake probability drop  $R$  may show that residual strain is more prevalent than initially indicated.

## 6.6. Identifying the Best Model

Identifying which probability model performs best is challenging. Prior analyses have shown (Matthews et al., 2002) that most statistical tests fail to differentiate between the models. There is evidence that the short-term memory models perform better than the memoryless exponential model (Scharer et al., 2010), but differentiating between short-term memory models is challenging. Prior proposals for “better” models have relied on qualitative arguments. The BPT was preferred over the lognormal distribution because the probability of an earthquake will eventually return to zero with the lognormal distribution if the quiescent period continues for a long time (Matthews et al., 2002). Unlike the lognormal, the BPT decreases but then asymptotically approaches a stable probability value in the long run. Although this is preferable to the lognormal model’s behavior, I still consider it unsatisfactory. Shouldn’t earthquake probability continue to increase

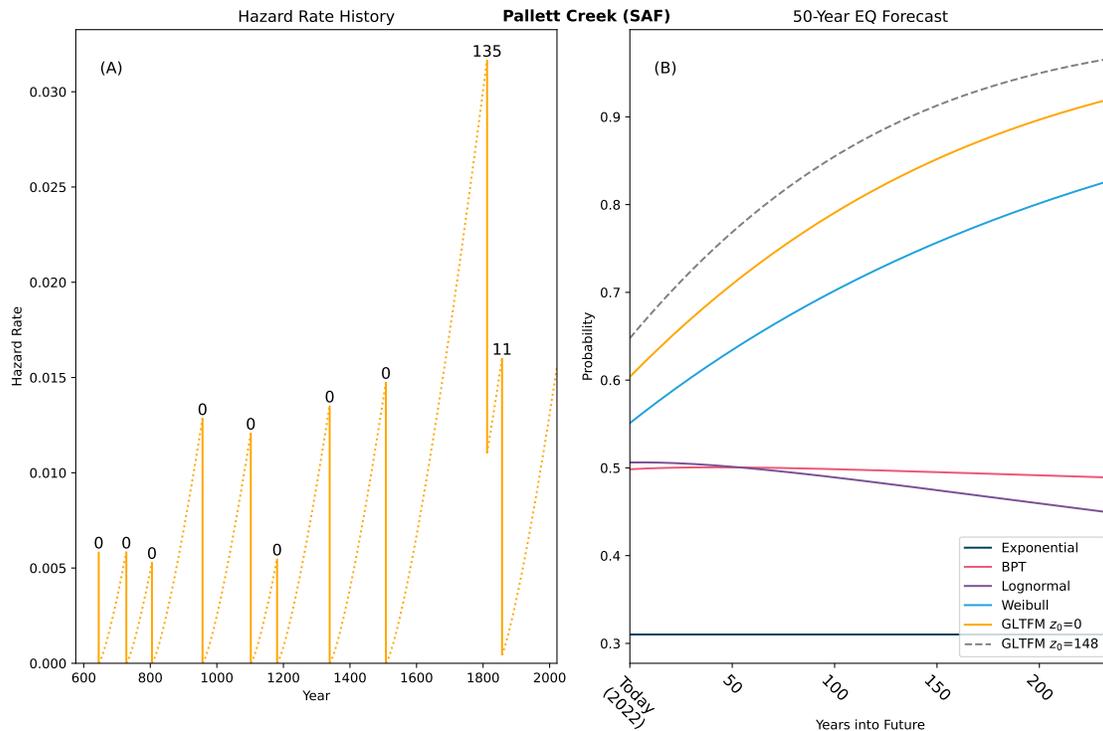


Figure 6.5. GLTFM applied to the Pallett Creek paleoseismic record. (A) Hazard rate curve history with the residual strain indicated above each earthquake. (B) Corresponding 50-year earthquake forecasts starting today for GLTFM and other models. The two GLTFM curves indicate different initial residual strain ( $z_0$ ) conditions.

with time? Strain is still accumulating so why should probability remain constant? In Chapter 5, I argued that the Markov LTFM's linear constant increase in probability increment made more physical sense. The GLTFM has this property as well and is even more flexible by allowing for any functional form of the change in probability with time (aka hazard rate function) although a monotonically increasing function (using the Weibull with  $\beta > 1$ ) is preferred.

Regardless of how exactly earthquake probability increases with time, the paleoseismic evidence (Wallace, 1970; Rockwell et al., 2000; Friedrich et al., 2003; Weldon et al., 2004;



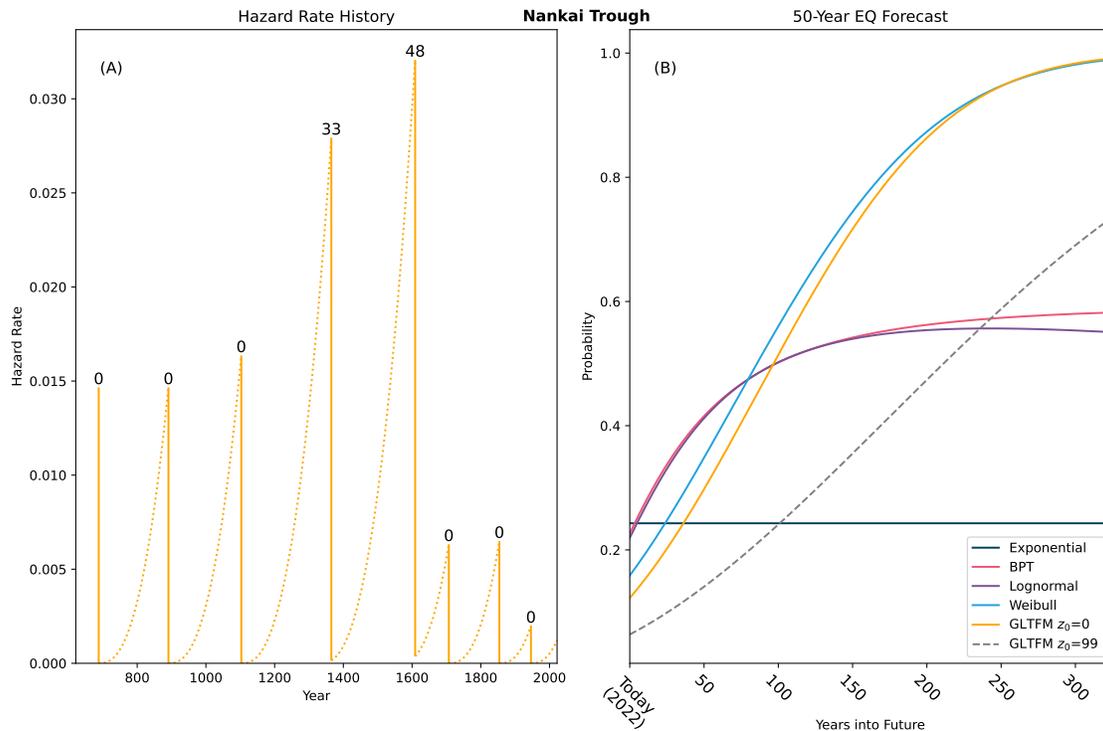


Figure 6.7. GLTFM applied to the Nankai Trough (section A-B) subduction zone paleoseismic record. (A) Hazard rate curve history with the residual strain indicated above each earthquake. (B) Corresponding 50-year earthquake forecasts starting today for GLTFM and other models. The two GLTFM curves indicate different initial residual strain ( $z_0$ ) conditions.

that LTFM forecasts the short 45-year inter-event time much better than the short-term memory models.

Ultimately, assessing which model performs best with these paleoseismic records runs into a fundamental problem: these records are simply too short and sparse to differentiate the models. Methods like the Brier Score (Brier, 1950)—a commonly used scoring method to assess probabilistic weather forecasts—do not perform well with very rare events (Benedetti, 2010) such as large earthquakes. One possible approach is to perform a joint likelihood analysis with all available paleoseismic records. This Bayesian approach

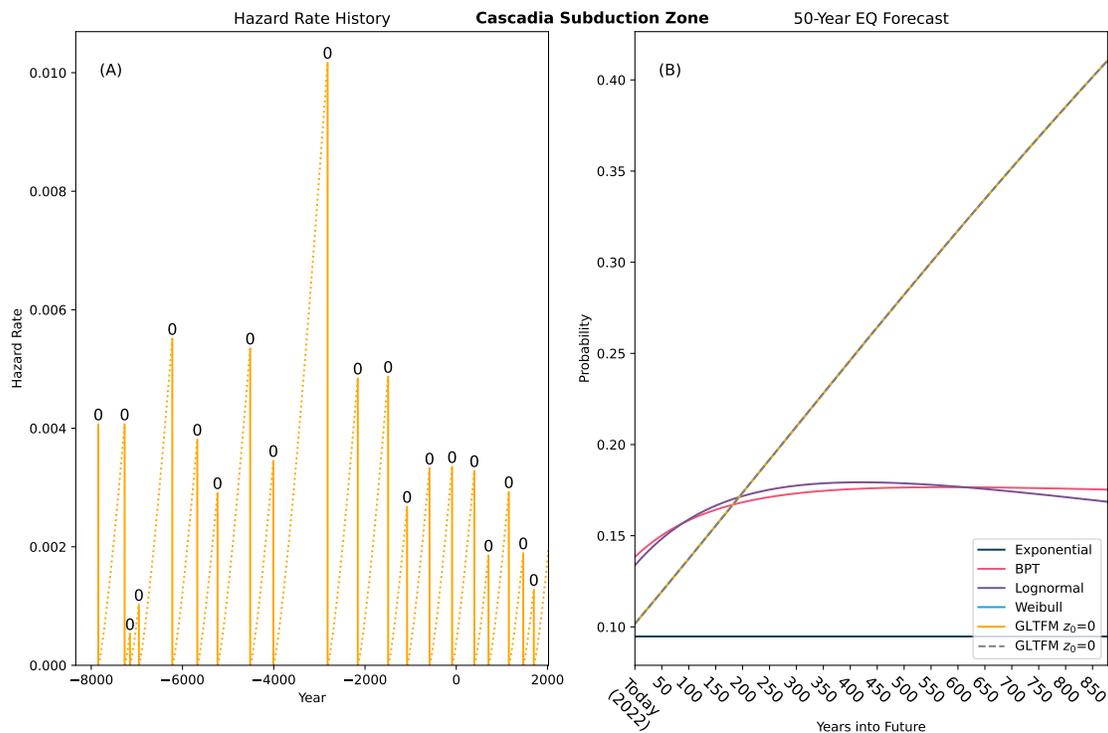


Figure 6.8. GLTFM applied to the Cascadia subduction zone paleoseismic record. (A) Hazard rate curve history with the residual strain indicated above each earthquake. (B) Corresponding 50-year earthquake forecasts starting today for GLTFM and other models. The two GLTFM curves indicate different initial residual strain ( $z_0$ ) conditions. The short-term memory Weibull model is indistinguishable from the GLTFM.

allows the specific parameters to vary for each paleoseismic record and then calculates the likelihood that a specific model produced all of the paleoseismic records. By jointly solving for all paleoseismic records, it leverages additional data to distinguish which model performs best.

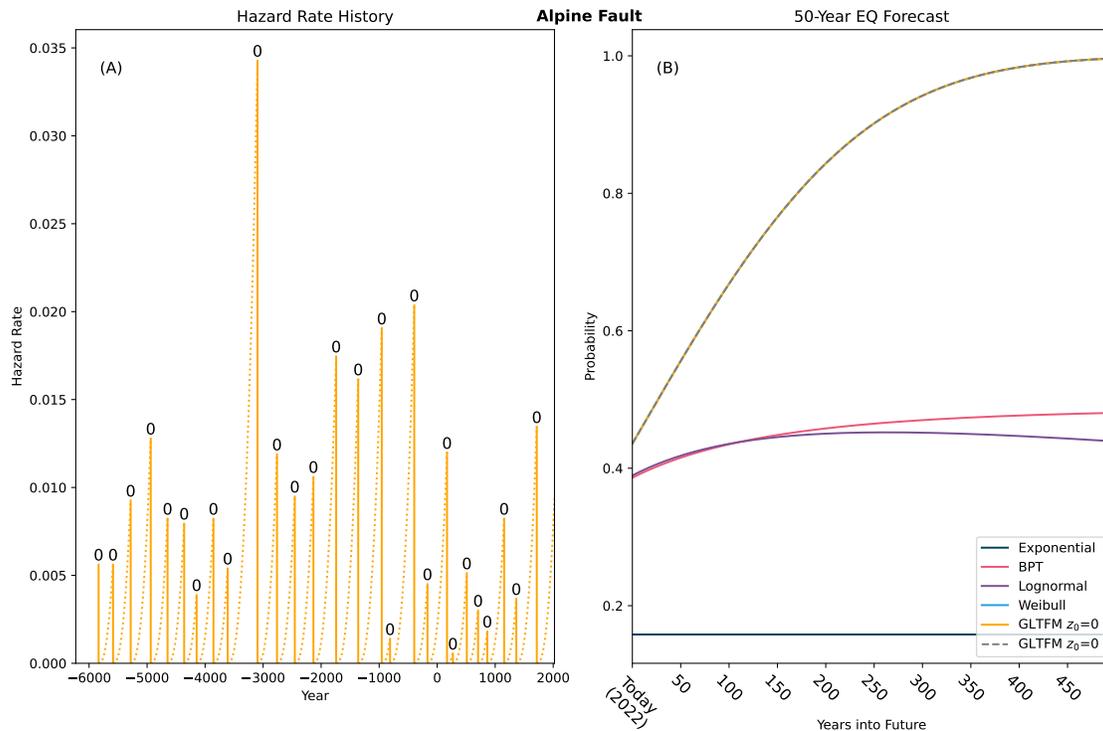


Figure 6.9. GLTFM applied to the Alpine fault paleoseismic record. (A) Hazard rate curve history with the residual strain indicated above each earthquake. (B) Corresponding 50-year earthquake forecasts starting today for GLTFM and other models. The two GLTFM curves indicate different initial residual strain ( $z_0$ ) conditions. The short-term memory Weibull model is indistinguishable from the GLTFM.

