NORTHWESTERN UNIVERSITY

Investigating Earthquake Recurrence and Hazard Models

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

Leah Marschall Salditch

EVANSTON, ILLINOIS

December 2021

ABSTRACT

Investigating Earthquake Recurrence and Hazard Models

Leah Marschall Salditch

How well do Probabilistic Seismic Hazard Analysis (PSHA) maps forecast ground shaking due to earthquakes? This question is central to ensuring the safety, security, and economic well-being of citizens. PSHA maps are an important product for users including seismologists, engineers, insurers, and policymakers. PSHA, which has been used worldwide for almost 50 years, uses estimates of the probability of future earthquakes and the resulting shaking to predict the shaking expected with a certain probability over a given time (Cornell, 1968; Field, 2014). Extensive research is ongoing into how well hazard maps perform relative to these expectations and how the maps can be improved.

This dissertation explores aspects of PSHA and its components – from the challenges of modelling earthquake histories with temporal clusters, to the task of comparing model predictions with observations of shaking intensity using performance metrics. The former proposes an alternative to the traditional earthquake cycle model in which a fault's past influences its future likelihood of experiencing an earthquake. The latter takes advantage of hindcasting – using past data to evaluate models which forecast the future. Ideally, we

would compare hazard maps to data collected after the maps were made, however, due to the long recurrence times of large earthquakes relative to post-map observation periods, this is not typically feasible.

Part of this research involved collecting, compiling, and consistently interpreting 162 years of seismic intensity data in the California Historical Intensity Mapping Project (CHIMP). Using CHIMP data and historical seismic intensity compilations from Italy, France, and Nepal, I compared the maximum observed shaking in an area to that predicted by PSHA models. Assuming the datasets to be correct, it appears that PSHA models overpredict shaking, even correcting for the time period involved. Assuming the PSHA models are correct, a shaking deficit exists between the model and observations. Possible reasons for this apparent discrepancy between the model and observations are threefold: 1) the observations could be biased low; 2) the observation period has been less seismically active than typical – either by random chance or temporal variability due to stress shadow effects; 3) the model overpredicts, due to limitations of either the earthquake rupture forecast or the ground motion models (GMMs).

Acknowledgements

"Things happens for a reason, and that reason is random chance."

– New Girl, Season 6, Episode 4

First and foremost, I thank my principal advisor Seth Stein for his unwavering support and guidance throughout my PhD and for always believing in me, even when I did not believe in myself. I also thank my advisors Bruce Spencer, Susan Hough, and Norm Abrahamson for their mentorship. Specifically, I thank Bruce for teaching me the principles of Zen and the Art of Statistical Modeling. I thank Sue for being an all-around inspiration and for generously sharing her creative vision. I thank Norm for his passion and patience. I consider myself extremely lucky to have formed a thesis committee made up of the most brilliant and, more importantly, kind people I have encountered in academia.

I thank my partner Art Zachary for his constant cheerleading, in-house math tutoring, and of course for help coining the CHIMP acronym. You are the real MVP.

Finally, I thank my family for their love and support. I could not have done this without you.

I am grateful for the random series of events that led me here.

Contents

List of Tables

List of Figures

CHAPTER 1

Thesis Overview

1.1. Introduction

In this thesis, I investigate earthquake recurrence and hazard models to identify weaknesses and potentially create improvements. I first propose an alternative to the traditional earthquake cycle model, which I call Long-Term Fault Memory (LTFM), that allows a fault's past earthquake history and present state of strain to influence its future probability of an earthquake. This model was necessary because historical and paleoseismic data show that earthquake histories are more complex than previous models could accommodate. This thesis also assesses the performance of various PSHA models using historical data. To do this, I compared earthquake hazard model predictions with observations of shaking intensity collected over hundreds of years, going back as far as the historical record allows. The comparison indicates that there may be a systematic bias either in the models or the observations. Assuming the data are without error and that they represent long-term hazard, the hazard maps overpredict shaking. Assuming the maps to be correct, there is a shaking deficit in the observed record.

1.2. Chapter 2: Earthquake Supercycles and Long-Term Fault Memory

Large earthquakes are known to occur in temporal clusters, also known as supercycles, based on paleoseismic records. This implies that not all strain accumulated on a fault is released in an earthquake, so strain can remain high, resulting in multiple earthquakes occurring close in time to one another. This is a problem for the commonly used statistical models of earthquake occurrence, which must assume that earthquakes occur independently of one another. This independence is accomplished by asserting that all strain built up in the inter-seismic period is completely released in each earthquake, so that the history of strain has no effect on future earthquakes. However, earthquake histories suggest that some strain remains on a fault after an earthquake. The amount of remaining strain depends on the complex distribution of stress in the earth's crust and properties of the specific source properties of an earthquake. The Long-Term Fault Memory model successfully simulate paleo-earthquake records that exhibit temporal clusters, helping us to understand the gross characteristics and behavior of faults in different tectonic regimes.

This chapter is published as Salditch, L., S. Stein, J. Neely, B.D. Spencer, E.M. Brooks, A. Agnon, M. Liu (2020). Earthquake Supercycles and Long-Term Fault Memory, Tectonophysics 774, 228289. doi: https://doi.org/10.1016/j.tecto.2019.228289.

1.3. Chapter 3: 1952 Kern County Earthquake

In chapter 3, I investigate the distribution of seismic intensity observations from the 1952 Kern County, California earthquake. I reinterpret these observations using the latest interpretation standards, which have changed since the observations were originally collected and assigned values. Using these intensity assignments and established Intensity Prediction Equations (IPEs), I compare the distribution of seismic intensity with distance from the White Wolf Fault, which ruptured in this earthquake. Inverting for the magnitude of the earthquake independently of instrumental data, which was of poor quality at the time, yields an estimate of MW 7.2 ± 0.2, on the lower end of previous magnitude estimates.

This chapter is published as Salditch, L., S. E. Hough, S. Stein, B. D. Spencer, E. M. Brooks, J. S. Neely, and M. C. Lucas (2018). The 1952 Kern County, California earthquake: A case study of issues in the analysis of historical intensity data for estimation of source parameters, Physics of the Earth and Planetary Interiors, 283, 140-151, doi: 10.1016/j.pepi.2018.08.007.

1.4. Chapter 4: California Historical Intensity Mapping Project (CHIMP)

Building off of the Kern County investigation, Chapter 4 details the ambitious California Historical Intensity Mapping Project (CHIMP). CHIMP is a dataset of reinterpreted seismic intensity observations from 62 of the largest earthquakes in California from 1857 to 2019. CHIMP brings together intensity information from disparate and often hard to access compilations. The result is a consistently reinterpreted history of shaking in California. I compare the maximum observed intensities with shaking levels predicted by the USGS 2018 seismic hazard maps using performance metrics developed by Stein et al., 2015. The primary metric is the fractional exceedance metric, M0, which is implicit in PSHA. I also use the M1 metric, which is akin to a visual comparison. The results show a discrepancy indicating either that the hazard map overpredicts shaking or that there is a shaking deficit in CHIMP.

This chapter is published as Salditch, L., M. M. Gallahue, M. C. Lucas, J. S. Neely, S. E. Hough, and S. Stein (2020). California Historical Intensity Mapping Project (CHIMP): A Consistently Reinterpreted Dataset of Seismic Intensities for the Past 162 Yr and Implications for Seismic Hazard Maps, Seismological Research Letters 91(5), 1 20, doi: 10.1785/0220200065.

1.5. Chapter 5: Comparing Maps

Performance metrics are useful because they allow direct comparisons to be made between maps of the same region. This is necessary because plausible alternative input parameters and geologic assumptions, based on the limited available data, can produce quite different maps. This chapter examines the performance of alternative maps by looking at two examples. In Nepal, differences in seismic source and ground-motion models impact the predicted hazard in a high seismicity region. In France, I investigate the difference between characterizing the hazard using the mean or median of the models. The results indicate that the median may be a better indicator of map performance than the mean, which has important implications for hazard assessment.

1.6. Chapter 6: Conclusion

Chapter 6 summarizes the results of hazard map performance assessments to date. Looking at regions from a variety of tectonic settings (Nepal, France, Japan, Italy, and intraplate and plate boundary zones of the U.S.A.), we find a consistent bias between the hazard map predictions and the historical observations of shaking. Theoretically, the longer the observation time is relative to the map return period, the better a map should predict observed exceedance levels. However, my results do not show this. In fact, the best performing hazard maps, the 2016 and 2017 1-year forecasts of natural and induced seismicity in the Central and Eastern United States, are the shortest in terms of observation period and map return period. The cause of this bias remains unclear, but has important implications for the future of earthquake hazard assessments.

CHAPTER 2

Earthquake Supercycles and Long-Term Fault Memory

2.1. Summary

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

probability estimates of the next earthquake will depend crucially on whether the cluster is treated as ongoing or over.

2.2. Introduction

Since the 1906 San Francisco earthquake, the dominant paradigm in earthquake seismology has been the earthquake cycle, in which strain accumulates between large earthquakes due to interseismic motion between the two sides of a locked fault and is released by coseismic slip on the fault when an earthquake occurs (Reid, 1910). Over time, this process should give rise to approximately periodic earthquakes and a steady accumulation of cumulative displacement (Figure 2.1a). The fact that earthquake sequences are only approximately periodic prompted a refinement of the model with"time-predictable" recurrence in which a specific strain level must accumulate for an earthquake, but the strain release in the earthquake is variable (Shimazaki and Nakata, 1980).