Laboratory-based experiments may also prove to be fruitful for further analysis. These experiments simulate earthquakes by sliding various materials across each other at a constant loading rate. Most of the time the materials are locked together (like the interseismic period), but eventually the accumulated strain overcomes the frictional strength of the materials, and the blocks move (like an earthquake). Unlike with natural earthquakes, these lab earthquake data sets can contain far more events for analysis. Prior

studies suggest that memoryless models often perform well at forecasting these inter-event times (Rubinstein et al., 2012) but others suggest that there may be memory in the observed earthquake sequences. Beeler et al. (2014) suggest the stress drops of consecutive earthquakes are strongly correlated possibly suggesting that there may be memory of prior slip in lab experiments. Although lab earthquakes are not perfect analogs for real earthquakes, they may provide some insights into influence of long-term fault memory on earthquake recurrence.

### 6.7. Conclusion

GLTFM is a powerful, yet simple to implement, earthquake probability model that better reflects our understanding of the strain accumulation and release processes that drive earthquakes. Unlike the most commonly used earthquake probability models, GLTM can estimate whether an earthquake likely left residual strain along a fault and can incorporate this knowledge into earthquake forecasts. Although the GLTFM suggests that residual strain is relatively rare along faults, when it does occur, it raises the estimated earthquake probability in the near term by making shorter inter-event times more likely. Paleoseismic records are too short to quantitatively prove that GLTFM performs better than other models, but joint likelihood approaches and analyses of simulated earthquakes from laboratory experiments may provide insight into this question.

## APPENDIX A

**Derivations for Markov LTFM Formulation****A.1. Fitting the LTFM to Data**

We calculate the LTFM probabilities in Chapter 5 using the following steps. The corresponding Appendix section or equations are indicated for each step.

1. Construct the matrix of transition probabilities  $\mathbf{P}$  using best fitting number of states and earthquake drop size (Section A2).

2. Calculate the vector  $\pi$  of stationary probabilities of being in different states (Equation A.2).

3. Calculate the values in Equations (A.5) through (A.10) and use these values as input to calculate the distribution of inter-event times  $r_{T|1}$  (Equation (A.12)). This is the long-run LTFM-LR PDF.

4. Calculate the values in Equations (A.13), (A.14), (A.17), and (A.19). Use these values as input to calculate the distribution of inter-event times conditioned on the most recent inter-event time  $r_{T|1, T_m}$  (Equation (A.21)). This is the short-term forecast LTFM-P1 for the next earthquake. To calculate the inter-event time distribution conditioned on additional prior inter-event times (LTFM-P2, LTFM-P3, etc.), use Equation (A.27). For the Pallett Creek data,  $T_m = 45$ .

5. To calculate the LTFM hazard curves, use Equation (A.38) with  $T' = 1$ . For the long-run LTFM-LR curve, use  $s_{k|T}$  from Equation (A.14) as input. For the LTFM-P1 curve use  $s_{k|T,T_m}$  from Equation (A.23) with  $T_m = 45$  for the Pallett Creek data.

6. To calculate the conditional probability of an earthquake in the next 30 years for Pallett Creek, use Equation (A.38) with  $T' = 30$ . For the long-run LTFM-LR curve, use  $s_{k|T}$  from Equation (A.14) as input. For the LTFM-P1 curve use  $s_{k|T,T_m}$  from Equation (A.23) with  $T_m = 45$ .

We used a maximum likelihood estimation (MLE) to find the best combination of LTFM parameters (number of states  $N$  and earthquake drop size  $R$ ) to fit the sequence of inter-event times in Table 5.1. We did not include the current half-open inter-event interval (1857 to today). The likelihood is found by evaluating Equation (A.27) for each inter-event time, starting with the first in the sequence, and then multiplying the probabilities. For the Pallett Creek data,  $m = 10$  and  $T_1 = 83, T_2 = 77, \dots, T_8 = 304, T_9 = 45$ . For the Pallett Creek sequence, the likelihood equals  $r_{T_1|1} \times r_{T_2|1,T_1} \times \dots \times r_{T_9|1,T_1,T_2,T_3,T_4,T_5,T_6,T_7,T_8}$ . To find the parameter values that maximize the likelihood we used a grid search, with number of states spanning the range from 1000 to 25000 in increments of 250 and earthquake drop size from 50 to 1000 in increments of 25. For the exponential, lognormal, and Brownian Passage Time models, we used the fit function from Python's `scipy.stats` package to find the best parameters through MLE. We again do not include the current open inter-event interval. Conditional probabilities for the exponential, lognormal, and Brownian Passage Time are calculated using Equation (17) from Matthews et al. (2002).

The following section provides the equations (and corresponding derivations) needed to calculate earthquake probabilities using the Markov LTFM described in Chapter 5.

## A.2. The Model

The process occurs in discrete time epochs, denoted by  $t = 0, 1, \dots$ . The state of the process refers to the amount of accumulated strain, denoted by  $S_t$  for epoch  $t$ . Strain is non-negative and increases by a constant amount  $A$  in each epoch unless an earthquake occurs, in which case some or all of the accumulated strain is released. We assume  $A = \frac{1}{N-1}$  for some integer  $N \geq 3$ . Once the accumulated strain is large enough, an earthquake is certain to occur and release some or all of the strain. Thus,  $S_t$  is bounded above, and without loss of generality we assume the measurement scale for strain is such that the upper bound equals 1. Strain is restricted to values  $0, A, 2A, 3A, \dots, 1$  with associated “states”  $1, 2, \dots, N$ . Thus, state 1 has strain 0, state 2 has strain  $A$ , state 3 has strain  $2A$ , and so on, up to state  $N$  with strain 1.

The probability of an earthquake for any state is assumed to be proportional to the strain, so the state variable can be considered to be either probability or strain.

To consider when earthquakes occur, note that the model allows for thresholds to limit when earthquakes can or cannot occur. When the strain  $S_t$  drops to 0 or below a threshold  $\delta \geq 0$ , the probability of an earthquake equals zero. Otherwise, it is equal to  $S_t$ .

Define the random variable  $I_t$  to indicate whether an earthquake occurs at epoch  $t$  ( $I_t = 1$ ) or not ( $I_t = 0$ ). Then the conditional probability of an earthquake, given that it is in state  $S_t$  is

$$(A.1) \quad \Pr [I_t = 1 | S_t] = \begin{cases} \min \{S_t, 1\} & \text{if } \delta \leq S_t \\ 0 & \text{if } S_t < \delta \end{cases}$$

If no earthquake occurs at epoch  $t$ , the strain increases by  $A$  so  $S_{t+1} = S_t + A$ . If an earthquake occurs at epoch  $t$ , strain is released. The strain released equals  $R$  unless  $R > S_t$  in which case the strain released is  $S_t$ . Thus we have

$$(A.2) \quad S_{t+1} = \begin{cases} \max\{0, S_t - R\} & \text{if } I_t = 1 \\ S_t + A & \text{if } I_t = 0 \end{cases}$$

In the numerical results in Chapter 5,  $R$  is fixed and equal to a positive integer multiple of  $A$ . However, the model allows a more general formulation, so that  $R$  can assume values  $A, 2A, 3A, \dots, S_t$  according to a specified probability distribution.

The following Table A.1 provides a glossary of notation.

Table A.1. LTFM notation glossary.

Symbol	Interpretation	Section or Equation
$I_t$	Indicator of earthquake (1) or no earthquake (0) at epoch $t$	A1
$A$	Strain increment if no earthquake	A1
$S_t$	Accumulated strain at epoch $t$	A1
$\delta$	Strain threshold for earthquake to occur, i.e., no earthquake if $S_t < \delta$	A1
$R$	Strain drop if earthquake occurs and $S_t \geq R$	A1
$p_{j,k}$	Transition probability, i.e., probability of going to state $k$ at next epoch given system is in state $j$	A2
$\mathbf{P}$	Matrix of transition probabilities	A2

Continued on next page

Table A.1 – continued from previous page

Symbol	Interpretation	Section or Equation
$\pi$	Vector of long-run probabilities of being in various states	A.3
$e_k$	Conditional probability of earthquake in epoch $t$ given system is in state $k$ in epoch $t$	A.4
$w_{k,T}$	Conditional probability that the next earthquake occurs exactly $T$ epochs from the current epoch given that the current state is $k$	A.5
$v_{k,T}$	Conditional probability that the next earthquake occurs $T$ or more epochs from the current epoch given that the current state is $k$	A.6
$w_T$	Unconditional probability that the next earthquake occurs exactly $T$ epochs from the current epoch	A.7
$v_T$	Unconditional probability that the next earthquake occurs $T$ or more epochs from the current epoch	A.8
$d_{k 0}$	Conditional probability of being in state $k$ at epoch $t$ given earthquake at epoch $t$	A.9
$d_{k 0,T}$	Conditional probability of being in state $k$ at epoch $t$ given last 2 earthquakes occurred in epochs $t$ and $t - T$	A.17
$d_{k 0,T_1,\dots,T_{m-1}}$	Conditional probability of being in state $k$ at epoch $t$ given last $m$ earthquakes occurred in epochs $t, t - T_1, t - T_1 - T_2, \dots, t - \sum_{i=1}^{m-1} T_i$ . Thus, the $T_i$ 's are inter-event times.	A.24

---

Continued on next page

Table A.1 – continued from previous page

Symbol	Interpretation	Section or Equation
$s_{k 1}$	Conditional probability of being in state $k$ at epoch $t$ given last earthquake at epoch $t - 1$	A.10
$s_{k T}$	Conditional probability of being in state $k$ at epoch $t$ given last earthquake at epoch $t - T$	A.14
$s_{k T, T_1, \dots, T_{m-1}}$	Conditional probability of being in state $k$ at epoch $t$ given last $m$ earthquakes prior to $t$ occurred at epochs $t - T, t - T - T_1, \dots, t - T - \sum_{i=1}^{m-1} T_i$ . Thus, the $T_i$ 's are inter-event times.	A.29
$r_{T 1}$	Conditional probability that next earthquake occurs in epoch $t + T$ given that last earthquake occurred in epoch $t$	A.12
$r_{T 1, T_1}$	Conditional probability that next earthquake occurs in epoch $t + T$ given that last 2 earthquakes occurred in epochs $t$ and $t - T_1$	A.21
$r_{T 1, T_1, \dots, T_{m-1}}$	Conditional probability that next earthquake occurs in epoch $t + T$ given that last $m$ earthquakes occurred in epochs $t, t - T_1, t - T_1 - T_2, \dots, t - \sum_{i=1}^{m-1} T_i$	A.27
$q_{T 1}$	Conditional probability that next earthquake occurs in epoch $t + T$ or later given that last earthquake occurred in epoch $t$	A.13
$q_{T 1, T_1}$	Conditional probability that next earthquake occurs in epoch $t + T$ or later given that last 2 earthquakes occurred in epochs $t$ and $t - T_1$	A.22

---

Continued on next page

Table A.1 – continued from previous page

Symbol	Interpretation	Section or Equation
$q_{T 1,T_1,\dots,T_{m-1}}$	Conditional probability that next earthquake occurs in epoch $t + T$ or later given that last $m$ earthquakes occurred in epochs $t, t - T_1, t - T_1 - T_2, \dots, t - \sum_{i=1}^{m-1} T_i$	A.28
$\tau_k$	Expected waiting time until the next earthquake given that the process in state $k$	A.30
$\sigma_{\tau,k}^2$	Conditional variance of the waiting time until the next earthquake given that the process in state $k$	A.31
$\tau$	Unconditional expected waiting time until the next earthquake	A.32
$\sigma_{\tau}^2$	Unconditional variance of waiting time until the next earthquake	A.33
$\tau_{ T,T_1,T_2,\dots,T_{m-1}}$	Conditional expectation of waiting time until next earthquake given that last $m$ earthquakes occurred in epochs $t, t - T_1, t - T_1 - T_2, \dots, t - \sum_{i=1}^{m-1} T_i$	A.34
$\sigma_{\tau T,T_1,T_2,\dots,T_{m-1}}^2$	Conditional variance of waiting time until next earthquake given that last $m$ earthquakes occurred in epochs $t, t - T_1, t - T_1 - T_2, \dots, t - \sum_{i=1}^{m-1} T_i$	A.35
$F_{T',k}$	Conditional probability of at least 1 earthquake in the next $T'$ epochs given the current state is $k$	A.36
$F_{T'}$	Unconditional probability of at least 1 earthquake in the next $T'$ epochs	A.37

---

Continued on next page

Table A.1 – continued from previous page

Symbol	Interpretation	Section or Equation
$F_{T' T, T_1, T_2, \dots, T_{m-1}}$	Conditional probability of at least 1 earthquake in the next $T'$ epochs given that the previous $m$ earthquakes occurred $T$ epochs ago, $T + T_1$ epochs ago, ..., $T + \sum_{i=1}^{m-1} T_i$	A.38
$\mu_{\text{renew}}$	Expected number of epochs for the process to renew (return to state 1 from state 1)	A.42
$\sigma_{\text{renew}}$	Standard deviation of number of epochs for the process to renew (return to state 1 from state 1)	A.43
$\text{cv}_{\text{renew}}$	Coefficient of variation (cv) of number of epochs for the process to renew (return to state 1 from state 1)	A.44

### A.3. Transition Probability and Stationary Distribution

The conditional probability of transition to state  $k$  at epoch  $t$  given that the process is in state  $j$  is called a transition probability, denoted by  $p_{jk}$  or  $p_{j,k}$ . The matrix  $\mathbf{P}$  of transition probabilities  $p_{jk}$  is called the transition probability matrix and its entries satisfy  $0 \leq p_{jk} \leq 1$  and  $1 = \sum_k p_{jk}$  for each  $j$ . The LTFM rules in Equations (A.1) and (A.2) specify that for  $k \geq j$ ,

$$p_{j,k} = \begin{cases} (j-1)A & \text{if } k = 1 \text{ and } j \leq D \\ (j-1)A & \text{if } k = j - D > 0 \\ 0 & \text{otherwise, and } k < j \end{cases}$$

From Markov chain theory, the conditional probability of transition to state  $k$  at epoch  $t + m$  given that the process is in state  $j$  at epoch  $t$  given by the  $(j, k)$  entry in the  $m$ -fold matrix product  $\mathbf{P}^m$ . Standard references on Markov chains include Feller (1968) and Çinlar (1975).

Another result from Markov chain theory applicable to our model is the existence of a long-run or stationary probability distribution over the states. The probability distribution is given by a row vector  $\pi$  with elements between 0 and 1 and summing to 1 such that

$$(A.3) \quad \pi \mathbf{P} = \pi$$

The vector  $\pi$  is called the stationary distribution of the Markov chain. We interpret  $\pi_j$  as the long-run probability of being in state  $j$ . Stationary distributions only exist if the Markov chain is irreducible and regular, meaning it is possible to go from every state to every other state (not necessarily in one epoch or “step”) and for some  $n \geq 1$  it is possible to go from any state to any other state in exactly  $n$  epochs (or steps). Equivalently, for some  $n \geq 1$  every element of  $\mathbf{P}^n$  is positive. Those conditions are met by the LTFM.

In application, the numerical value of an epoch,  $t$ , will depend not on how long the process has been ongoing but on an arbitrary reference period, such as the date of the earliest observed earthquake, which might be considered epoch 0, for example. This perspective is justified by the time-independence of the transition probability matrix  $\mathbf{P}$  and the existence of the stationary distribution  $\pi$ .

The conditional probability of an earthquake occurring given its state was shown in Equation A.1 and can be expressed in terms of the elements of  $\mathbf{P}$ . We denote the

conditional probability of an earthquake occurring in an epoch given it is in state  $k$  in the epoch by  $e_k$ . Because transitions to lower states indicate an earthquake occurred, it follows that

$$(A.4) \quad e_k = 1 - p_{k,k+1}$$

#### A.4. Conditional Probability of an Earthquake Occurring in $T$ Epochs

##### A.4.1. Conditional probability that the next earthquake occurs exactly $T$ epochs from the current epoch given system is currently in state $k$

We define  $w_{k,T}$  as the conditional probability that the next earthquake occurs exactly  $T$  epochs from the current epoch given that the current state is  $k$ . The maximum number of epochs to wait equals the number of total states,  $N$ , so  $w_{k,T} = 0$  for  $T > N$ . An earthquake first occurs at epoch  $T$  if one does not occur at intermediate epochs and then one occurs at  $T$ , so

$$(A.5) \quad w_{k,T} = \begin{cases} e_k & \text{if } T = 1 \\ e_{k+T-1} \prod_{t=0}^{T-2} (1 - e_{k+t}) & \text{if } 1 < T \leq N \end{cases}$$

The  $N \times N$  matrix  $\mathbf{W}$  with elements  $w_{k,T}$  ( $k = 1, \dots, N; T = 1, \dots, N$ ) will be called the conditional waiting time matrix. The  $k^{\text{th}}$  row of  $\mathbf{W}$  contains the conditional probability distribution for waiting times given that the process is in state  $k$ .

**A.4.2. Conditional probability that the next earthquake occurs  $T$  or more epochs from the current epoch given system is currently in state  $k$**

The conditional probability that the next earthquake occurs  $T$  or more epochs from the current epoch given that the current state is  $k$  is given by

$$(A.6) \quad v_{k,T} = \sum_{\tau \geq T} w_{k,T}$$

Notice that  $v_{k,1} = 1$  for any state  $k$ .

**A.5. Unconditional Probability of an Earthquake Occurring in  $T$  Epochs**

**A.5.1. Unconditional probability that the next earthquake occurs exactly  $T$  epochs from current epoch**

The unconditional probability that next earthquake occurs exactly  $T$  epochs from the current epoch is equal to the weighted average of  $w_{k,T}$  terms, with weights equal to the probability of being in state  $k$ . The latter probabilities are not known but can be approximated by the long-run probabilities  $\pi_k$ . This leads us to approximate the unconditional probability by

$$(A.7) \quad w_T = \sum_{k=1}^N \pi_k w_{k,T}$$

The vector  $\mathbf{w}$  values  $w_1, \dots, w_N$  will be referred to as the unconditional waiting time vector. It contains the approximate unconditional distribution of the waiting time until the next earthquake.

### A.5.2. Unconditional probability that the next earthquake occurs $T$ or more epochs from current epoch

The approximate unconditional probability  $v_T$  that the next earthquake occurs  $T$  or more epochs from the current epoch is given by

$$(A.8) \quad v_T = \sum_{\tau \geq T} w_\tau$$

Note that  $v_T = 1$ .

## A.6. Conditional Probabilities Over States Given Recent Earthquake

### A.6.1. Conditional probability of being in state $k$ at epoch $t$ given an earthquake at epoch $t$

We are interested in the conditional probability that the system was in state  $k$  at epoch  $t$  given that an earthquake occurred in epoch  $t$ . To calculate this probability, we use the relationship  $P(A|B) = P(B|A)P(A)/P(B)$  with  $A$  denoting the event that the system was in state  $k$  at epoch  $t$  and  $B$  denoting the event that an earthquake occurred at epoch  $t$ . We approximate  $P(A)$  by the long-run probability  $\pi_k$  and we approximate  $P(B)$  by  $w_1$ . The conditional probability of an earthquake  $P(B|A)$  is given by  $e_k$ . Putting this together yields the approximation  $d_{k|0}$  for the conditional probability that the system was in state  $k$  at epoch  $t$  given that an earthquake occurred in epoch  $t$ ,

$$(A.9) \quad d_{k|0} = \frac{e_k \pi_k}{w_1}$$

To verify that the approximation yields a conditional probability distribution, note that  $d_{k|0} > 0$  and  $\sum_k d_{k|0} = 1$ .

### A.6.2. Conditional probability of being in state $k$ at epoch $t + 1$ given an earthquake at epoch $t$

We also want to know the conditional probability of being in state  $k$  at epoch  $t + 1$  given occurrence of an earthquake epoch  $t$ . We approximate this probability by  $s_{k|1}$  given by

$$(A.10) \quad s_{k|1} = \sum_{j>k} p_{j,k} d_{j|0} / e_j = \sum_{j>k} p_{j,k} \pi_j / w_1$$

Note that  $s_{k|1}$  is also the approximate conditional probability of being in state  $k$  at epoch  $t$  given occurrence of an earthquake epoch in epoch  $t - 1$ . To verify that the approximation yields a conditional probability distribution, note that  $s_{k|1} \geq 0$  and  $\sum_k s_{k|1} = 1$ .

To derive Equation (A.10), note that to be in state  $k$  at epoch  $t + 1$  after an earthquake occurred at epoch  $t$ , the system must have been in a state  $j$  that allowed a transition to state  $k$  by release of strain. The law of total probability implies the relationship

$$(A.11) \quad P(A|B) = \sum_j P(A|B, C_j) P(C_j|B)$$

and we let  $A$  denote the event that the system is in state  $k$  at epoch  $t + 1$ ,  $B$  denote the event that an earthquake occurred in epoch  $t$ , and  $C_j$  denote the event that the system was in state  $j$  at epoch  $t$ . From the relationship  $P(A|B, C_j) = P(A, B|C_j) / P(B|C_j)$ , we approximate  $P(A|B, C_j)$  by  $p_{j,k} / e_j$  if  $j > k$  and  $P(A|B, C_j)$  if  $j \leq k$ . Using Equation

(A.9) to approximate  $P(C_j|B)$  by  $d_{j|0}$ , we approximate  $P(A|B)$  by  $\sum_{j>k} p_{j,k} d_{j|0}/e_j$  which equals  $\sum_{j>k} p_{j,k} \pi_j/w_1$ .

## A.7. Probability that the Next Earthquake Occurs $T$ or More Epochs After the Last Earthquake

### A.7.1. Probability that the next earthquake occurs exactly $T$ epochs after the last earthquake

Consider the conditional probability that the next earthquake occurs at epoch  $t+T$  given that the last earthquake occurred at epoch  $t$ . Since the process is stationary, this equals the conditional probability that the next earthquake occurs exactly  $T$  epochs after the last earthquake. This conditional probability is approximately equal to  $r_{T|1}$  defined as

$$(A.12) \quad r_{T|1} = \sum_j w_{j,T} s_{j|1}$$

Observe that  $r_{T|1} \geq 0$  for  $1 \leq T \leq N$ ,  $r_{T|1} = 0$  for  $T > N$  and  $T < 1$ , and  $\sum_T r_{T|1} = 1$ . Thus, we can treat the set of conditional probabilities  $r_T$  as a probability mass function for the different values of  $T$ , and comparable to discretized versions of distributions commonly used in renewal models for earthquake occurrence (exponential, lognormal Brownian Passage Time, etc.). The renewal model closest to the LTFM is the Weibull distribution with shape parameter equal to 2, for which probability increments for earthquake occurrences are constant and positive.

Derivation of Equation (A.12). Let  $A$  denote the event that an earthquake occurs at epoch  $t+T$  but not at epochs  $t+1, t+2, \dots, t+T-1$ . Let  $B$  denote the event that an earthquake occurred in epoch  $t$ . Thus, the probability of interest is  $P(A|B)$ . To obtain

this, we use Equation (A.11) with  $C_j$  denoting the event that the system was in state  $j$  at epoch  $t$ . Recall that  $P(A|B, C_j)$  is approximated by  $w_{j,T}$  and  $P(C_j|B)$  is approximated by  $s_j$ . It follows from Equation (A.11) that the probability of waiting exactly  $T$  epochs from the last earthquake until the next earthquake is approximated by  $r_{T|1}$  as given by Equation (A.12).

### A.7.2. Probability that the next earthquake occurs $T$ or more epochs after the last earthquake

The conditional probability that the next earthquake occurs  $T$  or more epochs after the last earthquake is approximately equal to  $q_{T|1}$  defined as

$$(A.13) \quad q_{T|1} = \sum_{\tau \geq T} r_{\tau|1}$$

Note that  $q_{1|1} = 1$ .

### A.8. Conditional Distribution over States Given Date of Last Earthquake

We are also interested in the conditional distribution over states given that the most recent earthquake occurred  $T$  epochs ago. We will calculate the conditional probability of being in state  $k$  at epoch  $t+T$  given that an earthquake occurred in epoch  $t$  and there was no earthquake in the  $T-1$  succeeding epochs,  $t+1, \dots, t+T-1$ . In other words, the most recent earthquake occurred  $T$  epochs ago. This conditional probability is approximately equal to  $s_{k|T}$  defined as

$$(A.14) \quad s_{k|T} = v_{k-T+1, T} s_{k-T+1|1} / q_{T|1}$$

Observe that for a given  $T$ ,  $s_{k|T} \geq 0$  and  $\sum_k s_{k|T} = 1$ .

We now derive Equation (A.14). Let  $A$  denote the event of being in state  $k$  at epoch  $t + T$ , let  $B$  denote the event of an earthquake occurring in epoch  $t$ , let  $C$  denote the event of no earthquake in the  $T$  epochs  $t + 1, \dots, t + T$ , and let  $D_j$  denote the event of being in state  $j$  in epoch  $t + 1$ . The probability of interest is  $P(A|B, C)$ . The law of total probability and the properties of conditional probability imply

$$\begin{aligned}
 (A.15) \quad P(A|B, C) &= \sum_j P(A|B, C, D_j)P(D_j|B, C) \\
 &= \sum_j P(A|B, C, D_j) \frac{P(C|B, D_j)P(D_j|B)P(B)}{P(C|B)P(B)}
 \end{aligned}$$

Given that there was no earthquake for  $T$  epochs following epoch  $t$ , the only way to be in state  $k$  at epoch  $t + T$  is to have been in state  $k - 1$  at epoch  $t + T - 1$ , in state  $k - 2$  at epoch  $t + T - 2, \dots$ , and in state  $k - T + 1$  at epoch  $t + 1$ . Furthermore, given that there was no earthquake for  $T$  epochs following epoch  $t$ , if the system was in state  $k - T + 1$  at epoch  $t + 1$ , it was necessarily in state  $k$  at epoch  $t + T$ . Thus,  $P(A|B, C, D_j)$  equals 1 if  $j = k - T + 1$  and it equals 0 if  $j \neq k - T + 1$ . It follows from Equation (A.15) that

$$(A.16) \quad P(A|B, C) = \frac{P(C|B, D_{k-T+1})P(D_{k-T+1}|B)}{P(C|B)}$$

The denominator,  $P(C|B)$  is approximated by  $q_t$ . In the numerator,  $P(C|B, D_{k-T+1})$  is approximated by  $v_{k_{T+1}, T}$ . The other term in the numerator,  $P(D_{k-T+1}|B)$  is approximated by  $s_{k-T+1|1}$ . Thus,  $P(A|B, C)$  is approximately equal to Equation (A.14).

## A.9. Conditional Probabilities Given Dates of Past 2 Earthquakes

### A.9.1. Conditional probability of being in state $k$ at epoch $t$ given that the last 2 earthquakes occurred in epochs $t$ and $t - T_1$

Because the Markov LTFM does not reset after each earthquake, the conditional probabilities based on the two most recent earthquakes are not equal to those conditional on just the most recent earthquake. The conditional probability that the system was in state  $k$  at epoch  $t$  given that an earthquake occurred in epoch  $t$  (0 epochs ago) and the previous earthquake occurred in epoch  $t - T_1$  (or  $T_1$  epochs ago, i.e.,  $T_1$  epochs between earthquakes) is approximately equal to  $d_{k|0,T_1}$  defined as

$$(A.17) \quad d_{k|0,T_1} = \frac{e_k s_{k|T_1}}{\sum_j e_j s_{j|T_1}}$$

To verify that the approximation yields a conditional probability distribution for each  $T_1$ , note that for any fixed  $T_1$ ,  $d_{k|0,T_1} \geq 0$ , and  $\sum_k d_{k|0,T_1} = 1$ .

Derivation of Equation (A.17). Let  $A$  denote the event of being in state  $k$  at epoch  $t$ , let  $B$  denote the event that an earthquake occurs in epoch  $t$ , let  $C$  denote the event of there are exactly  $T_1$  epochs between two successive earthquakes, and let  $D$  denote the event that an earthquake occurs in epoch  $t - T_1$ . We have

$$(A.18) \quad \begin{aligned} P(A|B, C, D) &= \frac{P(B|A, C, D)P(A|C, D)P(C, D)}{P(B|C, D)P(C, D)} \\ &= \frac{P(B|A, C, D)P(A|C, D)}{P(B|C, D)} \end{aligned}$$

The conditional probability  $P(B|A, C, D)$  is equal to  $P(B|A)$  by the Markov property, and is equal to  $e_k$ . Next,  $P(A|C, D)$  is approximately equal to  $s_{k|T_1}$ . The numerator is

thus approximately  $e_k s_{k|T_1}$ . The denominator  $P(B|C, D)$  equals the sum of the numerator over all states, and dividing the numerator by the denominator yields Equation (A.17).