However, long earthquake records often show more complex behavior (Figure 2.1b). Wallace (1987) found that faults and groups of faults in the Western U.S.'s Great Basin often showed "grouping, a series of displacement events, each being followed by a period of quiescence. Slip rates during a group of events along a segment of fault, thus, could be considerably greater than the long-term average slip rate. During quiescent periods, the slip rate would be lower than the average rate and might even be zero... Additionally, if grouping is real, the concept that accumulated elastic strain is released at some regular interval by a single displacement event in a seismic cycle should be reexamined. Perhaps strain that has accumulated at a more or less constant rate is released in a stuttering, spasmodic manner in a group of displacement events." Subsequent investigation (Friedrich

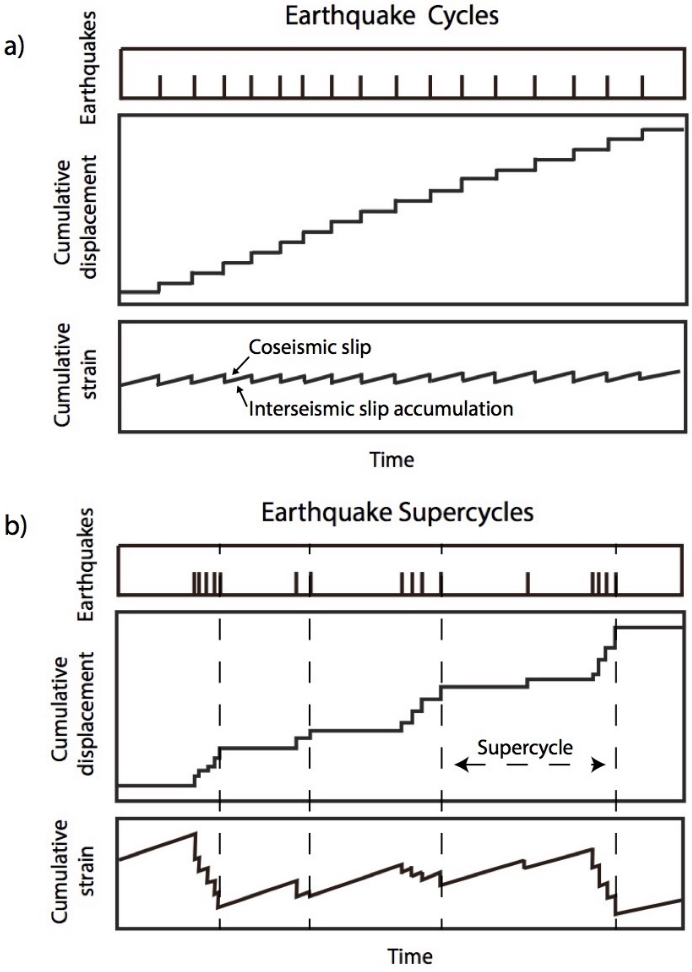

Figure 2.1. Schematic comparison of the histories of earthquake ocurrence, cumulative displacement, and cumulative strain for a fault without supercycles (a) and a fault with supercycles (b). Adapted from Wallace (1987) and Friedrich et al. (2003).

et al., 2003) supported this analysis, finding that “seismic strain release may be clustered on the 10-kyr timescale... with comparatively low, uniform strain accumulation rates on the 100-kyr timescale.” They suggested calling the conventional earthquake cycles “Reid-type” behavior and the longer period variations “Wallace-type” behavior.

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

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

Conceptually, the history of strain accumulation and release is the underlying process that gives rise to patterns in the resulting earthquakes and cumulative displacement. In the schematic example of Figure 2.1b, supercycles appear as patterns longer than individual earthquake cycles in the earthquake history, cumulative displacement, and cumulative strain records. The fullest picture is given by the strain record. This infers the strain by combining data about strain release via slip in earthquakes over time with the interseismic strain accumulation inferred from the slip between earthquakes taken from present-day geodetic, long-term geological, or other data. The cumulative displacement record shows the dates of earthquakes and the coseismic slip in each, whereas the earthquake history gives only the dates. The displacement record can be viewed as the time derivative of the strain record, and the earthquake history can be viewed as the time derivative of the displacement record, with each differentiation involving a loss of information. Conversely, constructing a displacement record requires supplementing an earthquake history with coseismic slip data, and constructing the strain record then involves also including data or assumptions about the interseismic strain accumulation. Hence the earthquake history has the least uncertainty, and the displacement and strain records have progressively larger uncertainties.

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

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

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

2.3. Examples of Supercycles

Supercycles and/or clustering have been observed in many tectonic environments (Figure 2.2). The best data come from earthquake histories at plate boundaries, because the relatively rapid plate motion (typically < 5 mm/yr) gives shorter and hence easier-to-observe cycles. In some cases, the slip and strain history also show evidence for supercycles.

Weldon et al. (2004) used the dates and offset in paleoearthquakes since 500 CE across the San Andreas fault near Wrightwood, California, together with the interseismic slip rate observed from present-day geodesy and long-term geological rates, to reconstruct the history of strain accumulation and release (Figure 2.2a). They argue that *"it is hard to escape the conclusion that strain accumulated over many earthquake cycles was responsible for the flurry of large slip events."* Nearby, at Pallett Creek, Sieh et al. (1989) find that paleoearthquakes occurred in clusters within which they were *"separated by periods of several decades, but the clusters are separated by dormant periods of two to three centuries."* To the south, where the San Jacinto fault takes up some of the motion between the Pacific and North America plates, Rockwell et al. (2015) find that *"for much of the past 4,000 years the fault ruptured in a quasi-periodic fashion. In the past 1,000*

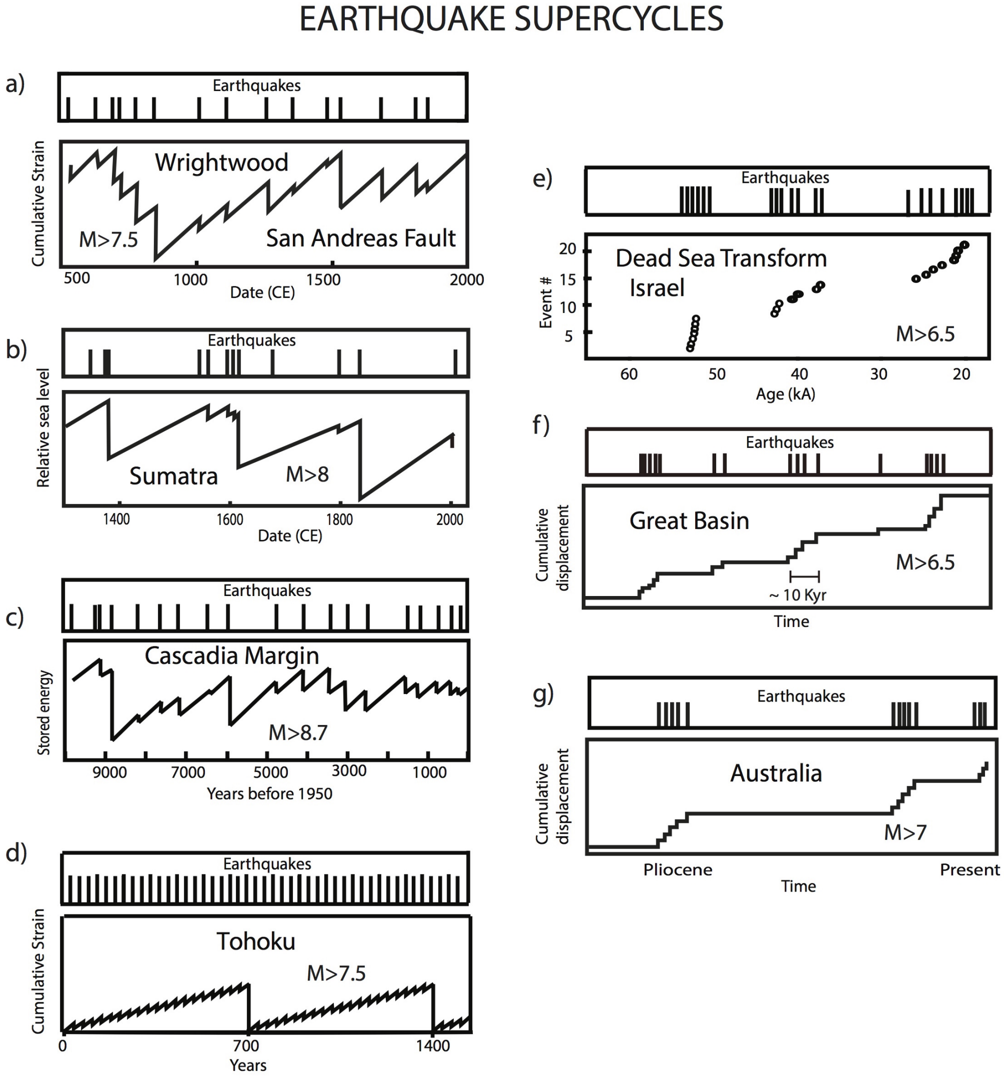

Figure 2.2. Examples of reported supercycles. a) Strain accumulation and release inferred from paleoseismic data across the San Andreas fault (Weldon et al., 2004). b) Supercycles on the Sumatra megathrust inferred from corals (Sieh et al., 2008). c) Long-term energy cycling inferred from turbidites on the Cascadia megathrust (Goldfinger et al., 2013). d) Schematic earthquake history for the Japan Trench off Tohoku (Satake, 2015). e) Earthquake history on the Dead Sea transform (Agnon, 2014). f) Schematic earthquake history for faults and groups of faults in the Western U.S.'s Great Basin (Wallace, 1987). g) Schematic earthquake history for faults in Australia (Clark et al., 2012).

years, in contrast, a flurry or cluster of four earthquakes occurred in a 150-year period, and the overall recurrence interval is much shorter."

Sieh et al. (2008) analyzed relative sea level changes recorded by corals from Sumatra, which show interseismic subsidence and coseismic uplift (Figure 2.2b). They infer that *"this 700-kilometer-long section of the Sunda megathrust has generated broadly similar sequences of great earthquakes about every two centuries for at least the past 700 years... Because each of the three past episodes of emergence consists of two or more discrete events, we refer to the broad periods of strain accumulation and relief as supercycles rather than merely cycles."* To the north along the subduction zone, Rubin et al. (2017) studied a 4500 year sequence of at least 11 tsunami deposits and find that *"the average time period between tsunamis is about 450 years with intervals ranging from a long, dormant period of over 2,000 years, to multiple tsunamis within the span of a century... these variable recurrence intervals suggest that long dormant periods may follow Sunda megathrust ruptures as large as that of the 2004 Indian Ocean tsunami."*

The dates and volumes of turbidite deposits, assumed to have been generated by great earthquakes on the Cascadia megathrust (Adams, 1990), show evidence for supercycles. Using these to infer the history of strain energy accumulation and release (Figure 2.2c), Goldfinger et al. (2013) find that *"the resulting sawtooth pattern reveals what we interpret as a complex pattern of long-term energy cycling on the Cascadia megathrust... Overall, what is suggested by this pattern is that some events release less energy, whereas others release more energy than available from plate convergence (slip deficit) and may have borrowed stored energy from previous cycles".* Although an additional event has been identified (Goldfinger et al., 2017), the inferred strain energy history would not be substantially altered. Kelsey et al.'s (2005) analysis of coastal deposits that record local tsunamis and seismic shaking finds that *"over the 4600 yr period when Bradley Lake was an optimum tsunami recorder, tsunamis from Cascadia plate-boundary earthquakes came in clusters."*

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

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

In other areas, supercycles have been inferred from the earthquake history, even though the strain history requires data on the slip in individual events. Agnon's (2014) analysis of a long record of seismites, sediment records of earthquake shaking (Marco et al., 1996), along the Dead Sea Transform in Israel (Figure 2.2e) finds *"a pattern of long quiescence periods between quasi-periodic clusters. During each cluster of seismicity the recurrence interval is quite uniform, varying among clusters between 200 and 1,400 years. Quiescence periods may linger 3-10,000 years."* Further north on this transform, Wechsler et al. (2014) find that *"the interevent time of surface-rupturing earthquakes varies by a factor of two to four during the past 2 ka at our site, and the fault?s behavior is not time predictable."*

Supercycles and/or clustering have also been identified in plate boundary zones, where diffuse deformation is spread over multiple faults with long-term slip rates typically slower than on the primary plate boundary faults, and in continental interiors, which typically deform at < 1 mm/yr. As noted earlier, paleoseismic data from faults and groups of faults in the Western U.S.'s Great Basin (Wallace, 1987), part of the broad boundary zone that accommodates motion between the Pacific and North American plates, often show *"clustered strain release and uniform, low strain accumulation"* (Friedrich et al., 2003), shown schematically in Figure 2.2f.

Topographic data within the Australian plate, where erosion is very slow, provide some of the best evidence available of how continental intraplate faults slip over time, shown schematically in Figure 2.2g. Clark et al. (2012) found that *"a common characteristic of morphogenic earthquake occurrence in Australia appears to be temporal clustering. Periods of earthquake activity comprising a finite number of large events are separated by much longer periods of seismic quiescence, at the scale of a single fault and of proximal faults. In several instances there is evidence for deformation at scales of several hundred kilometers switching on and off over the last several million years."* As result, *"assigning an 'active/inactive' label to a fault in a slowly deforming area based upon the occurrence (or non-occurrence) of an event in the last few thousands to tens of thousands of years is not a useful indicator of future seismic potential"* (Clark et al., 2011) and *"it is debatable whether a 'recurrence interval' on individual faults applies"* (Clark, 2003).

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

2.4. Earthquake recurrence models

The most easily studied aspect of supercycles is that they often - but need not always - cause variability in earthquake recurrence interval times, notably temporal clusters (Figure 2.3), which have important consequences for hazard estimation. As a result, many studies focus on possible clusters in a fault's earthquake history and their implications for the recurrence of future large earthquakes.

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

One model treats earthquake occurrence as a Poisson process, in which the probability of a large earthquake is constant with time (Figure 2.3a). This probability depends on the mean recurrence interval μ, such that the probability of at least one event in a time interval t that is short compared to μ is t/μ. In this model the occurrence of a large earthquake does not reduce the probability of another. Hence the fault has no "memory," the dates of previous earthquakes have no effect on when the next will occur, and any clusters resulting from short intervals between events arise purely by chance. As the earthquake record's length increases, the standard deviation of the recurrence intervals approaches the mean. Equality of the mean and standard deviation of inter-event times is a property of a Poisson process, but - as shown later in the paper - other stochastic processes can also have this property.

Because the Poisson model is the simplest recurrence model, it is traditionally used in earthquake hazard modeling and provides a null hypothesis against which other models can be tested (Rundle and Jackson, 1977; Smalley et al., 1987; Kagan and Jackson, 1991; Michael, 1997; Biasi et al., 2002).

An alternative class of probability models is based on the concept of an earthquake cycle (Figure 2.1), in which strain accumulates between large earthquakes and is completely released when one occurs (Reid, 1910; Savage and Burford, 1973; Sykes and Nishenko, 1984; Matthews et al., 2002; Field et al., 2015). In these models, the probability of a large earthquake increases with time until one occurs, at which point the probability drops to

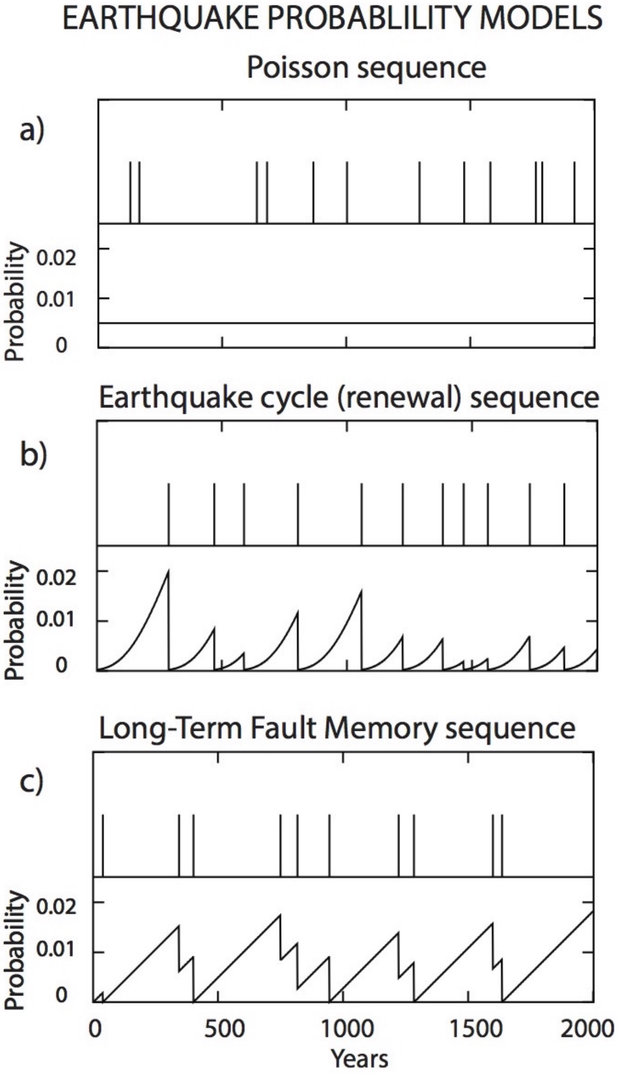

Figure 2.3. Comparison of earthquake recurrence models. a) Poisson process, in which the probability of a large earthquake is constant with time, so the fault has no memory and any clusters resulting from short intervals between events arise purely by chance. b) Earthquake cycle, in which the probability of a large earthquake increases with time until one occurs, at which point the probability drops to zero. The fault has only “short-term memory” because the probability of a large earthquake depends only on the time since the past one. c) Modified earthquake cycle in which after an earthquake the probability decreases, but not necessarily to zero. The fault has ”long-term memory” because the probability depends on the earthquake history over previous cycles.

zero and the cycle begins again (Figure 2.3b). This assumption corresponds to the fault releasing all the strain accumulated on it in each cycle, so strain would not accumulate on timescales longer than individual cycles.