### A.9.2. Conditional probability of being in state $k$ at epoch $t + 1$ given that the last 2 earthquakes occurred in epochs $t$ and $t - T_1$

We also want to know the conditional probability of being in state  $k$  at epoch  $t + 1$  given an earthquake occurred in epoch  $t$  and the earthquake before that occurred at epoch  $t - T_1$ . Notice that the distance between the current epoch and the epochs with earthquakes on which the conditioning occurs takes the values 1 and  $T_1 + 1$ . This conditional probability is approximately equal to  $s_{k|1, T_1+1}$  given by

$$(A.19) \quad \begin{aligned} s_{k|1, T_1+1} &= \sum_{j>k} p_{j,k} d_{k|0, T_1} / e_j \\ &= \frac{\sum_{j>k} p_{j,k} s_{j|T_1}}{\sum_l e_l s_{l|T_1}} \end{aligned}$$

To explain the notation, we use  $s_{k|T, T_1, \dots, T_{m-1}}$  to denote the probability of being in state  $k$  given that the latest earthquake occurred  $T$  epochs ago, the second-to-latest earthquake before that occurred  $T + T_1$  epochs ago, and so on, and the  $m^{\text{th}}$ -to-latest earthquakes occurred  $T + T_1 + \dots + T_{m-1}$  epochs ago. Thus, the spacings between the earthquakes, going backward in time, are  $T, T_1, T_2, \dots, T_{m-1}$ .

To derive Equation (A.19), note that in order to be in state  $k$  at epoch  $t + 1$  after an earthquake occurred at epoch  $t$ , the system must have been in a state  $j$  that allowed a transition to state  $k$  by release of strain. Let  $A$  denote the event that the system is in state  $k$  at epoch  $t + 1$ ,  $B$  denote the event that the last 2 earthquakes occurred 1 epoch

ago and  $T_1 + 1$  epochs ago, and  $C_j$  denote the event that the system was in state  $j$  at epoch  $t$ . From Equation (A.11) it now follows that

$$(A.20) \quad P(A|B) = \sum_{j>k} P(A|B, C_j)P(C_j|B)$$

From the relationship  $P(A|B, C_j) = P(A, B|C_j)/P(B|C_j)$  we approximate  $P(A|B, C_j)$  by  $p_{j,k}/e_j$  if  $j > k$  and  $P(A|B, C_j) = 0$  if  $j \leq k$ . Using Equation (A.17) to approximate  $P(C_j|B)$  by  $d_{j|0, T_1}$ , we approximate  $P(A|B)$  by  $\sum_{j>k} p_{j,k}d_{j|0, T_1}/e_j$ , which is equal to  $\text{sum}_{j>k} p_{j,k}s_{j|T_1} / \sum_l e_l s_{l|T_1}$ .

### A.9.3. Conditional probability the next earthquake occurs in epoch $t+T$ given that the last 2 earthquakes occurred in epochs $t$ and $t - T_1$

Consider the conditional probability that the next earthquake occurs in epoch  $t+T$  given that the last earthquake occurred in epoch  $t$  and the earthquake prior to that occurred in epoch  $t - T_1$ . This probability is approximated by  $r_{T|1, T_1}$  defined as

$$(A.21) \quad r_{T|1, T_1} = \sum_k w_{k, T} s_{k|1, T_1}$$

Derivation of Equation (A.21). The conditional probability of an earthquake in epoch  $t+T$  given that the system is in state  $k$  at epoch  $t+1$  and that the past earthquake occurred in epoch  $t$  and the earthquake prior to that occurred in epoch  $t - T_1$  is equal to  $w_{k, t}$ , by the Markov property. The conditional probability that the system is in state  $k$  at epoch  $t+1$  given that the last earthquake occurred in epoch  $t$  and the earthquake prior to that occurred in epoch  $t - T_1$  is approximately equal to  $s_{k|1, T_1}$ . The product  $w_{k, T} s_{k|1, T_1}$  approximates the conditional probability that the system is in state  $k$  at epoch

$t + 1$  and an earthquake occurs in epoch  $t + T$  given that the past earthquake occurred in epoch  $t$  and the earthquake prior to that occurred in epoch  $t - T_1$ . By the law of total probability,  $r_{T|1,T_1}$  is approximated by the sum of  $w_{k,T}s_{k|1,T_1}$  over all states  $k$ , thus establishing Equation (A.21).

**A.9.4. Conditional probability that the next earthquake occurs in epoch  $t + T$  or later given the past 2 earthquakes occurred in epochs  $t$  and  $t - T_1$**

The conditional probability that the next earthquake occurs in epoch  $t + T$  or later given that the past earthquake occurred in epoch  $t$  and the earthquake prior to that occurred in epoch  $t - T_1$  is approximately equal to  $q_{T|1,T_1}$  defined as

$$(A.22) \quad q_{T|1,T_1} = \sum_{\tau \geq T} r_{\tau|1,T_1}$$

Because there must always be 1 epoch between earthquakes,  $q_{1|1,T_1} = 1$ .

**A.9.5. Conditional probability of being in state  $k$  at epoch  $t$  given that the past 2 earthquakes occurred in epochs  $t - T$  and  $t - T - T_1$**

The conditional probability of being in state  $k$  at epoch  $t$  given the past 2 earthquakes occurred at epochs  $t - T$  and  $t - T - T_1$  is approximately equal to  $s_{k|T,T_1}$  defined as

$$(A.23) \quad s_{k|T,T_1} = \frac{v_{k-T+1,T}s_{k-T+1|1,T_1}}{q_{T|1,T_1}}$$

For a given  $T$ ,  $\sum_k s_{k|T,T_1} = 1$ .

Derivation of Equation (A.23). Let  $A$  denote the event of being in state  $k$  at epoch  $t$  let  $B$  denote the event that no earthquake occurs in epochs  $t - T + 1, t - T + 2, \dots, t$ , let

$C$  denote the event that earthquakes occurred at epochs  $t - T$  and  $t - T - T_1$ , but not in between, and let  $D_j$  denote the event that the system is in state  $j$  at epoch  $t - T + 1$ .

The probability of interest is  $P(A|B, C, D_j)$  and it can be expressed as Equation (A.15). To be in state  $k$  at epoch  $t$  given  $B$  and  $C$ , the system must have been in state  $j = k - T + 1$  in epoch  $t - T + 1$ , which is the epoch immediately following the past earthquake. Thus,  $P(A|B, C, D_j) = 0$  unless  $j = k - T + 1$ , in which case  $P(A|B, C, D_j) = 1$ . It follows that equation (A.16) holds. Observe that  $P(B|C, D_{k-T+1})$  is equal to  $P(B|D_{k-T+1})$  by the Markov property and is thus approximated by  $v_{k-T+1, T}$ . Next,  $P(D_{k-t+1}|C)$  is approximated by  $s_{k-T+1|1, T_1}$ . Finally,  $P(B|C)$  is approximated by  $q_{T|1, T_1}$ . Equation (A.23) follows from substitution into Equation (A.16).

## A.10. Conditional Probabilities Based on the Past $m$ Earthquakes

We have derived  $d_{k|0, T_1}$ ,  $r_{T|1, T_1}$ ,  $q_{T|1, T_1}$ , and  $s_{k|T, T_1}$  in Equations (A.17), (A.21), (A.22), and (A.23). Those probabilities were conditional on 2 past earthquakes. Derivation of conditional probabilities given additional numbers of earthquakes proceeds inductively.

### A.10.1. Conditional probability of being in state $k$ at epoch $t$ given that the past $m$ earthquakes occurred in epochs $t, t - T_1, t - T_1 - T_2, \dots, t - \sum_{i=1}^{m-1} T_i$

Assume that conditional probabilities given the past  $m - 1$  earthquakes are known. Then the conditional probability of being in state  $k$  at epoch  $t$  given the past  $m$  earthquakes occurred in epochs  $t, t - T_1, t - T_1 - T_2, \dots, t - \sum_{i=1}^{m-1} T_i$  is approximately equal to  $d_{k|0, T_1, \dots, T_{m-1}}$

given by

$$(A.24) \quad d_{k|0,T_1,\dots,T_{m-1}} = \frac{e_k s_{k|T_1,\dots,T_{m-1}}}{\sum_j e_j s_{j|T_1,\dots,T_{m-1}}}$$

Derivation of Equation (A.24). Let  $A$  denote the event of being in state  $k$  at epoch  $t$ , let  $B$  denote the event that an earthquake occurs in epoch  $t$ , and let  $C$  denote the event that  $m - 1$  earthquakes occurred in the epochs from  $t - 1$  back to  $t - T_1 - T_2 \cdots T_{m-1}$ , and the epochs in which they occurred are  $t - T_1, t - T_1 - T_2, \dots, t - T_1 - T_2 \cdots T_{m-1}$ . We have

$$(A.25) \quad \begin{aligned} P(A|B, C) &= \frac{P(B|A, C)P(A|C)P(C)}{P(B|C)P(C)} \\ &= \frac{P(B|A, C)P(A|C)}{P(B|C)} \end{aligned}$$

The conditional probability  $P(B|A, C)$  is equal to  $P(B|A)$  by the Markov property, and is approximately equal to  $e_k$ . Next,  $P(A|C)$  is approximately equal to  $s_{k|T_1,\dots,T_{m-1}}$ . The numerator is thus approximately  $e_k s_{k|T_1,\dots,T_{m-1}}$ . The denominator  $P(B|C)$  equals the sum of the numerator over all states, and dividing the numerator by the denominator yields Equation (A.24).

### A.10.2. Conditional probability of being in state $k$ at epoch $t$ given that the past $m$ earthquakes occurred in epochs $t - 1, t - 1 - T_1, \dots, t - 1 - \sum_{i=1}^{m-1} T_i$

The conditional probability of being in state  $k$  at epoch  $t$  given that the past  $m$  earthquakes occurred at epochs  $t - 1, t - 1 - T_1, \dots, t - 1 - \sum_{i=1}^{m-1} T_i$  is approximately equal to

$s_{k|1,T_1,\dots,T_{m-1}}$  given for  $m > 1$  by

$$(A.26) \quad \begin{aligned} s_{k|1,T_1,\dots,T_{m-1}} &= d_{k|0,T_1,\dots,T_{m-1}} \sum_{j>k} p_{j,k}/e_j \\ &= \frac{\sum_{j>k} p_{j,k} s_{j|T_1,\dots,T_{m-1}}}{\sum_l e_l s_{l|T_1,\dots,T_{m-1}}} \end{aligned}$$

The case when  $m = 1$  is given by Equation (A.10) for  $s_{k|1}$ .

Derivation of Equation (A.26). Let  $A$  denote the event that the system is in state  $k$  at epoch  $t$  and let  $B$  denote the event that the past  $m$  earthquakes occurred in the epochs from  $t - 1$  back through  $t - 1 - \sum_{i=1}^{m-1} T_i$  and the epochs in which they occurred are  $t - 1, t - 1 - T_1, \dots, t - 1 - \sum_{i=1}^{m-1} T_i$ . Let  $C_j$  denote the event that the system was in state  $j$  at epoch  $t - 1$ . In order to be in state  $k$  at epoch  $t$  after an earthquake occurred at epoch  $t - 1$ , the system must have been in a state  $j$  that allowed a transition to state  $k$  by release of strain. From Equation (A.11) it thus follows that Equation (A.20) holds. From the relationship  $P(A|B, C_j) = P(A, B|C_j)/P(B|C_j)$ , we approximate  $P(A|B, C_j)$  by  $p_{j,k}/e_j$  if  $j > k$  and  $P(A|B, C_j) = 0$  if  $j \leq k$ . Using Equation (A.17) to approximate  $P(C_j|B)$  by  $d_{j|0,T_1,\dots,T_{m-1}}$ , we approximate  $P(A|B)$  by  $\sum_{j>k} p_{j,k} d_{j|0,T_1,\dots,T_{m-1}}/e_j$  which is equal to  $\sum_{j>k} p_{j,k} s_{j|T_1,\dots,T_{m-1}}/\sum_l e_l s_{l|T_1,\dots,T_{m-1}}$ .

### A.10.3. Conditional probability the next earthquake occurs in epoch $t+T$ given that the past $m$ earthquakes occurred in epochs $t, t - T_1, \dots, t - \sum_{i=1}^{m-1} T_i$

Consider the conditional probability that the next earthquake occurs in epoch  $t + T$  given that the past  $m$  earthquakes occurred in epoch  $t, t - T_1, \dots, t - \sum_{i=1}^{m-1} T_i$ . This probability

is approximated by  $r_{T|1,T_1,\dots,T_{m-1}}$  defined as

$$(A.27) \quad r_{T|1,T_1,\dots,T_{m-1}} = \sum_k w_{k,T} s_{k|1,T_1,\dots,T_{m-1}}$$

Derivation of Equation (A.27). The conditional probability of an earthquake in epoch  $t+T$  given that the system is in state  $k$  at epoch  $t+1$  and the last  $m$  earthquakes occurred in epoch  $t, t-T_1, \dots, t-\sum_{i=1}^{m-1} T_i$  is equal to  $w_{k,T}$  by the Markov property. The conditional probability that the system is in state  $k$  at epoch  $t+1$  given that the past  $m$  earthquakes occurred in epoch  $t, t-T_1, \dots, t-\sum_{i=1}^{m-1} T_i$  is approximately equal to  $s_{k|1,T_1,\dots,T_{m-1}}$ . The product  $w_{k,T} s_{k|1,T_1,\dots,T_{m-1}}$  approximates the conditional probability that the system is in state  $k$  at epoch  $t+1$  and the past  $m$  earthquakes occurred in the epochs from  $t$  back through  $t-\sum_{i=1}^{m-1} T_i$  and the epochs in which they occurred are  $t, t-T_1, \dots, t-\sum_{i=1}^{m-1} T_i$ . By the law of total probability,  $r_{T|1,T_1,\dots,T_{m-1}}$  is approximated by the sum of  $w_{k,T} s_{k|1,T_1,\dots,T_{m-1}}$  over all states  $k$  thus establishing Equation (A.27).