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

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

However, the fact that strain accumulation and/or clusters are observed on many fault systems has led to proposals that they are, at least in part, a real effect due to intrinsic properties of the faulting process (Ben-Zion et al., 1999). Hence in this paper, we take the view that observations of clustering likely reflect supercycles. We thus explore the possibility that faults have "long-term memory," such that the occurrence of large earthquakes depends on earthquake history over multiple previous earthquake cycles (Figure 2.3c).

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

2.5. Characterizing Earthquake Sequences

2.5.1. Aperiodicity

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

A related characterization uses the burstiness parameter

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

$M = \frac{1}{(N-1)} \sum_{i=1}^{N-1} (\tau_i - \mu_1)(\tau_{i+1} - \mu_2)/\sigma_1\sigma_2$ where N is the number of recurrence intervals τ_i , μ_1 and σ_1 are the mean and standard deviations of τ_i $(i = 1, 2, ..., N - 1)$, and μ_2 and σ_2 are the mean and standard deviations of τ_{i+1} $(i = 1, 2, ..., N - 1)$. M ranges from -1 to 1, with $M > 0$ when short interevent times are generally followed by short ones, and long interevent times are generally followed by long ones. $M < 0$ when short interevent times are generally followed by long ones, and vice versa. These arise because M is a normalized form of the autocorrelation of lag one, i.e. the crosscorrelation between the series of interevent times and that series shifted by one.

Figure 2.4 shows the aperiodicities for the earthquake sequences in Figure 2.2. The Wrightwood and Cascadia (Figures 2.2a and 2.2c) sequences have $\alpha = 0.47$ and 0.51, so the time series alone do not indicate the supercycle behavior shown by the strain records. In contrast, the Sumatra and Dead Sea transform (Figures 2.2b and 2.2e) sequences have $\alpha = 1.05$ and 1.6, indicating the supercycle behavior. The Great Basin and Australia sequences (Figures 2.2f and 2.2g) were described schematically without specific dates, so the aperiodicity illustrated is also schematic. Also shown is the global result from Goes (1996), who compiled 52 earthquake sequences from the San Andreas fault and the Middle America, Alaska, Chile, and Japan trenches. She found aperiodicities varying from 0.0 to 1.7, with "a large average aperiodicity" of 0.72 ± 0.36 that she interpreted as showing that earthquake recurrence is more irregular "than often assumed in hazard analysis."

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

i) *Sequences with the same aperiodicity can be quite different.* Because the aperiodicity depends only on the mean and standard deviation of the interevent times, it does not depend on the order of events. Thus quite different sequences can have the same aperiodicity (Cowie et al., 2012). Figure 5a shows a sequence of paleoearthquakes composed of clusters of events several decades apart, separated by gaps of two to three centuries. The sequence has $\alpha = 0.79$, showing strong aperiodicity. Grouping the short-interval events together (Figure 2.5b) does not change α, but we would probably view the sequence as

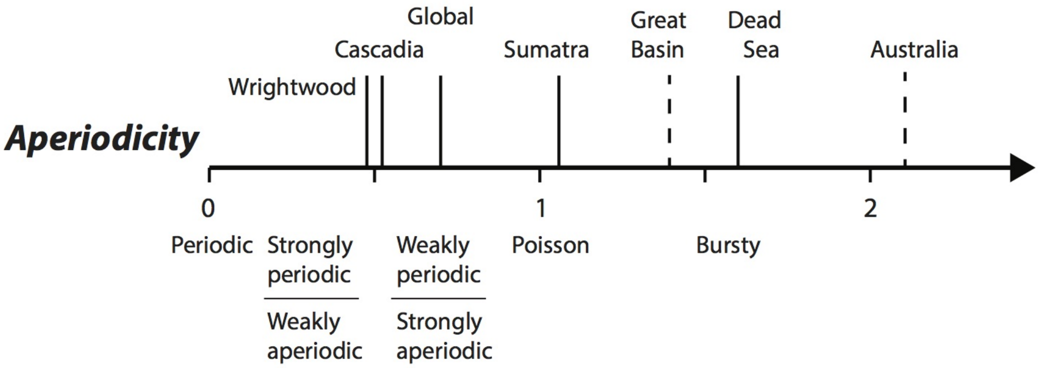

Figure 2.4. Illustration of characterizing earthquake sequences by their aperiodicity, which measures the extent that a sequence differs from perfectly periodic. $\alpha = 1$ is expected for an ideal Poisson process, but can arise from other stochastic processes. Values are shown for examples in Figure 2.2. Solid bars show sequences with dates and dashed bars show schematic sequences with approximate aperiodicites. Also shown is the result from Goes' (1996) global compilation.

showing a change from longer recurrence times in the past to more recent short recurrence times. The memory parameter illustrates the difference, in that the more clustered sequence has a negative value, $M = -0.28$, whereas the grouped sequence has $M = 0.70$. This difference between the two sequences can also be seen in the interevent time plots shown to the right of each sequence. In these, major gaps appear as interevent times longer than the mean, which is shown by a horizontal line. In the first sequence, short and long intervals generally alternate, giving clusters and negative values of M. In the second sequence, short and long intervals are grouped, giving a positive memory.

ii) *Sequences with "quasiperiodic" aperiodicity can be quite clustered.* Earthquake sequences that we would consider clustered can fall below the nominal burstiness criterion of $\alpha > 1$. Figure 2.6b shows that lengthening the three major gaps in Figure 2.6a by 100 years increases the aperiodicity from 0.79 to 0.92, making the clustering stronger and the

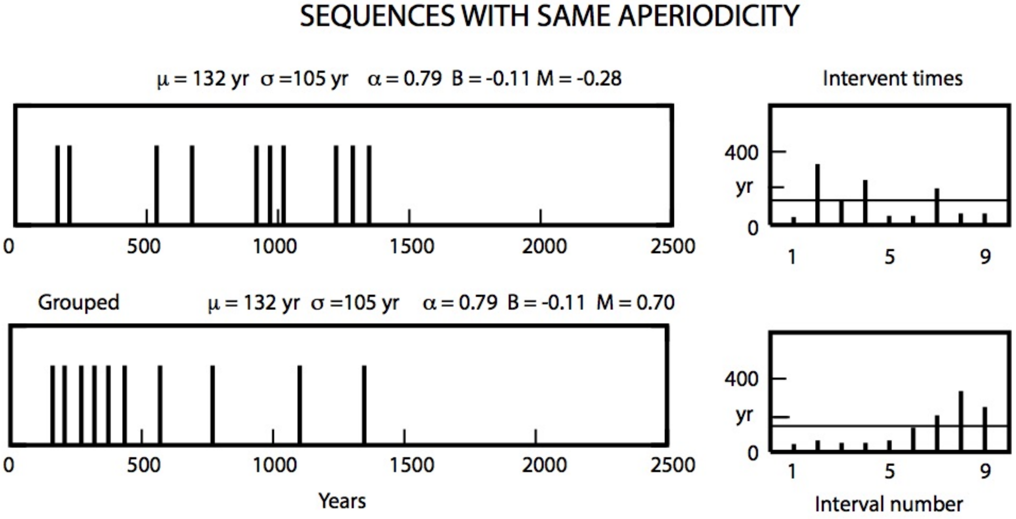

Figure 2.5. Comparison of two sequences with the same aperiodicity. a) Sequence with strong aperiodicity ($\alpha = 0.79$) showing clustering. (b) Same sequence with the short-interval events grouped together, which does not show clustering.

weak periodicity even weaker. Lengthening the gaps by 300 years (Figure 2.6c) increases the aperiodicity to 1.08. In all three panels we assume that observations begin at the earliest observed event (at the right side of the time axis showing years before present), so no gap is observed prior to the earliest event. This example illustrates that a sequence must be very strongly clustered to be bursty.

iii) *Sequences with aperiodicity close to 1 need not result from a Poisson process.* Earthquake records with aperiodicity close to 1 could resemble those that would be generated by a Poisson process. However, other stochastic processes, including the Long-Term Fault Memory process discussed later in this paper, can also generate earthquake records with interevent times whose mean and standard deviation are similar. Hence given the evidence in some areas of an underlying process involving strain supercycles, we think it useful to

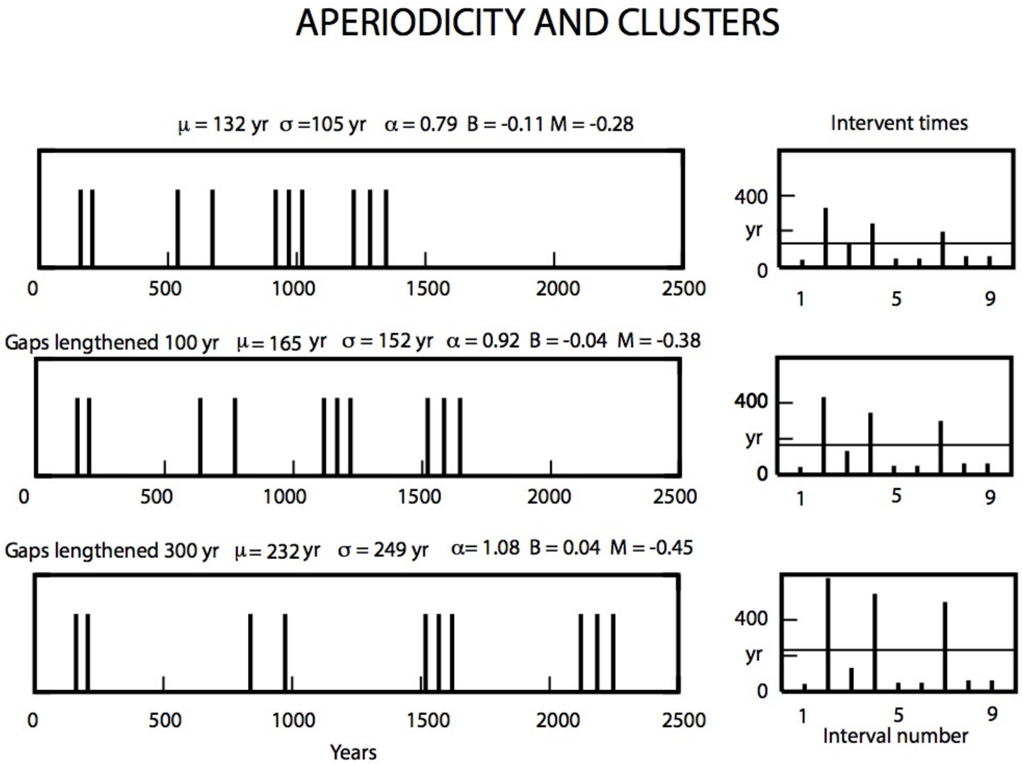

Figure 2.6. Illustration of the fact that "quasiperiodic" ($\alpha < 1$) sequences can be quite clustered. a) Initial sequence with $\alpha = 0.79$. b) Same sequence with major gaps lengthened by 100 years, raising aperiodicity to $\alpha = 0.92$. c) Same sequence with major gaps lengthened by 300 years, raising aperiodicity further to $\alpha = 1.08$. Only c) has aperiodicity above the nominal burstiness criterion of $\alpha > 1$.

consider such sequences as clustered in many senses. In particular, considering clustering in such cases means that estimates of the probability that the next earthquake will occur within a given time window will depend crucially on whether the cluster is treated as ongoing or over.

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

2.5.2. Cluster Analysis

Another way to characterize earthquake sequences is through clustering. The statistical literature provides several criteria for defining a cluster and how many exist in a sequence. Categorizing clusters could facilitate definition of a supercycle, for example one cluster plus one gap. Hence we briefly review different clustering methods which either assign events to a cluster or choose the number of clusters. Clustering algorithms are broadly classified as either partitioning or hierarchical. To illustrate, we use Sieh et al.'s (1989) record from Pallett Creek, California (Figure 2.7a).

Partitioning methods such as the popular $k-$means algorithm are used to divide a sequence of observations, forming a given number of clusters, k, each observation assigned to one cluster. Other methods, discussed later, are used to determine the number of clusters for a given sequence. In our application, the observations in a sequence are the dates, in years, of n earthquakes in an earthquake record and the clusters are defined as time intervals encompassing the range of dates. In the Pallett Creek record, $n = 10$.

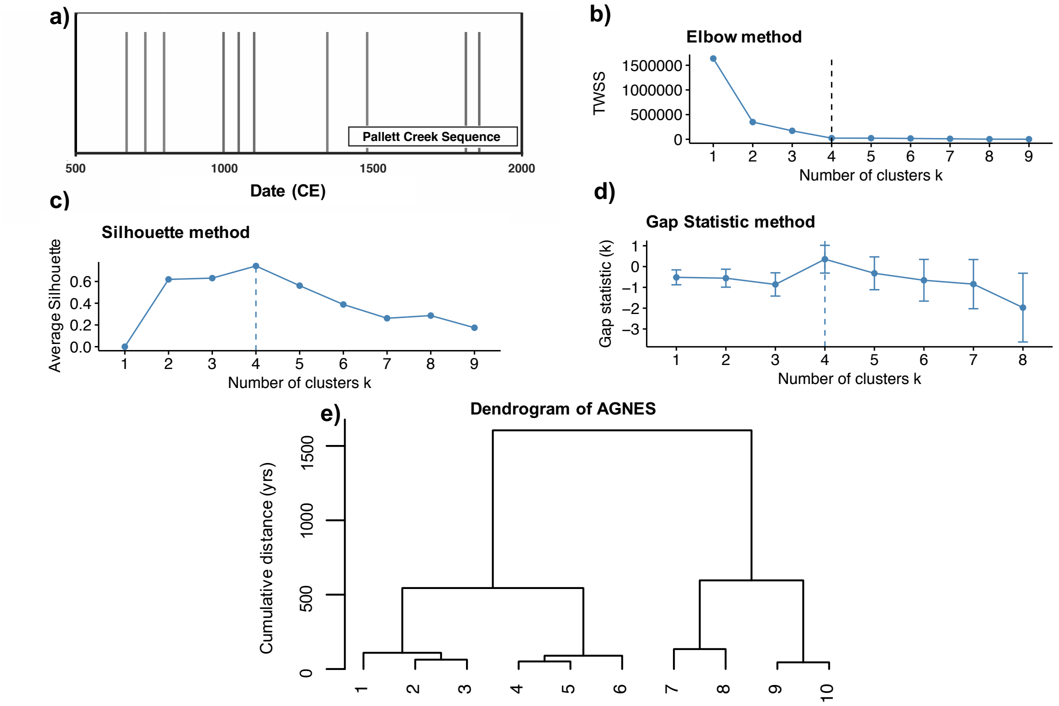

Figure 2.7. Results of different methods to determine the number of clusters in a) Pallett Creek record of Sieh et al., 1989, with event order corresponding to the figure in part e). b) Elbow method where number of clusters is the largest k before increasing k creates only minor improvements of TWSS. c) Silhouette method where maximum value indicates number of clusters. d) Gap statistic method where maximum value indicates number of clusters. e) Hierarchical clustering method using agglomerative nesting (AGNES) with Ward's method; vertical axis shows the cumulative length of time between cluster centers being merged at each step.

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

Some methods, such as the Elbow method, examine the graph of TWSS versus k (Figure 2.7b) and choose the value of k corresponding to a kink in the plot resembling a bent elbow (Tibshirani et al., 2001). Increasing k beyond this value conveys a lesser reduction in TWSS. By this method, Pallett Creek has 4 clusters. The Silhouette method compares the tightness (length of clusters) and separation (distance between clusters) to determine whether the cluster lengths are small compared to the distances between-clusters (Rousseeuw, 1987). Each observation receives a silhouette value, ranging from -1 to +1, indicating the extent to which the observation is well matched to its assigned cluster and poorly matched to the others. The number of clusters k is chosen to maximize the average values for all observations; again $k = 4$ for Pallett Creek (Figure 7c). The Gap method plots two curves that are functions of k, the logarithm of TWSS and its expected value under a uniform distribution of earthquake dates within the record (Tibshirani et al., 2001). The Gap statistic is the distance (gap) between the curves. The chosen value for k has the maximum Gap statistic, which again is $k = 4$ (Figure 2.7d).

Hierarchical methods do not rely on advance specification of the number of clusters, k, but rather create clusters for all $k = 1, ..., n$ possibilities. This process is illustrated by dendrogram plots (tree diagrams), showing the order in which different clusters are merged through connecting branches (Figure 2.7e). The vertical axis shows the cumulative difference in dates (in years) between cluster centers being merged. Hierarchical methods are of two kinds: 1) Divisive, in which all observations start in a single cluster, $k = 1$, and are iteratively separated until $k = n$. 2) Agglomerative nesting (termed AGNES), in which all observations start in their own cluster ($k = n$), with the closest clusters (defined here by years between cluster centers) iteratively joined until $k = 1$ (Kaufman and Rousseeuw, 1990). AGNES may be better at identifying small clusters, while divisive methods may be better at identifying large clusters, although this choice makes no difference for our example. A popular AGNES algorithm, Ward's (1963) method, minimizes the within-cluster sum of squares using an update formula which assigns a new cluster's height on the vertical axis as the cumulative distance between the cluster centers being merged at that step and each step below it (Murtagh and Legendre, 2014). We show Ward's method because it is intended for interval-scaled data such as the dates of earthquakes (Kaufman and Rousseeuw,1990). Figure 2.7e shows the tree resulting from applying Ward's method to Pallett Creek. Clusters that merge at high levels on the vertical axis (indicating large distances between cluster centers being merged at that step) relative to the level of the clusters within them can be interpreted as a 'natural' number of clusters (Hastie et al., 2009). This determination is subjective, so in this example one could reasonably choose 2 or 4 clusters (Figure 2.7e). The four clusters [1,2,3],[4,5,6],[7,8],[9,10] are the same as obtained from k-means with $k = 4$.

Goldfinger's (2012) hierarchical clustering analysis on the 10,000-year-long Cascadia earthquake record found either four or five clusters, using AGNES with complete linkage (furthest neighbor) method. Furthest neighbor defines the distance between two clusters as the distance between two observations, one in each cluster, that are farthest away from one another (Tibrishani et al., 2001b). The two clusters with the shortest distance between them are merged at each step. Applying the complete linkage method to the Pallett Creek record yields the same tree as shown in Figure 2.7e using Ward's method. Goldfinger performed several tests of the statistical significance of the clusters with most resulting in a rejection of an underlying Poisson distribution. He cautions, however, "there is no requirement that physical systems pass statistical tests" (Goldfinger et al., 2012).