**A.10.4. Conditional probability the next earthquake occurs in epoch  $t+T$  or late given that the past  $m$  earthquakes occurred in epochs  $t, t-T_1, \dots, t-\sum_{i=1}^{m-1} T_i$**

The conditional probability that the next earthquake occurs in epoch  $t+T$  or later given that the past  $m$  earthquakes occurred in epoch  $t, t-T_1, \dots, t-\sum_{i=1}^{m-1} T_i$  is approximately equal to  $q_{T|1,T_1,\dots,T_{m-1}}$ , defined as

$$(A.28) \quad q_{T|1,T_1,\dots,T_{m-1}} = \sum_{\tau \geq T} r_{\tau|1,T_1,\dots,T_{m-1}}$$

Because there must always be 1 epoch between earthquakes,  $q_{1|1,T_1,\dots,T_{m-1}} = 1$ .

**A.10.5. Conditional probability of being in state  $k$  at epoch  $t$  given that the past  $m$  earthquakes occurred in epochs  $t - T, t - T - T_1, t - T - T_1 - T_2, \dots, t - T - \sum_{i=1}^{m-1} T_i$**

The conditional probability of being in state  $k$  at epoch  $t$  given that the past  $m$  earthquakes occurred at epochs  $t - T, t - T - T_1, t - T - T_1 - T_2, \dots, t - T - \sum_{i=1}^{m-1} T_i$  is approximately equal to  $s_{k|T, T_1, T_2, \dots, T_{m-1}}$  given by

$$(A.29) \quad s_{k|T, T_1, \dots, T_{m-1}} = \frac{v_{k-T+1, T} s_{k-T+1|1, T_1, \dots, T_{m-1}}}{q_{T|T_1, \dots, T_{m-1}}}$$

The probability of interest is  $P(A|B, C, D_j)$  and can be expressed as Equation (A.15). To be in state  $k$  at epoch  $t$  given  $B$  and  $C$ , the system must have been in state  $j = k - T + 1$  in epoch  $t - T + 1$ , which is the epoch immediately following the past earthquake. Thus,  $P(A|B, C, D_j) = 0$  unless  $j = k - T + 1$  in which case  $P(A|B, C, D_j) = 1$ . It follows that Equation (A.16) holds. Observe in Equation (A.16) that  $P(B|C, D_{k-T+1})$  is equal to  $P(B|D_{k-T+1})$  by the Markov property and is thus approximated by  $v_{k-T+1, T}$ . Next,  $P(D_{k-T+1}|C)$  is approximated by  $s_{k-T+1|1, T_1, \dots, T_{m-1}}$ . Finally,  $P(B|C)$  is approximated by  $q_{T|T_1, \dots, T_{m-1}}$ . Equation (A.29) follows from substitution into Equation (A.16).

## A.11. Waiting Time Until Next Earthquake

### A.11.1. Conditional Mean and Variance of Waiting Time Until Next Earthquake Given State $k$

The conditional probability that waiting time until the next earthquake is  $T$  epochs, given that the process is in state  $k$  is equal to  $w_{k, T}$ . The expected waiting time until the next

earthquake given that the process is in state  $k$ , say  $\tau_k$ , is thus

$$(A.30) \quad \tau_k = \sum_{T \geq 1} T w_{k,T}$$

The conditional variance of the waiting time until the next earthquake is approximately

$\sigma_{\tau,k}^2$  defined as

$$(A.31) \quad \begin{aligned} \sigma_{\tau,k}^2 &= \sum_{T \geq 1} (T - \tau_k)^2 w_{k,T} \\ &= \sum_{T \geq 1} [T^2 w_{k,T} - \tau_k^2] \end{aligned}$$

### A.11.2. Unconditional Mean and Variance of Waiting Time Until Next Earthquake

The unconditional expected waiting time until the next earthquake is approximately  $\tau$ , defined as

$$(A.32) \quad \tau = \sum_{T \geq 1} T w_T$$

The unconditional variance of the waiting time until the next earthquake is approximately

$\sigma_{\tau}^2$  defined as

$$(A.33) \quad \begin{aligned} \sigma_{\tau}^2 &= \sum_{T \geq 1} (T - \tau)^2 w_T \\ &= \sum_{T \geq 1} [T^2 w_T - \tau^2] \end{aligned}$$

### A.11.3. Conditional Mean and Variance of Waiting Time Given Earthquake History

We can improve our estimation of the mean and variance of the expected waiting time by conditioning on the earthquake history. The expected waiting time given that the previous  $m$  earthquakes occurred  $T$  epochs ago,  $T + T_1$  epochs ago, . . . ,  $T + \sum_{i=1}^{m-1} T_i$  epochs ago is approximately equal to  $\tau_{|T, T_1, T_2, \dots, T_{m-1}}$  which is a weighted sum of  $\tau_k$  values

$$(A.34) \quad \tau_{|T, T_1, T_2, \dots, T_{m-1}} = \sum_k \tau_k s_{k|T, T_1, T_2, \dots, T_{m-1}}$$

The corresponding variance for the expected waiting time given that earthquake history is approximately equal to  $\sigma_{\tau|T, T_1, T_2, \dots, T_{m-1}}^2$  defined as

$$(A.35) \quad \begin{aligned} \sigma_{\tau|T, T_1, T_2, \dots, T_{m-1}}^2 &= \sum_k \left( \tau_k s_{k|T, T_1, T_2, \dots, T_{m-1}} - \tau_{|T, T_1, T_2, \dots, T_{m-1}} \right)^2 \\ &= \sum_k \left( \tau_k^2 s_{k|T, T_1, T_2, \dots, T_{m-1}} - \tau_{|T, T_1, T_2, \dots, T_{m-1}}^2 \right) \end{aligned}$$

### A.11.4. Conditional Probability of an Earthquake in Next $T'$ Epochs Given State $k$

The conditional probability of at least 1 earthquake in the next  $T'$  epochs given that the current state is  $k$  is approximately equal to  $F_{T', k}$  defined as

$$(A.36) \quad F_{T', k} = \sum_{\tau \leq T'} w_{k, \tau}$$

### A.11.5. Unconditional Probability of an Earthquake in Next $T'$ Epochs

Similarly, the unconditional probability of at least 1 earthquake in the next  $T'$  epochs is approximately equal to  $F_{T'}$  defined as

$$(A.37) \quad F_{T'} = \sum_{\tau \leq T'} w_{\tau}$$

### A.11.6. Conditional Probability of an Earthquake in Next $T'$ Epochs Given Earthquake History

The conditional probability of at least 1 earthquake in the next  $T'$  epochs given that the previous earthquakes occurred  $T$  epochs ago,  $T + T_1$  epochs ago, . . . ,  $T + \sum_{i=1}^{m-1} T_i$  epochs ago is approximately equal to

$$(A.38) \quad F_{T'|T, T_1, T_2, \dots, T_{m-1}} = \sum_k F_{T', k} s_{k|T, T_1, T_2, \dots, T_{m-1}}$$

Equation (A.38) can also be rewritten as

$$(A.39) \quad F_{T'|T, T_1, T_2, \dots, T_{m-1}} = \sum_k (1 - v_{k, T'+1}) s_{k|T, T_1, T_2, \dots, T_{m-1}}$$

## A.12. Waiting Time Between Renewals of System

Although the times between earthquakes are not independent, the system can be viewed as a renewal process from the perspective of a sufficiently long time scale. For example, the LTFM process can be considered to renew itself each epoch  $t$  at which the accumulated strain  $S_t$  equals 0. As described in Section A1, state 1 has accumulated strain  $S$  equal to 0, so we would then consider the time for the system from state 1 back

to state 1. More generally, we could consider the time (number of epochs) needed for the accumulated strain to move from at or below a threshold  $\kappa_1$  to above another threshold  $\kappa_2$  and then back to  $\kappa_1$  or below. For simplicity, we consider here the time needed for the system to move from state 1 back to state 1, or from zero accumulated strain aback to zero accumulated strain.

The number of epochs,  $y$ , required to move from state 1 back to state 1 is equal to 1 plus the number of epochs required to move from state 2 to state 1, because the probability of an earthquake in state 1 equals 0. To find the distribution of  $y$ , we consider a modified process that is identical to the original one except that once the modified process enters state 1, it remains there. Denote the transition matrix of the modified process by  $\bar{\mathbf{P}}$ . The matrix  $\bar{\mathbf{P}}$  is identical to  $\mathbf{P}$  except that the entries in the first row are  $\bar{p}_{1,1} = 1$  and  $\bar{p}_{1,k} = 0$  for  $k > 1$ .

Let  $G_{1,2}(t)$  denote the conditional probability that the modified process is in state 1  $t$  epochs after being in state 2. The theory of Markov chains implies

$$(A.40) \quad G_{1,2}(t) = (0, 1, 0, \dots, 0) \bar{\mathbf{P}}^t (1, 0, 0, \dots, 0)^T$$

Notice that  $G_{1,2}(T)$  gives the conditional probability that the modified process has entered state 1 within  $t$  epochs after being in state 2. Let  $g_{1,2}(t) = G_{1,2}(t) - G_{1,2}(t - 1)$  denote the conditional probability that the modified process enters state 1 exactly  $t$  epochs after being in state 2. Observe that

$$(A.41) \quad g_{1,2}(t) = (0, 1, 0, \dots, 0) \left( \bar{\mathbf{P}}^t - \bar{\mathbf{P}}^{t-1} \right) (1, 0, 0, \dots, 0)^T$$

The expected number of epochs for the original process to return to state 1 from state 1 thus has expected value  $\mu_{\text{renew}}$  given by

$$(A.42) \quad \mu_{\text{renew}} = 1 + \sum_{t \geq 0} t g_{1,2}(t)$$

The standard deviation  $\sigma_{\text{renew}}$  of the time between renewals given by

$$(A.43) \quad \sigma_{\text{renew}} = \sqrt{\sum_{t \geq 0} (1 + t - \mu_{\text{renew}})^2 g_{1,2}(t)}$$

and coefficient of variation (cv) for the time between renewals,

$$(A.44) \quad cv_{\text{renew}} = \sigma_{\text{renew}} / \mu_{\text{renew}}$$

## References

- [1] Abercrombie, R. E., 1995. Earthquake source scaling relationships from -1 to 5  $M_L$  using seismograms recorded at 2.5-km depth, *J. Geophys. Res.* 100 24015-24036. doi:10.1029/95JB02397
- [2] Abercrombie, R. E., 2014. Stress drops of repeating earthquakes along the San Andreas Fault at Parkfield, *Geophys. Res. Lett.* 41 8784-8791. doi:10.1002/2014GL062079
- [3] Adams, J., 2011. Seismic hazard maps for the national building code of Canada. Canadian Soc. Civ. Eng. General Conference, Jag Humar Symposium Paper 1 (10 pp), 2011, Ottawa. Available at [http://www.earthquakescanada.nrcan.gc.ca/hazard-alea/2011CSCE/2011CSCE\\_JHS-1.pdf](http://www.earthquakescanada.nrcan.gc.ca/hazard-alea/2011CSCE/2011CSCE_JHS-1.pdf)
- [4] Aki, K., 1965. Maximum likelihood estimate of  $b$  in the formula  $\log(N) = a - bM$  and its confidence limits. *Bull. Earthq. Res. Inst. Univ. Tokyo* 43, 237-239.
- [5] Allmann, B. P., Shearer, P. M., 2007. Spatial and temporal stress drop variations in small earthquakes near Parkfield, California, *J. Geophys. Res.: Solid Earth* 112 B04305. doi:10.1029/2006JB004395
- [6] Allmann, B. P., Shearer, P. M., 2009. Global variations of stress drop for moderate to large earthquakes, *J. Geophys. Res.: Solid Earth* 114 B01310. doi:10.1029/2008JB005821
- [7] Ambraseys, N., 1991. The Rukwa earthquake of 13 December 1910 in East Africa, *Terra Nova* 3(2), 202-211. doi:10.1111/j.1365-3121.1991.tb00873.x
- [8] Anagnos, T., Kiremidjian, A. S., 1985. A stochastic earthquake recurrence model with temporal and spatial dependence, Report No. 76, John A. Blume Earthquake Engineering Center, Stanford University, Stanford, CA

- [9] Anagnos, T., Kiremidjian, A. S., 1988. A review of earthquake occurrence models for seismic hazard analysis, *Probabilistic Engineering Mechanics* 3, 3–11. doi:10.1016/0266-8920(88)90002-1
- [10] Anderson, J.G., Biasi, G.P., Wesnousky, S.G., 2017. Fault-scaling relationships depend on the average fault-slip rate, *Bull. Seismol. Soc. Am.* 107(6), 2561–2577. doi:10.1785/0120160361
- [11] Ando, M., 1975. Source mechanisms and tectonic significance of historical earthquakes along the Nankai Trough, Japan, *Tectonophysics* 27, 119-140.
- [12] Atkinson, G. M., Beresnev, I., 1997. Don't call it stress drop, *Seismol. Res. Lett.* 68 3-4. doi:10.1785/gssrl.68.1.3
- [13] Axen, G.J., 1999. Low-angle normal fault earthquakes and triggering. *Geophys. Res. Lett.* 26(24), 3683-3696. doi: 10.1029/1999GL005405
- [14] Ayele, A., Kulhánek, O., 2000. Reassessment of source parameters for the three major earthquakes in the East African rift system from historical seismograms and bulletins. *Annali Di Geofisica* 43(1). 81-94. doi:10.4401/ag-3627
- [15] Baltay, A. S., Hanks, T. C., 2014. Understanding the magnitude dependence of PGA and PGV in NGA-West 2 Data, *Bull. Seismol. Soc. Am.* 104 2851-2865. doi:10.1785/0120130283
- [16] Baltay, A. S., Hanks, T. C., Abrahamson, N. A., 2017. Uncertainty, variability, and earthquake physics in ground-motion prediction equations, *Bull. Seismol. Soc. Am.* 107 1754-1772. doi:10.1785/0120160164
- [17] Baltay, A. S., Hanks, T. C., Beroza, G. C., 2013. Stable stress-drop measurements and their variability: Implications for ground-motion prediction, *Bull. Seismol. Soc. Am.* 103 211-222. doi:10.1785/0120120161
- [18] Baltay, A. S., Ide, S., Prieto, G., Beroza, G., 2011. Variability in earthquake stress drop and apparent stress, *Geophys. Res. Lett.* 38 L06303. doi:10.1029/2011GL046698
- [19] Basham, P.W, Adams, J., 1983. Earthquakes on the continental margin of eastern Canada: need future large events be confined to the locations of large historical events? in "The 1886 Charleston earthquake and its implications for today" U.S. Geol. Survey Open File Report 83-843, 456–467.

- [20] Beeler, N. M., Tullis, T., Junger, J., Kilgore, B., Goldsby D., 2014. Laboratory constraints on models of earthquake recurrence, *J. Geophys. Res.: Solid Earth* 119, 8770–8791. doi:10.1002/ 2014JB011184.
- [21] Bell, A. F., Naylor, M., Main, I. G., 2013. Convergence of the frequency-size distribution of global earthquakes. *Geophys. Res. Lett.* 40(11), 2585–2589. doi:10.1002/grl.50416
- [22] Benedetti, R., 2010. Scoring rules for forecast verification, *Monthly Weather Review* 138, 203–211. doi:10.1175/2009MWR2945.1
- [23] Biasi, G.P., Weldon, R. J., Dawson, T., 2013. Distribution of slip in ruptures, U.S. Geol. Surv. Open-File Rept. 2013-1165, Uniform California Earthquake Rupture Forecast Version 3 (UCERF3)–The Time-Independent Model, Appendix F, 41 pp.
- [24] 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, 2761–2781. doi:10.1785/0120000605
- [25] Biasi, G.P., Wesnousky, S.G., 2016. Steps and gaps in ground ruptures: Empirical bounds on rupture propagation, *Bull. Seismol. Soc. Am.* 106(3), 1110–1124. doi:10.1785/0120150175
- [26] Biasi, G.P., Wesnousky, S.G., 2017. Bends and ends of surface ruptures. *Bull. Seismol. Soc. Am.* 107(6), 2543–2560. doi: 10.1785/0120160292
- [27] Bignami, C., Valerio, E., Caminati, E., Doglioni, C., Petricca, P., Tizzani, P., Lanari, R., 2020. Are normal fault earthquakes due to elastic rebound or gravitational collapse?, *Annals of Geophysics* 63(2), SE123. doi:10.4401/ag-8455
- [28] Bilek, S.L., Lay, T., 2018. Subduction zone megathrust earthquakes, *Geosphere* 14(4), 1468–1500. doi:10.1130/GES01608.1.
- [29] Boore, D. M., 1983. Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, *Bull. Seismol. Soc. Am.* 73 1865–1894.
- [30] Borrero, J. C., Weiss, R., Okal, E. A., Hidayat, R., Arcas, D., Titov, V. V., 2009. The tsunami of 2007 September 12, Bengkulu province, Sumatra, Indonesia: post-tsunami field survey and numerical modelling, *Geophys. J. Int.* 178, 180–194. doi:10.1111/j.1365-246X.2008.04058.x

- [31] Brier, G.W., 1950. Verification of forecasts expressed in terms of probability, *Monthly Weather Review* 78, 1-3. doi:10.1175/1520-0493(1950)0782.0.CO;2
- [32] Brillinger, D. R., 1982. Seismic risk assessment: Some statistical aspects, *Earthquake Prediction Research* 1, 183–195.
- [33] Brune, J. N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.* 75 4997-5009. doi:10.1029/JB075i026p04997
- [34] Burov, E.B., 2011. Rheology and strength of the lithosphere, *Marine and Petroleum Geology* 28, 1402-1443. doi:10.1016/j.marpetgeo.2011.05.008
- [35] 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.*, 43(20), 10621-10637. doi:10.1002/2016GL070815
- [36] 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 assessments of long-term earthquake activity and maximum magnitude in intraplate northwest Europe, between the Lower Rhine Embayment and the North Sea, in *Continental Intraplate Earthquakes: Science, Hazard, and Policy Issues*, edited by S. Stein and S. Mazzotti, 193-224, Geol. Soc. Amer. Inc, Boulder.
- [37] Campo, E.A., 2008. *Selection of Polymeric Materials: How to Select Design Properties from Different Standards*, 1st ed., Published by William Andrew
- [38] Chapman, M. C., Beale, J. N., Hardy, A. C., Wu, Q., 2016. Modern seismicity and the fault responsible for the 1886 Charleston, South Carolina, earthquake. *Bull. Seismol. Soc. Am.*, 106(2), 364-372. doi:10.1785/0120150221
- [39] Chou, I. H., Fischer, J. A., 1975. Earthquake hazard and confidence, *Proceedings of U.S. National Conference on Earthquake Engineering 1975*, 34–42.
- [40] Çinlar, E., 1975. *Introduction to Stochastic Processes*, Prentice Hall.
- [41] Clark, K.J., Cochran, U.A., Berryman, K.R., Biasi, G., Langridge, R., Villamor, P., Bartholomew, T., Litchfield, N., Pantosti, D., Marco, S., Van Dissen, R., 2013. Deriving a long paleoseismic record from a shallow-water Holocene basin next to the Alpine fault, New Zealand, *GSA Bulletin* 125, 811–832. doi:10.1130/B30693.1

- [42] Clark, D., McPherson, A., Collins, C. D. N., 2011. Australia's seismogenic neotectonic record: a case for heterogeneous intraplate deformation. *Geoscience Australia Record* 2011/11.
- [43] Clark, D., McPherson, A., Cupper, M., Collins, C. D. N., Nelson, G. 2017. The Cadell Fault, southeastern Australia: a record for temporally clustered morphogenic seismicity in a low-strain intraplate region, in *Seismicity, Fault Rupture and Earthquake Hazards in Slowly Deforming Regions*, edited by A. Landgraf, S. Kübler, E. Hintersberger, S. Stein, 163-185, Geol. Soc., London, Special Publications. doi:10.1144/SP432.2
- [44] Cloetingh, S., Beekman, F., Ziegler, P. A., Van Wees, J. D., Sokoutis, D., 2008. Post-rift compressional reactivation potential of passive margins and extensional basins, in *Nature and Origin of Compression in Passive Margins*, edited by H. Johnson, T. G. Dore, R. W. Gatliff, R. W. Holdsworth, E. R. Lundin and J. D. Ritchie, 27-70, Geol. Soc. Publishing House, Bath. doi:10.1144/SP306.2
- [45] Cocco, M., Rovelli, A., 1989. Evidence for the variation of stress drop between normal and thrust faulting earthquakes in Italy, *J. Geophys. Res.* 95 9399-9416. doi:10.1029/JB094iB07p09399
- [46] Cocco, M., Tinti, E., Cirella, A., 2016. On the scale dependence of earthquake stress drop, *J. Seismol.* 20 1151-1170. doi:10.1007/s10950-016-9594-4
- [47] Cochran, U.A., Clark, K.J., Howarth, J.D., Biasi, G.P., Langridge, R.M., Villamor, P., Berryman, K.R., Vandergoes, M.J., 2017. A plate boundary earthquake record from a wetland adjacent to the Alpine fault in New Zealand refines hazard estimates. *Earth and Planet. Sci. Lett.* 464, 175–188. doi:10.1016/j.epsl.2017.02.026.
- [48] Coffin, M.F., Gahagan, L.M., Lawver, L.A., 1998. Present-day plate boundary digital data compilation, University of Texas Institute for Geophysics Technical Report No. 174.
- [49] Collettini, C., 2011. The mechanical paradox of low-angle normal faults: Current understanding and open questions, *Tectonophysics* 510(3-4), 253-268. doi:10.1016/j.tecto.2011.07.015
- [50] Collettini, C., Sibson, R.H., 2001. Normal faults, normal friction? *Geology* 29(10), 927-930. doi:10.1130/0091-7613(2001)029;0927:NFNF;2.0.CO;2
- [51] Cornell, C.A., 1968. Engineering seismic risk analysis. *Bull. Seism. Soc. Am.*, 58, 1583-1606.

- [52] Cornell, C. A., Winterstein, S. R., 1988. Temporal and magnitude dependence in earthquake recurrence models, *Bull. Seismol. Soc. Am.* 78, 1522–1537. doi:10.1785/BSSA0780041522
- [53] Cotton, F., Archuleta, R., Causse, M., 2013. What is sigma of the stress drop? *Seismol. Res. Lett.* 84 42-48. doi:10.1785/0220120087
- [54] Courboux, F., Vallée, M., Causse, M., Chounet, A., 2016. Stress-drop variability of shallow earthquakes extracted from a global database of source time functions, *Seismol. Res. Lett.* 87 912-918. doi:10.1785/0220150283
- [55] Craig, T.J., Copley, A., Jackson, J., 2014. A reassessment of outer-rise seismicity and its implications for the mechanics of oceanic lithosphere, *Geophys. J. Int.* 197(1), 63-89. doi:10.1093/gji/ggu013
- [56] Crone, A. J., De Martini, P. M., Machette, M. N., Okumura, K., Prescott, J. R., 2003. Paleoseismicity of two historically quiescent faults in Australia: Implications for fault behavior in stable continental regions. *Bull. Seismol. Soc. Am.*, 93(5), 1913-1934.
- [57] Davis, P. M., Jackson, D. D., Kagan, Y. Y., 1989. The longer it has been since the last earthquake, the longer the expected time till the next? *Bull. Seismol. Soc. Am.* 79, 1439–1456. doi:10.1785/BSSA0790051439
- [58] Dempsey, D., Ellis, S., Archer, R., Rowland, J., 2012. Energetics of normal earthquakes on dip-slip faults, *Geology* 40, 279-282. doi:10.1130/G32643.1.
- [59] Deng, Q.D., Sung, F.M., Zhu, S.L., Li, M.L., Wang, T.L., Zhang, W.Q., Burchfiel, B.C., Molnar, P., Zhang, P.Z., 1984. Active faulting and tectonics of the Ningxia-Hui autonomous region, China. *J. of Geophys. Res.* 89, 4427–4445. doi:10.1029/JB089iB06p04427
- [60] Déverchère, J., Petit, C., Gileva, N., Radziminovitch, N., Melnikova, V., San'kov, V., 2015. Depth distribution of earthquakes in the Baikal rift system and its implications for the rheology of the lithosphere, *Geophys. J. Int.* 146(3), 714–730. doi:10.1046/j.0956-540x.2001.1484.484.x
- [61] Di Giacomo, D., Bondar, I., Storchak, D.A., Engdahl, E.R., Bormann, P., Harris, J., 2015. ISC-GEM: Global Instrumental Earthquake Catalog (1900-2009), III. Re-computed  $M_s$  and  $m_b$ , proxy  $M_w$ , final magnitude composition and completeness assessment, *Phys. of the Earth and Planet. Interiors* 239, 33-47. doi:10.1016/j.pepi.2014.06.005

- [62] Doglioni, C., Carminati, E., Petricca, P., Riguzzi, F., 2015. Normal fault earthquakes or graviquakes, *Scientific Reports* 5, 12110. doi:10.1038/srep12110
- [63] Doser, D.I., 1985. Source parameters and faulting processes of the 1959 Hebgen Lake, Montana, Earthquake Sequence. *J. Geophys. Res.* 90(B6), 4537-4555. doi:10.1029/JB090iB06p04537
- [64] Doser, D.I., 1986. Earthquake processes in the Rainbow Mountain-Fairview Peak-Dixie Valley, Nevada, region 1954-1959, *J. Geophys Res.* 91, 2572-2586. doi:10.1029/JB091iB12p12572
- [65] Doser, D.I., 1988. Source parameters of earthquakes in the Nevada Seismic Zone, 1915-1943, *J. Geophys. Res.* 93(B12), 15001-15015. doi:10.1029/JB093iB12p15001
- [66] Doser, D.I., 1991. Faulting within the eastern Baikal rift as characterized by earthquake studies, *Tectonophysics* 196(1-2), 109-139. doi:10.1016/0040-1951(91)90292-Z
- [67] Doser, D.I., Yarwood, D.R.. 1990. Strike-slip faulting in continental rifts: example from Sabukia, East Africa (1928), and other regions, *Tectonophysics* 197(2-4), 213-224. doi:10.1016/0040-1951(91)90042-Q
- [68] DuRoss, C.B., Personius, S.F., Crone, A.J., Olig S.S, Hylland, M.D., Lund, W.R., Schwartz, D.P., 2016. Fault segmentation: New concepts from the Wasatch Fault Zone, Utah, USA, *J. Geophys. Res.: Solid Earth* 121, 1131-1157. doi:10.1002/2015JB012519
- [69] Dziewonski, A.M., Chou, T.-A., Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.: Solid Earth* 86, 2825-2852. doi:10.1029/JB086iB04p02825
- [70] Ebel, J. E., Bonjer, K.-P., Oncescu, M. C., 2000. Paleoseismicity: Seismicity evidence for past large earthquakes. *Seismol. Res. Lett.*, 71(2), 283-294. doi:10.1785/gssrl.71.2.283
- [71] Ebel, J. E., Chambers, D. W., Kafka, A. L., Baglivo, J. A., 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
- [72] Ekström, G., Nettles, M., Dziewonski, A. M., 2012. The global CMT project 2004-2010: centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. In.* 200-201 1-9. doi:10.1016/j.pepi.2012.04.002