Hierarchical methods are complementary to partitioning methods such as k-means. For example, one can use the cluster centers from Ward's method as the initial cluster centers in k -means. As discussed above, results from $k-$means for chosen k can be compared to the results of AGNES. Our results are moderately robust to slight changes in dates, as illustrated by comparing the slightly differing dates of Pallett Creek from Sieh et al. (1989), Biasi et al. (2002), and Scharer et al. (2011). The tree diagrams are the same, because they largely reflect only the events' order. Differences in the k-means evaluations are shown in Table 1. The Gap statistic for Scharer et al.'s dates yields 1 cluster, and the next best number is 4, with the difference between their statistics being quite small compared to the differences between other numbers of clusters. A similar situation occurs in the silhouette for the Sieh et al. dates (Figure 2.7c) where one could argue for 2, 3, or 4 clusters because of the similar values. The Elbow method is the most stable between these different records and the Gap statistic is the least.

Table 2.1. Differences in the number of clusters indicated by three methods for records of earthquakes at Pallett Creek with slightly differing dates.

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

2.6. Long-Term Fault Memory Model

To explore how earthquake supercycles and clusters arise, we use a simple Long-Term Fault Memory (LTFM) model, which is a modified version of the standard earthquake cycle model. In it, the probability of an earthquake reflects the accumulated strain. This increases steadily with time until an earthquake happens, after which it decreases, but not necessarily to zero (Figure 2.8). Hence, the probability of an earthquake depends on the earthquake history over multiple prior cycles. Clusters happen because after a gap, a period of quiescence, the probability can remain higher than the long-term average for several cycles. The model simulates large earthquakes releasing only part of the strain accumulated on the fault, in contrast to the standard model in which all of the accumulated strain is released.

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

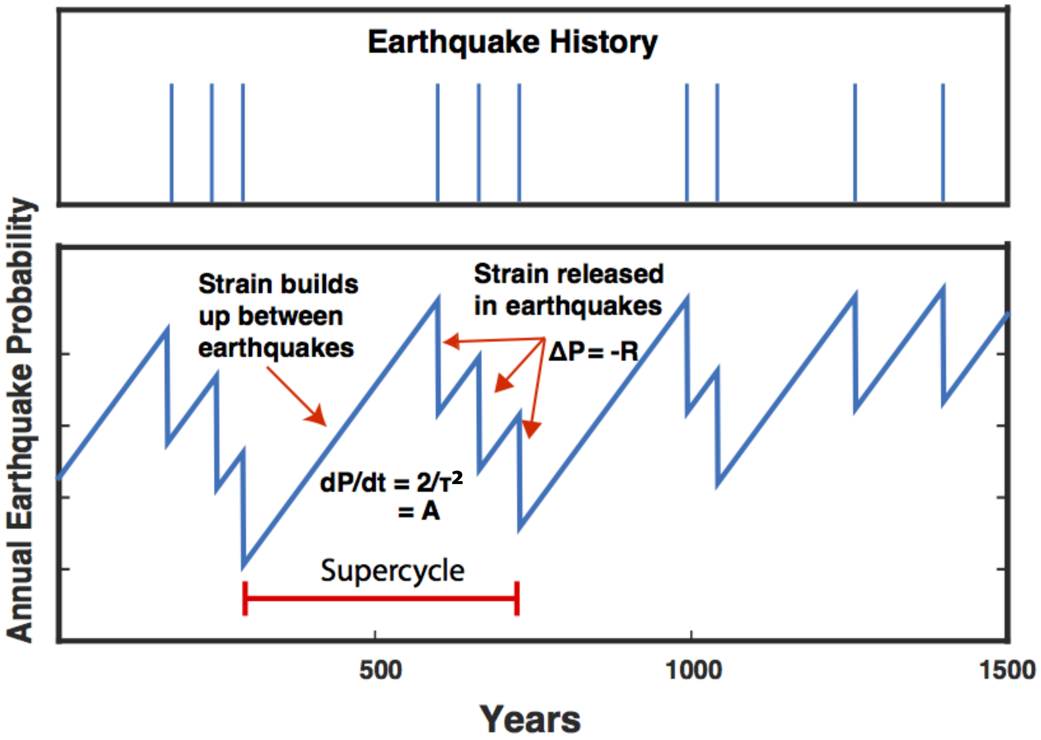

Figure 2.8. Long-Term Fault Memory model. (Top) Simulated earthquake history. (Bottom) Earthquake probability versus time.

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

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

The saw-tooth behavior of LTFM simulates the proposed long-term variations in stored elastic strain or strain energy (Figure 2.2). Supercycles and clusters arise because longer intervals between earthquakes generally begin at times of low probability, consistent with the pattern noted in terms of cumulative strain by Weldon et al. (2004). A lower probability corresponds to the fault having less memory of previous earthquakes. Thus, as the probability (i.e. cumulative strain) approaches zero, the corresponding supercycle can be viewed as approaching a renewal process.

Because LTFM is a stochastic model, the resulting earthquake sequences depend on both the model parameters and chance. As a result (Figure 2.9) sequences can appear strongly periodic, weakly periodic, Poissonian, or bursty. The four sequences in this example have the same probability (i.e. strain) accumulation rate ($A = 2/125^2$) but different release parameters ($R = 200A, 175A, 80A, 50A$). As shown, the aperiodicity increases as R decreases. The strongly periodic sequence arises in a way similar to a standard earthquake cycle model because R is so large that the probability drops to zero after each earthquake, so the fault has no memory. The effects of fault memory increase for successively smaller values of R, making the sequences less periodic. However, A and R control only the overall sequence properties via the probability of earthquake occurrence, because when earthquakes occur is random. As a result, the aperiodicity varies between different portions of the sequence.

In some cases, we use two thresholds, $\delta_2 > \delta_1$ and corresponding probability drops, $R_2 > R_1$, to describe the earthquakes with larger and smaller strain changes implied

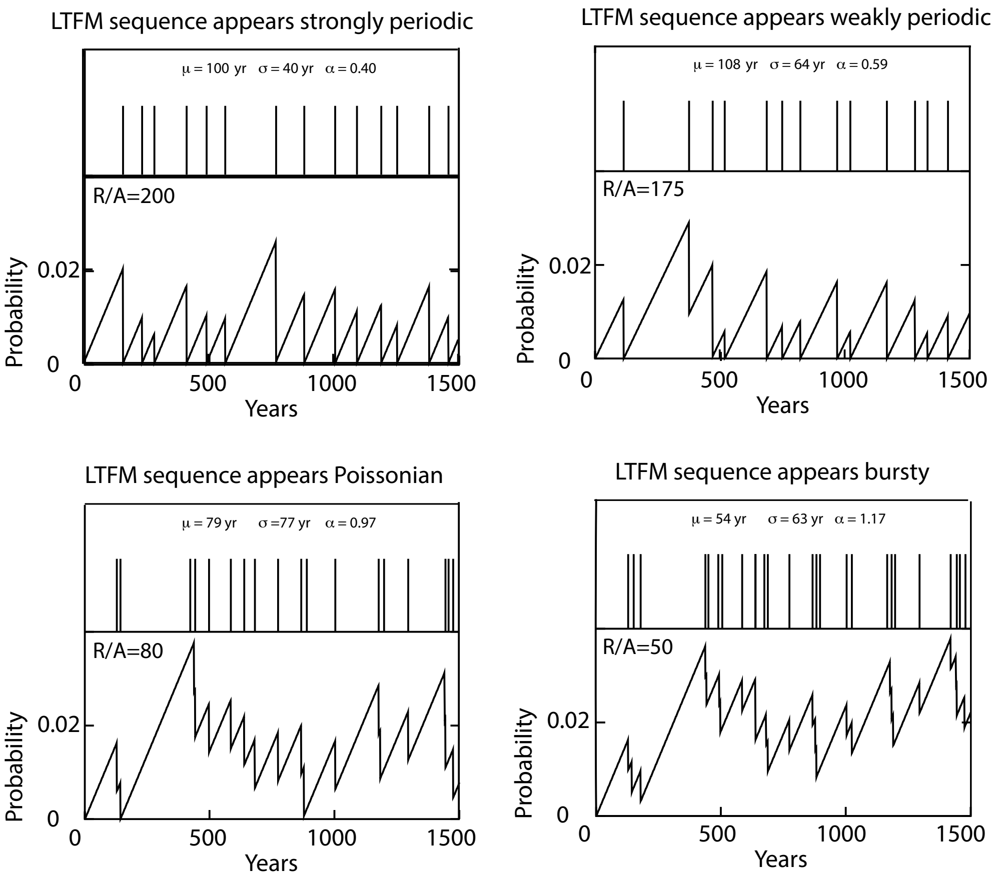

Figure 2.9. Sequences produced by the LTFM model can appear a) strongly periodic, b) weakly periodic, c) Poissonian, or d) bursty, depending on the model parameters. The four sequences shown have the same probability accumulation rate but different release parameters, so the aperiodicity increases as R/A decreases.

by some records (Figure 2.2). Hence if $P(t) > \delta_i$, the probability drops by R_i. Using two probability thresholds and probability drops to describe both rare larger and more frequent smaller strain changes allows LTFM to simulate the range of observed supercycle behavior (Figure 2.10). The higher threshold and probability drop simulate infrequent events that have slip and strain release much greater than typical events, and so end a supercycle. Using two similar thresholds simulates a Sumatra-style earthquake history with long gaps separating clusters, because earthquakes can occur only late in a supercycle. This case would correspond to a very strong fault. Conversely, a low threshold for smaller earthquakes and a much higher one for larger earthquakes simulates a Tohoku-style record where smaller earthquakes occur frequently between the largest events, so the supercycles in the strain record do not appear in the earthquake history as gaps and clusters. The threshold and drop parameters can be chosen to simulate the very long gaps associated with intraplate and plate boundary zone earthquakes. In such situations, because strain accumulates slowly relative to plate boundaries, the lower threshold is quite low. Hence this threshold can be used in most applications with higher strain rates (e.g., Figure 2.10b), since it would have essentially the same effect as a zero threshold (e.g., Figure 2.9).