- [73] Eshelby, J. D., 1957. The determination of the elastic field of an ellipsoid, and related problems, *Proc. R. Soc. Lond. A* 241 376–396. doi:10.1098/rspa.1957.0133
- [74] Feller, W., 1968. *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed., Wiley
- [75] Felzer, K. R., 2008. Calculating California seismicity rates, Appendix I in *The Uniform California Earthquake Rupture Forecast, version 2 (UCERF 2)*: U.S. Geol. Survey Open File Report 2007-1437I and California Geological Survey Special Report 203I, 42.
- [76] Field, E. H., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. A., Johnson, K. M., Jordan, T. H., Madden, C., Michael, A. J., Milner, K., Page, M. T., Parsons, T. E., Powers, P., Shaw, B. E., Thatcher, W. R., Weldon, R. J. Zeng, Y., 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
- [77] Field, E. H., Jordan, T. H., Page, M. T., Milner, K. R., Shaw, B. E., Dawson, T. E., Biasi, G. P., Parsons, T., Hardebeck, J. L., Michael, A. J., Weldon, R. J., Powers, P. M., Johnson, K. M., Zeng, Y., Felzer, K. R., van der Elst, N., Madden, C., Arrowsmith, R., Werner, M. J., Thatcher, W. R., 2017. A synoptic view of the Third Uniform California Earthquake Rupture Forecast (UCERF3). *Seismol. Res. Lett.*, 88(5), 1259-1267. doi:10.1785/0220170045
- [78] 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 implication for the spectral character of Earth deformation at periods of 10 to 10 million years, *J. Geophys. Res.: Solid Earth* 108. doi:10.1029/2001JB000682
- [79] Frohlich, C., 1992. Triangle diagrams: ternary graphs to display similarity and diversity of earthquake focal mechanisms, *Phys. Earth Planet. Int.* 75 193-198. doi:10.1016/0031-9201(92)90130-N
- [80] Geller, R. J., 2011. Shake-up time for Japanese seismology. *Nature*, 472, 407-409. doi:10.1038/nature10105
- [81] Ghosh, A., Holt, W. E., Wen, L., 2013. Predicting the lithospheric stress field and plate motions by joint modeling of lithosphere and mantle dynamics. *J. Geophys. Res.: Solid Earth*, 118(1), 346-368. doi:10.1029/2012JB009516
- [82] Girardin, V., Limnios, N., 2018. *Applied Probability: From Random Sequences to Stochastic Processes*, Springer

- [83] Godano, M., Bernard, P., Dublanchet, P., 2015. Bayesian inversion of seismic spectral ratio for source scaling: application to a persistent multiplet in the western Corinth rift, *J. Geophys. Res.: Solid Earth* 120 7683-7712. doi:10.1002/2015JB012217
- [84] 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 Geol.*, 384, 4-46, doi:10.1016/j.margeo.2016.06.008
- [85] Goldfinger, C., Ikeda, Y., Yeats, R. S., Ren, J., 2013. Superquakes and supercycles, *Seismol. Res. Lett.* 84, 1–9. doi:10.1785/0220110135
- [86] Goldfinger, C., Nelson, C. H., Morey, A. E., Johnson, J. E., Patton, J. R., Karabanov, E. B., Gutierrez-Pastor, J., Eriksson, A. T., Gracia, E., Dunhill, G., Enkin, R. J., Dallimore, A., Vallier, T., 2012. Turbidite event history—Methods and implications for Holocene paleoseismicity of the Cascadia subduction zone, U.S. Geological Survey Professional Paper 1661-F, 170 p. doi:10.3133/pp1661F
- [87] Gordon, R. G., Stein, S., 1992. Global tectonics and space geodesy, *Science* 256, 333–342. doi:10.1126/science.256.5055.333
- [88] Gulkan, P., 2013. A dispassionate view of seismic-hazard assessment. *Seismol. Res. Lett.*, 84(3), 413-416. doi:10.1785/0220130005
- [89] Gupta, H., Rao, N., Rastogi, B., Sarkar, D., 2001. The deadliest intraplate earthquake, *Science* 291(5511), 2101-2102. doi:10.1126/science.1060197
- [90] 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
- [91] Hanks, T. C., McGuire, R. K., 1981. The character of high-frequency strong ground motion, *Bull. Seismol. Soc. Am.* 71 2071-2095.
- [92] Hashimoto, M., 2021. Is the Long-Term Probability of the Occurrence of Large Earthquakes along the Nankai Trough Inflated? Conflict between Science and Risk Management, *Seismol. Res. Lett.* 93, 1971–1973. doi:10.1785/0220220068

- [93] Hecker, S., DeLong, S. B., Schwartz, D. P., 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
- [94] Hough, S. E., 2014. Shaking from injection-induced earthquakes in the central and eastern United States, *Bull. Seismol. Soc. Am.* 104 2619-2626. doi:10.1785/0120140099
- [95] Hough, S. E., 2016. Predicting the unpredictable: The tumultuous science of earthquake prediction, Princeton University Press
- [96] Huang, Y., Ellsworth, W. L., Beroza, G. C., 2017. Stress drops of induced and tectonic earthquakes in the central United States are indistinguishable, *Sci. Adv.* 3 e1700772. doi:10.1126/sciadv.1700772
- [97] Imanishi, K., Ellsworth, W. L., 2006. Source scaling relationships of microearthquakes at Parkfield, CA, determined using the SAFOD pilot hole seismic array, in *Earthquakes: Radiated Energy and the Physics of Faulting*. Vol. 170, pp. 81-90, eds. R. Abercrombie, A. McGarr, G. D. Toro, H. Kanamori, American Geophysical Union. doi:10.1029/170GM10
- [98] Jackson, J., Blenkinsop, T., 1993. The Malawi earthquake of March 10, 1989: deep faulting within the East African rift system, *Tectonics* 12(5), 1131-1139. doi:10.1029/93TC01064
- [99] Jackson, J., White, N.J., 1989. Normal faulting in the upper continental crust: observations from regions of active extension, *J. of Structural Geology* 11(1-2), 15-36. doi:10.1016/0191-8141(89)90033-3
- [100] Johnson, K.L., Nissen, E., Lajoie, L., 2018. Surface rupture morphology and vertical slip distribution of the 1959  $M_w$  7.2 Hebgen Lake (Montana) earthquake from airborne Lidar topography, *J. Geophys. Res.: Solid Earth* 123(9), 8229-8248. doi:10.1029/2017JB015039
- [101] Johnston, A. C., 1989. The seismicity of ‘stable continental interiors’. In: *Earthquakes at North Atlantic Passive Margins: Neotectonics and Postglacial Rebound*. NATO ASI Series (Series C: Mathematical and Physical Sciences), vol 266. Springer, Dordrecht. doi:10.1007/978-94-009-2311-18
- [102] Kane, D. L., Prieto, G. A., Vernon, F. L., Shearer, P. M., 2011. Quantifying seismic source parameter uncertainties, *Bull. Seismol. Soc. Am.* 101 535-543. doi:10.1785/0120100166

- [103] Kaneko, Y., Shearer, P. M., 2015. Variability of seismic source spectra, estimated stress drop, and radiated energy, derived from cohesive-zone models of symmetrical and asymmetrical circular and elliptical ruptures, *J. Geophys. Res.: Solid Earth* 120 1053-1079. doi:10.1002/2014JB011642
- [104] Kagan, Y. Y., 2002. Seismic moment distribution revisited: I. Statistical results. *Geophys. J. Int.*, 148, 520-541. doi:10.1046/j.1365-246x.2002.01594.x
- [105] Kagan, Y.Y., 2003. Accuracy of modern global earthquake catalogs, *Phys. of the Earth and Planet. Interiors* 135(2-3), 173-209. doi:10.1016/S0031-9201(02)00214-5
- [106] Kagan, Y. Y., Jackson, D. D., 2013. Tohoku earthquake: A surprise? *Bull. Seismol. Soc. Amer.*, 103(2B), 1181-1194. doi:10.1785/0120120110
- [107] Kanamori, H., 2011. Prepare for the unexpected. *Nature*, 473, 147. doi:10.1038/473146a
- [108] Kanamori, H., Anderson D. L., 1975. Theoretical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am.* 65 1073-1095.
- [109] Kijko, A., 2004. Estimation of the maximum earthquake magnitude,  $M_{max}$ . *Pure and Appl. Geophys.*, 161(8), 1655-1681. doi:10.1007/s00024-004-2531-4
- [110] Kohlstedt, D.L., Evans, B., Mackwell, S.J., 1995. Strength of the lithosphere: Constraints imposed by laboratory measurements, *J. Geophys. Res.* 100, 17587-17602.
- [111] Kulikova, G., Krüger, F., 2015. Source process of the 1911 M8.0 Chon-Kemin earthquake: Investigation results by analogue seismic records, *Geophys. J. Int.* 201, 1891-1911. doi:10.1093/gji/ggv091
- [112] Leguillon, D., Martin, É., Lafarié-Frenot, M.C., 2015. Flexural vs. tensile strength in brittle materials, *Comptes Rendus Mécanique* 343(4), 275-281. doi:10.1016/j.crme.2015.02.003.
- [113] Leonard, M., Burbidge, D., Allen, T. I., Clark, D., Robinson, D., McPherson, A., Collins, C., 2014. The challenges of probabilistic seismic hazard assessment in stable continental interiors: an Australian example. *Bull. Seism. Soc. Amer.*, 104(6), 3008-3028. doi:10.1785/0120130248
- [114] Lienkaemper, J. J., Williams, P. L., 2007. A Record of Large Earthquakes on the Southern Hayward Fault for the Past 1800 Years, *Bull. Seismol. Soc. Am.* 97, 1803-1819. doi:10.1785/0120060258

- [115] Liu, M., Stein, S., 2016. Mid-continental earthquakes: Spatiotemporal occurrences, causes, and hazards. *Earth-Science. Rev.*, 162, 364-386. doi:10.1016/j.earscirev.2016.09.016
- [116] 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. doi:10.1130/L129.1
- [117] Lomnitz-Adler, J., 1983. A statistical model of the earthquake process, *Bull. Seismol. Soc. Am.* 73, 853–862. doi:10.1785/BSSA0730030853
- [118] Madariaga, R., 1976. Dynamics of an expanding circular fault, *Bull. Seismol. Soc. Am.* 66 639-666.
- [119] Malagnini, L., Mayeda, K., Nielsen, S., Yoo, S.-H., Munafo, I., Rawles, C., Boschi, E., 2014. Scaling transition in earthquake sources: a possible link between seismic and laboratory measurements, *Pure Appl. Geophys.* 171 2685-2707. doi:10.1007/s00024-013-0749-8
- [120] Manaker, D. M., Calais, E., Freed, A. M., Ali, S. T., Przybylski, P., Mattioli, G., Jansma, P., Pr epetit, C., De Chabalier, J. B., 2008. Interseismic plate coupling and strain partitioning in the Northeastern Caribbean. *Geophys. J. Int.*, 174(3), 889-903. doi:10.1111/j.1365-246X.2008.03819.x
- [121] Manighetti, I., Caulet, C., Barros, L.D., Perrin, C., Cappa, F., Gaudemer, Y., 2015. Generic along-strike segmentation of Afar normal faults, East Africa: Implications on fault growth and stress heterogeneity on seismogenic fault planes, *Geochem., Geophys., Geosys.* 16, 443–467. doi:10.1002/2014GC005691.
- [122] Matthews, K.J., Maloney, K.T., Zahirovic, S., Williams, S.E., Seton, M., M uller, R D., 2016. Global plate boundary evolution and kinematics since the late Paleozoic, *Global and Planet. Change* (146), 226-250. doi:10.1016/j.gloplacha.2016.10.002.
- [123] Matthews, M. V., Ellsworth, W. L., Reasenber, P. A., 2002. A Brownian model for recurrent earthquakes, *Bull. Seismol. Soc. Am.* 92, 2233–2250. doi:10.1785/0120010267
- [124] Mayeda, K., Walter, W., 1996. Moment, energy, stress drop, and source spectra of western United States earthquakes from regional coda envelopes, *J. Geophys. Res.* 101 11195-11208. doi:10.1029/96JB00112
- [125] Mazzotti, S., James, T. S., Henton, J., Adams, J., 2005. GPS crustal strain, postglacial rebound, and seismic hazard in eastern North America: The

- Saint Lawrence valley example. *J. Geophys. Res.: Solid Earth*, 110(B11), 16. doi:10.1029/2004JB003590
- [126] McCaffrey, R. 2008. Global frequency of magnitude 9 earthquakes. *Geology*, 36(3), 263-266. doi:10.1130/G24402A.1
- [127] McGuire, R. K., 1977. Effects of uncertainty in seismicity on estimates of seismic hazard for the east coast of the United States. *Bull. Seismol. Soc. Amer.*, 67(3), 827-848.
- [128] Nazareth, J.J., Hauksson, E., 2004. The seismogenic thickness of the Southern California crust, *Bull. Seismol. Soc. Am.* 94(3), 940–960. doi:10.1785/0120020129
- [129] Neely, J.S., Salditch, L., Spencer, B.D., Stein S., A more realistic model for the probability of large earthquakes. *Bull. Seismol. Soc. Am.* In Review.
- [130] Okal, E.A., 1976. Surface-wave investigation of rupture mechanism of Gobi-Altai (December 4, 1957) earthquake, *Phys. of the Earth and Planet. Interiors* 12, 319–328. doi:10.1016/0031-9201(76)90027-3
- [131] Okal, E.A., Romanowicz, B.A., 1994. On the variation of b-values with earthquake size, *Phys. of the Earth and Planet. Interiors* 87(1-2), 55-76. doi:10.1016/0031-9201(94)90021-3
- [132] Okal, E.A., Synolakis, C.E., Uslu, B., Kalligeris, N., Voukouvalas, E., 2009. The 1956 earthquake and tsunami in Amorgos, Greece, *Geophys. J. Int.* 178(3), 1533-1554. doi:10.1111/j.1365-246X.2009.04237.x
- [133] Okuwaki, R., Yagi, Y., 2017. Rupture process during the  $M_w$  8.1 2017 Chiapas, Mexico earthquake: shallow intraplate normal faulting by slab bending, *Geophys. Res. Lett.* 44, 11,816–11,823. doi:10.1002/2017GL075956
- [134] Papazachos B.C. Delibassis N.D., 1969. Tectonic stress field and seismic faulting in the area of Greece, *Tectonophysics*, 7, 231–255.
- [135] Pazzaglia, F. J., Zeitler, P. K., Idleman, B. D., McKeon, R., Berti, C., Enkelmann, E., Laucks, J., Ault, A., Elasmir, M., Becker, T., 2010. Tectonics and topography of the Cenozoic Appalachians, in *Tectonics of the Susquehanna Piedmont*, Symposium associated with Field Conference of Pennsylvania Geologists, 75, 111-126.
- [136] Peresan, A., Panza, G. F., 2012. Improving earthquake hazard assessments in Italy: an alternative to “Texas sharpshooting”. *Eos, Trans. Amer. Geophys. Union*, 93(51), 538-538. doi:10.1029/2012EO510009