2.6.1. Example

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

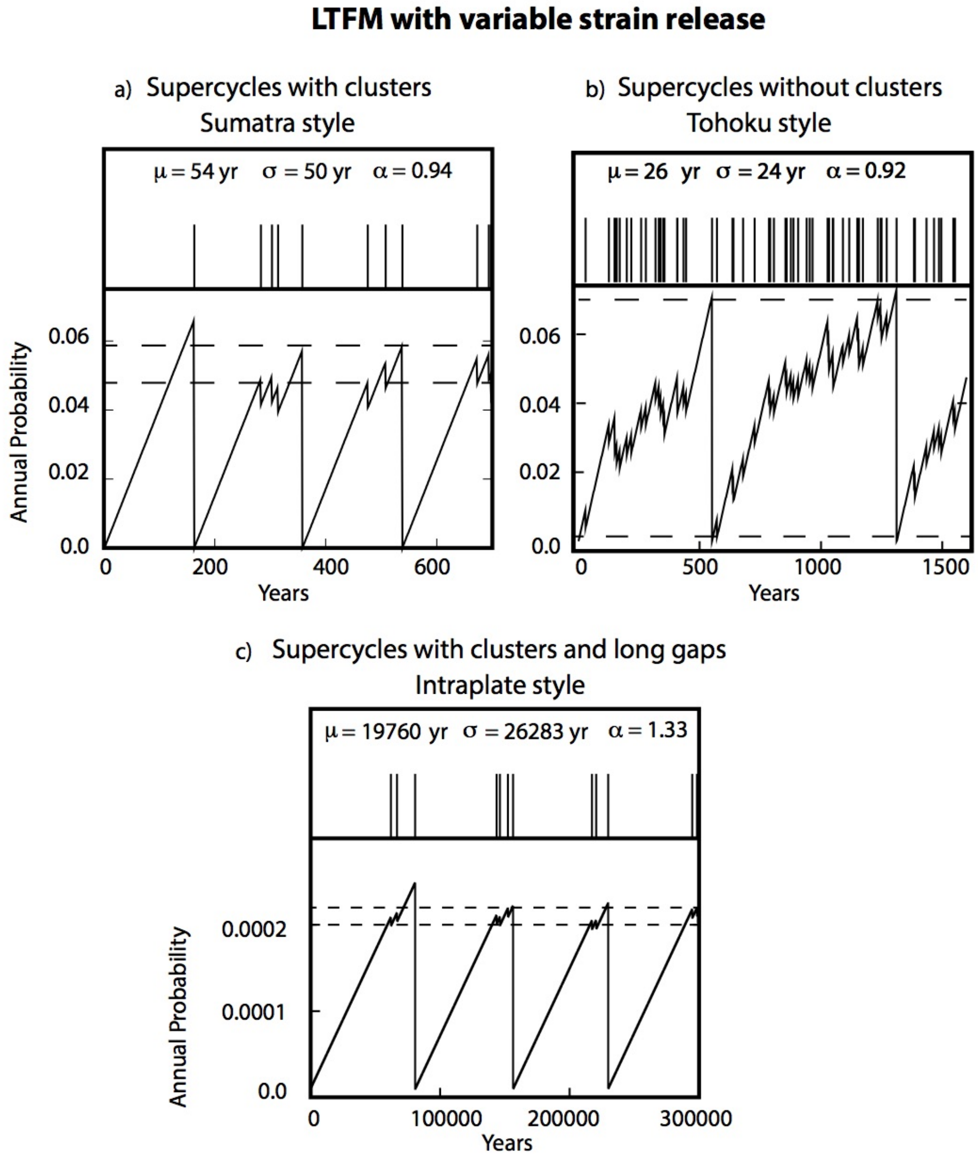

Figure 2.10. Using two probability thresholds (dashed lines) and probability drops to describe rare larger events and more frequent smaller events allows LTFM to simulate a wide range of observed supercycle behavior.

We then searched these regions for parameters giving a memory parameter close to that observed.

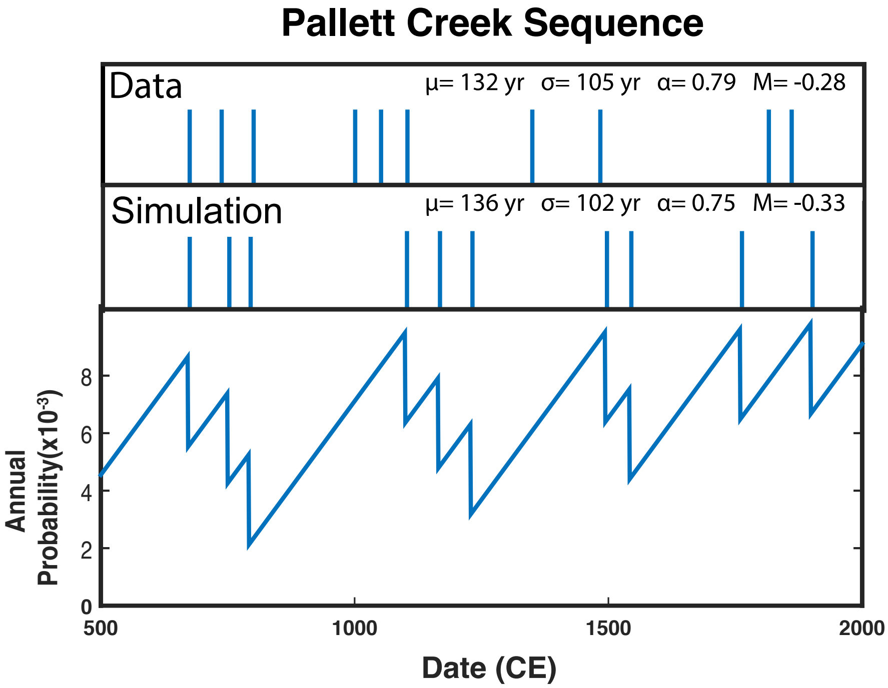

Figure 2.11. LTFM simulation for Pallett Creek, California. Top: Paleoseismic record (Sieh et al., 1989). Center and bottom: Simulation giving clustering similar to that observed. The event timing differs between the simulation and the observed record due to the model?s stochastic nature.

A simulation with accumulation rate $A = 2/289^2$ and release parameter R $= 130$A that gives clustering behavior similar to that observed is shown in Figure 2.11. The data have $\mu = 132yr$, $\sigma = 105yr$, $\alpha = 0.79$ and $M = -0.28$, and the simulation has $\mu = 136yr$, $\sigma = 102yr$, $\alpha = 0.75$ and $M = -0.33$, indicating weak periodicity. The event timing differs between the simulation and the observed record due to the model's stochastic nature. The longer intervals between earthquakes begin at times of low probability, consistent with the pattern noted in terms of cumulative strain by Weldon et al. (2004).

We used the LTFM model to explore the long-term variability of fault behavior by creating simulations much longer than paleoseismic records, and then sampling them for intervals corresponding to paleoseismic records. Figure 2.12 shows results for a 50,000 year long simulation using parameters appropriate for Pallet Creek. The mean and standard deviation of recurrence times averaged over a moving 1345-year window, corresponding to a paleoseismic record, are relatively stable over long time periods. This stability would be consistent with the idea of steady loading and unloading by plate motion and large earthquakes. However the mean and standard deviation of recurrence times vary somewhat. The aperiodicity shows that the simulated paleoseismic record sometimes appears strongly periodic (standard deviation small relative to the mean) implying a seismic cycle model, while at other times it looks weakly periodic, Poissonian (standard deviation similar to the mean), or bursty. This variability is illustrated by the earthquake history between model years 19,000 and 22,000. Hence the recurrence variability due to long-term fault memory can give rise to paleoseismic records that at different times appear to have different underlying statistical distributions. Thus a given paleoseismic or instrumental window may give a biased view of the long-term seismicity.