- [137] Petrucci A., Schorlemmer. D., Tormann, T., Rinaldi, A.P., Wiemer, S., Gasperini, P., and Vanucci, G., 2019. The influence of faulting style on the size-distribution of global earthquakes, *Earth and Planet. Science Lett.* 527. doi:10.1016/j.epsl.2019.115791
- [138] Rabiner, L. R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of the IEEE* 77, 257–286. doi:10.1109/5.18626
- [139] 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, DC.
- [140] Ritsema, A.R., 1974. Earthquake mechanisms of the Balkan region, *Kon. Ned. Meteor. Inst. Repts.* 74(4), 36 pp.
- [141] 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, 1200–1236. doi:10.1785/0119990023
- [142] Rubinstein, J. L., Ellsworth, W. L., Beeler, N. M., Kilgore, B. D., Lockner, D. A., Savage, H. M., 2012. Fixed recurrence and slip models better predict earthquake behavior than the time- and slip-predictable models: 2. Laboratory earthquakes, *J. Geophys. Res.: Solid Earth* 117. doi:10.1029/2011JB008723
- [143] Rundle, J.B., 1989. Derivation of the complete Gutenberg-Richter magnitude-frequency relation using the principle of scale invariance, *J. of Geophys. Res.: Solid Earth* 94(B9), 12337-12342. doi:10.1029/JB094iB09p12337
- [144] Sagiya, T., 2011. Integrate all available data, *Nature*, 473, 146-147. doi:10.1038/473146a
- [145] Salditch, L., Stein, S., Neely, J. S., Spencer, B. D., Brooks, E. Agnon, M. A., Liu M., 2020. Earthquake supercycles and Long-Term Fault Memory, *Tectonophysics* 774. doi:10.1016/j.tecto.2019.228289
- [146] Sato, T., Hirasawa, T., 1973. Body wave spectra from propagating shear cracks, *J. Phys. Earth* 21 415-431. doi:10.4294/jpe1952.21.415
- [147] Scharer, K. M., Biasi, G. P., Weldon, R. J., 2011. A reevaluation of the Pallett Creek earthquake chronology based on AMS radiocarbon dates, San Andreas fault, California, *J. Geophys. Res.: Solid Earth* 116, B12111. doi:10.1029/2010JB008099

- [148] Scharer, K. M., Biasi, G. P., Weldon, R. J., Fumal, T. E., 2010. Quasi-periodic recurrence of large earthquakes on the southern San Andreas fault, *Geology* 38, 555–558. doi:10.1130/G30746.1
- [149] Scholz, C.H., 1968. The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes, *Bull. Seismol. Soc. Am.* 58(1), 399–415.
- [150] Scholz, C.H., 2015. On the stress dependence of the earthquake  $b$  value. *Geophys. Res. Lett.* 42, 1399-1402. doi:10.1002/2014GL062863
- [151] Schorlemmer, D., Werner, M. J., Marzocchi, W., Jordan, T. H., Ogata, Y., Jackson, D. D., Mak, S., Rhoades, D. A., Gerstenberger, M. C., Hirata, N., Liukis, M., Maechling, P. J., Strader, A., Taroni, M., Wiemer, S., Zechar, J. D., Zhuang, J., 2018. The Collaboratory for the Study of Earthquake Predictability: Achievements and Priorities, *Seismol. Res. Lett.* 89, 1305–1313. doi:10.1785/0220180053
- [152] Schorlemmer, D., Wiemer, S., Wyss, M., 2005. Variations in earthquake-size distribution across different stress regimes, *Nature* 437(7058), 539-542. doi:10.1038/nature04094
- [153] Schulte, S. M., Mooney, W. D., 2005. An updated global earthquake catalogue for stable continental regions: reassessing the correlation with ancient rifts. *Geophys. J. Int.*, 161(3), 707-721. doi:10.1111/j.1365-246X.2005.02554.x
- [154] Sella, G. F., Stein, S., Dixon, T. H., Craymer, M., James, T. S., Mazzotti, S., Dokka, R. K., 2007. Observation of glacial isostatic adjustment in “stable” North America with GPS. *Geophys. Res. Lett.*, 34(2), 6. doi:10.1029/2006GL027081
- [155] Shi, Y., Bolt, B.A., 1982. The standard error of the magnitude-frequency  $b$  value, *Bull. Seismol. Soc. Amer.* 72, 1677-1687.
- [156] Shimazaki, K., Nakata T., 1980. Time-predictable recurrence model for large earthquakes, *Geophys. Res. Lett.* 7, 279–282. doi:10.1029/GL007i004p00279
- [157] Sibson, R.H., 1977. Fault rocks and fault mechanisms, *J. of the Geological Society* 133(3), 191–213. doi:10.1144/gsjgs.133.3.0191
- [158] Sibson, R.H., 1986. Earthquakes and rock deformation in crustal fault zones, *Annual Review of Earth and Planetary Sciences* (14), 149-175. doi:10.1146/annurev.ea.14.050186.001053
- [159] Sieh, K., Natawidjaja, D. H., Meltzner, A. J., Shen, C.-C., Cheng, H., Li, K.-S., Suwargadi, B. W., Galetzka, J., Philiposian, B., Edwards, R. L., 2008. Earthquake

- supercycles inferred from sea-level changes recorded in the corals of West Sumatra, *Science* 322, 1674–1678. doi:10.1126/science.1163589
- [160] Sieh, K., Stuiver, M., Brillinger, D., 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
- [161] Silver, P., 1983. Retrieval of source-extent parameters and the interpretation of corner frequency, *Bull. Seismol. Soc. Am.* 73 1499–1511.
- [162] Stein, R. S., Barka, A. A., Dieterich, J. H., 1997. Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering, *Geophys. J. Int.* 128, 594–604. doi:10.1111/j.1365-246X.1997.tb05321.x
- [163] Stein, S., Cloetingh, S., Sleep, N. H., Wortel, R., 1989. Passive margin earthquakes, stresses and rheology, in *Earthquakes at North Atlantic Passive Margins: Neotectonics and Postglacial Rebound*. NATO ASI Series, 266, 231-259, Kluwer Academic Publ, Dordrecht. doi:10.1007/978-94-009-2311-9\_14
- [164] Stein, S., Geller, R. J., Liu, M., 2012. Why earthquake hazard maps often fail and what to do about it. *Tectonophysics*, 562–563, 1-25. doi:10.1016/j.tecto.2012.06.047
- [165] Stein, S., Liu, M., 2009. Long aftershock sequences within continents and implications for earthquake hazard assessment. *Nature*, 462(5), 5-12. doi:10.1038/nature08502
- [166] Stein, S., Liu, M., Calais, E., Quingsong, L., 2009. Mid-continent earthquakes as a complex system. *Seismol. Res. Lett.*, 80(4), 551-553. doi:10.1785/gssrl.80.4.551
- [167] Stein, S., Newman, A., (2004), Characteristic and uncharacteristic earthquakes as possible artifacts: applications to the New Madrid and Wabash seismic zones. *Seismol. Res. Lett.*, 75(2), 173-187. doi:10.1785/gssrl.75.2.173
- [168] Stein, S., Okal, E. A., 2007. Ultralong period seismic study of the December 2004 Indian Ocean earthquake and implications for regional tectonics and subduction process. *Bull. Seismol. Soc. Amer.* 97(1A), 279-295. doi:10.1785/0120050617
- [169] Stein, S., Okal, E. A., 2011. The size of the 2011 Tohoku earthquake needn't have been a surprise. *Eos Trans. AGU*, 92(27), 227-228. doi:10.1029/2011EO270005
- [170] Stein, S., Sleep, N. H., Geller, R. J., Wang, S. C., Kroeger, G. C., 1979. Earthquakes along the passive margin of eastern Canada. *Geophys. Res. Lett.*, 6(7), 537-540. doi:10.1029/GL006i007p00537

- [171] Stein, S., Wysession, M., 2003. An introduction to seismology, earthquakes, and earth structure, Blackwell Publishing.
- [172] Styron R., Pagani, M., 2020. The GEM Global Active Faults Database (GAF-DB), *Earthquake Spectra* 36, 160-180. doi:10.1177/8755293020944182
- [173] Suter, M., 2015. Rupture of the Pitáycachi Fault in the 1887  $M_w$  7.5 Sonora, Mexico earthquake (southern Basin-and-Range Province): Rupture kinematics and epicenter inferred from rupture branching patterns, *J. Geophys. Res.: Solid Earth* 120, 617–641. doi:10.1002/2014JB011244
- [174] Swafford, L., Stein, S., 2007. Limitations of the short earthquake record for seismicity and seismic hazard studies, in *Continental Intraplate Earthquakes: Science, Hazard, and Policy Issues*, edited by S. Stein and S. Mazzotti, 49-58, Geolog. Soc Amer Inc, Boulder.
- [175] Toda, S., Stein, R. S., Reasenber, P. A., Dieterich, J. H., Yoshida, A., 1998. Stress transferred by the 1995  $M_w = 6.9$  Kobe, Japan, shock: Effect on aftershocks and future earthquake probabilities, *J. Geophys. Res.: Solid Earth* 103, 24543–24565. doi:10.1029/98JB00765
- [176] Todes J. P., Okal, E. A., Kirby, S. H., 2021. Frequency-size distributions of Wadati-Benioff zone and near-boundary intraplate earthquakes: Implications for intermediate and deep seismicity, *Phys. Earth Planet. In.* 321, 106707. doi:10.1016/j.pepi.2021.106707
- [177] Turcotte, D.L., Schubert, G., 2014, *Geodynamics*, Cambridge University Press, USA.
- [178] U.S. Nuclear Regulatory Commission, 2012. Central and eastern United States seismic source characterization for nuclear facilities (NUREG-2115). EPRI, Palo Alto, CA, U.S. DOE, and U.S. NRC, <http://www.ceus-ssc.com>.
- [179] Velasco, A.A., Ammon, C.J., Lay, T., Hagerty, M., 1996. Rupture process of the 1990 Luzon, Philippines ( $M_w = 7.7$ ), earthquake, *J. of Geophys. Res.* 101, 22,419–22,434. doi:10.1029/96JB02290
- [180] Vallée, M., Douet D., 2016. A new database of source time functions (STFs) extracted from the SCARDEC method, *Phys. Earth Planet. In.* 257, 149-157. doi:10.1016/j.pepi.2016.05.012

- [181] Vanneste, K., Vleminckx, B., Stein, S., Camelbeeck, T., 2016. Could  $M_{max}$  be the same for all continental interiors? *Seismol. Res. Lett.*, 87(5). doi:10.1785/0220150203
- [182] Votsi, I., Limnios, N., Tsaklidis, G., Papadimitriou, E., 2013. Hidden Markov models revealing the stress field underlying the earthquake generation, *Physica A: Statistical Mechanics and its Applications* 392, 2868–2885. doi:10.1016/j.physa.2012.12.043
- [183] Wallace, R. E., 1970. Earthquake recurrence intervals on the San Andreas fault, *GSA Bulletin* 81, 2875–2890. doi:10.1130/0016-7606(1970)81[2875:ERIOTS]2.0.CO;2
- [184] Warren, L., Shearer, P., 2002. Mapping lateral variations in upper mantle attenuation by stacking P and PP spectra, *J. Geophys. Res.* 107 2342. doi:10.1029/2001JB001195
- [185] Weichert, D. H., 1980. Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. *Bull. Seismol. Soc. Amer.*, 70(4), 1337-1346.
- [186] Wiens, D. A., Stein, S., 1984. Intraplate seismicity and stresses in young oceanic lithosphere, *J. Geophys. Res.* 89, 11442-11464. doi:10.1029/JB089iB13p11442
- [187] Weldon, R. J., Fumal, T. E., Biasi, G. P., 2004. Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work, *GSA Today* 14, 4–10. doi:10.1130/1052-5173(2004)014;4:WATECW;2.0.CO;2
- [188] 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
- [189] Wells, D. L., Coppersmith, K. J., 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull. Seismol. Soc. Amer.*, 84(4), 974-1002.
- [190] Wernicke, B., 1995. Low-angle normal faults and seismicity: A review, *J. Geophys. Res.: Solid Earth* 100(B10), 20159– 20174. doi:10.1029/95JB01911
- [191] Wesnousky, S.G., 1988. Seismological and structural evolution of strike-slip faults, *Nature* 335, 340-343. doi:10.1038/335340a0
- [192] Wesnousky, S.G., 2008. Displacement and geometrical characteristics of earthquake surface ruptures: Issues and implications for seismic-hazard analysis and

- the process of earthquake rupture, *Bull. Seismol. Soc. Am.* 98 (4), 1609–1632. doi:10.1785/0120070111
- [193] Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J. F., Wobbe, F., 2013. Generic Mapping Tools: Improved version released, *EOS Trans. AGU* 94 409-410. doi:10.1002/2013EO450001
- [194] Wheeler, R. L., 2009. Methods of  $M_{max}$  estimation east of the Rocky Mountains. U.S. Geol. Survey Open File Report 2009-1018.
- [195] Whitney, J.M., Knight, M., 1980. The relationship between tensile strength and flexure strength in fiber-reinforced composites, *Experimental Mechanics* 20, 211-216.
- [196] Wolin, E., Stein, S., Pazzaglia, F., Meltzer, A., Kafka, A., Berti, C., 2012. Mineral, Virginia, earthquake illustrates seismicity of a passive-aggressive margin. *Geophys. Res. Lett.*, 39(2). doi:10.1029/2011GL050310
- [197] Wyss, M., Nekrasova, A., Kossobokov, V., 2012. Errors in expected human losses due to incorrect seismic hazard estimates. *Nat. Hazards*, 62(3), 927-935.
- [198] Yu, G., Xu, X., Klinger, Y., Diao, G., Chen, G., Feng, X., Li, C., Zhu, A., Yuan, R., Guo, T., Sun, X. et al. (2010) Fault-scarp features and cascading-rupture model for the  $M_w$  7.9 Wenchuan earthquake, eastern Tibetan plateau, China. *Bulletin of the Seismological Society of America* 100, 2590–2614. doi:10.1785/0120090255
- [199] Zhuang J., Touati, S., 2015. Stochastic simulation of earthquake catalogs. Community Online Resource for Statistical Seismicity Analysis. doi:10.5078/corssa-43806322