2.6.2. LTFM and intraplate earthquakes

Long-term fault memory may also be an important contributor to the space-time variability of continental intraplate earthquakes. Considerable recent attention (reviewed by Liu and Stein, 2016, Calais et al., 2016, and Stein et al., 2017a) has been directed to how and why earthquakes within continents behave differently in space and time from those on plate boundaries. Faults at plate boundaries are loaded at constant rates by relatively

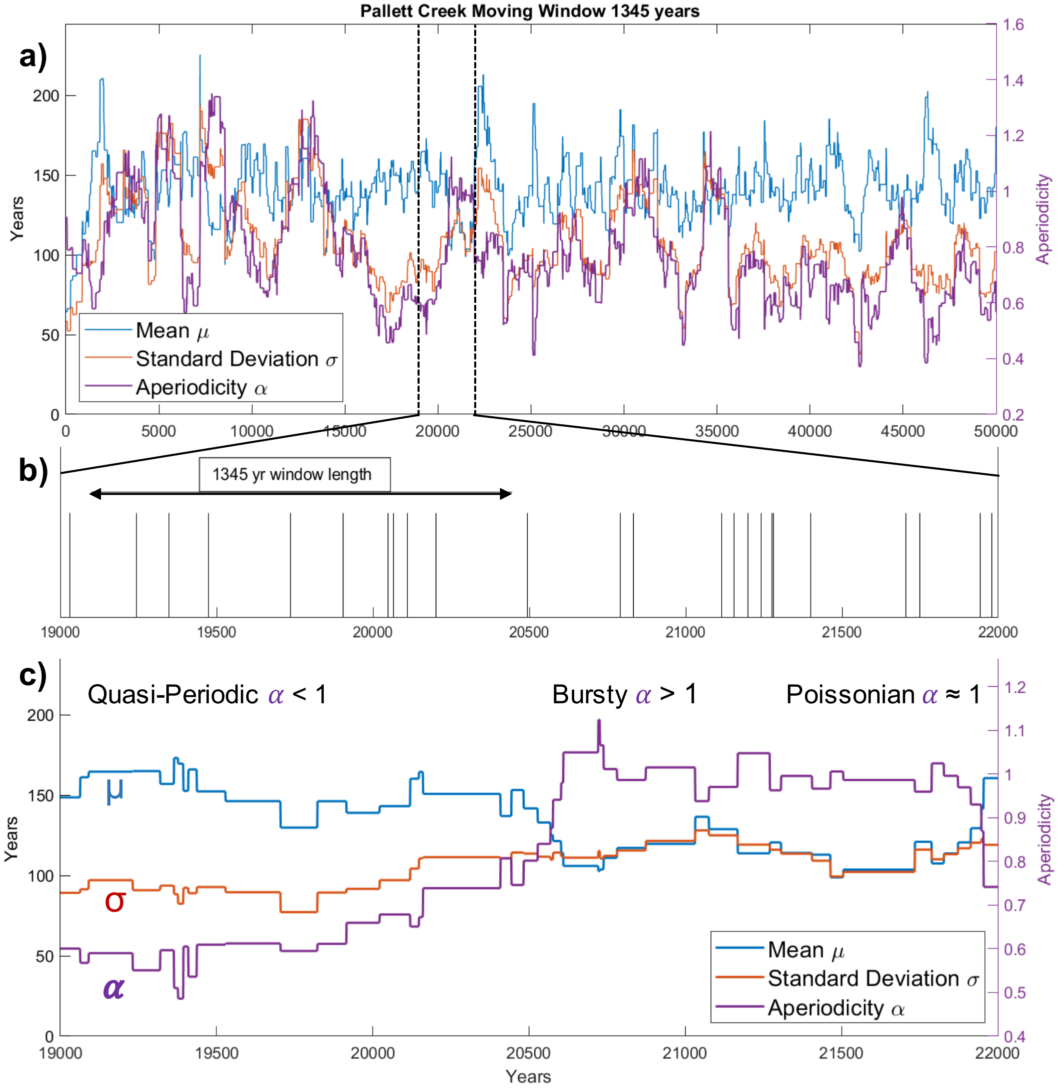

Figure 2.12. a) 50,000 year LTFM simulation using Pallett Creek parameters. The mean and standard deviation of recurrence times are averaged over a moving 1345-year window, corresponding to a paleoseismic record. b, c) 3,000 year section of simulation above between dashed lines in a). The aperiodicity shows that the simulated paleoseismic record sometimes appears strongly periodic ($\alpha < 0.5$) , while at other times it looks weakly periodic ($0.5 < \alpha < 1$), Poissonian ($\alpha = 1$), or bursty ($\alpha > 1$).

rapid and steady relative plate motion. Consequently, earthquakes concentrate along the plate boundary faults and show quasi-periodic (relative to intraplate earthquakes) occurrences, although the actual temporal patterns are often complicated. The spatial gaps that appear are filled in over time.

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

Topographic data from Australia, where erosion is very slow, provide some of the best evidence available of how intraplate faults slip over time. Figure 2.2f illustrates this pattern of clusters of activity separated by much longer and irregular intervals of quiescence. Liu and Stein (2016) note that the pattern of displacement accumulated over time is similar to the Devil's Staircase function, a fractal property of chaotic dynamic systems (Devaney et al., 1989; Turcotte, 1997). The apparent long-distance roaming of large mid-continental earthquakes also suggests dynamic system behavior. In such a system, change of any part of the system (such as rupture of a fault) could impact nonlinearly the behavior of the whole system.

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

A noticeable difference between the clustering in Australia and that on plate boundaries is that in Australia the gap durations are more than ten times as long as the clusters, whereas on plate boundaries the gaps are only 2-3 times as long as the clusters. As shown in Figure 2.9, LTFM can describe this effect via assuming the level of strain accumulation required for an earthquake. A proposed alternative is that clusters of large intraplate earthquakes reflect the fault weakening after the first major event, so as to permit repeated failure (Li et al., 2009). Models have been proposed for how weakening and subsequent healing might occur (Sibson, 1992; Lyakhovsky et al., 2001). In Lyakhovsky et al.'s model, as the rate ratio between loading and healing increases, behavior changes from regular to clustered. This is because healing tends to purge long-term memory. These models were developed with a view toward describing the evolution of fault properties over multiple earthquake cycles, i.e. a different type of long-term fault memory.

2.6.3. LTFM model discussion

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

In this spirit, we have used a simple model that simulates general properties of supercycles. We plan to explore its possible applicability to paleoseismic records in other areas and in different tectonic regimes. For example, clusters have been observed in paleoseismic data in plate boundary zones, where diffuse deformation is spread over multiple faults and long-term slip rates are slower than on primary plate boundary faults (which typically move at > 10 mm/yr) but higher than in continental interiors (which typically deform at < 1 mm/yr) (Wallace, 1987; Rockwell et al., 2000; Friedrich et al., 2003; Oskin et al., 2008; Dolan et al., 2016; Gold et al., 2017). Some clusters seem to arise on individual faults, whereas others involve groups of faults. The Wasatch fault and adjacent faults show a strain release and slip pattern similar to that in Australia (Figure 2.2e) (Wallace, 1987; Friedrich et al., 2003). In the Eastern California shear zone, regional strain release appears to occur via *"distinct periods or bursts of seismic activity punctuated by periods of relative quiescence. Individual faults, however, appear to behave in a quasiperiodic fashion, with the clustering produced by the in-phase earthquake generation of the system of faults"* (Rockwell et al., 2000). Hence LTFM may be involved in plate boundary zone faults, but fault interactions and changes in loading across the zone may also contribute.

Additional features could be added to the model without overcomplicating it. Its current form allows for two classes of earthquakes causing different probability decreases, or strain releases. In some cases, only one may be needed, as motivated by observations that slip in large events on individual fault segments appears similar (Schwartz and Coppersmith, 1984) and Weldon et al.'s (2004) observation that on the area of the San Andreas they studied *"there appears to be no relationship between strain level and the size of earthquakes."* However, Goldfinger et al. (2013) note a "weak tendency" for clusters to terminate with an "outsized" event, as found for the Tohoku and Sumatra records (Figure 2.2). Moreover, some of the strain release may occur via slow slip events (Rogers and Dragert, 2003; Jiang et al., 2017) that may not appear in the paleoseismic record.

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

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

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

2.6.4. Mathematics of the LTFM model

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

---

[1] Possible values of P(t) have the form

$$min\Big\{(\alpha A - \beta R_1 - \gamma(\lambda - 1)R_1)^+, 1\Big\}$$

where

$$\lambda = R_2/R_1$$
$$(x^+) = max\{x, 0\}$$

and α, β, γ taking non-negative integer values.

[2] If the probability at time s is $P(s) = C_s$ then the conditional probability of an earthquake at time t is equal to

$$min\Big\{(A + C_{t-1}) \times \chi_{pos}(A + C_{t-1} - \delta_1), 1\Big\}$$

with $\chi_{pos}(x)$ equal to 1 if $x > 0$ and equal to zero otherwise.

expected number of earthquakes in a distant time span of length T can be calculated, along with the approximate standard deviation.

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

2.7. Implications for hazard assessment

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

Current earthquake probability estimates depend on assuming a probability density function for the recurrence intervals with input parameters inferred from the available earthquake history. Figure 2.13 illustrates the resulting uncertainties for Cascadia. Figure 2.13a shows the effects of additional paleoseismic data. Goldfinger et al.'s (2012) chronology yielded a mean recurrence interval of 530 yrs and a standard deviation of 271 yrs for the entire 10,000 year record, and a mean recurrence interval of 326 yrs and a standard deviation of 88 yrs for the most recent cluster. Including a newly-identified event in a revised chronology (Goldfinger et al., 2017) has a small effect on the 10,000-year record's parameters, changing the mean recurrence interval to 502 yrs and a standard deviation of 239 yrs. However, adding this event makes all events in the past 5,000 years part of the same cluster, with a recurrence interval of 452 yrs and a standard deviation of 142 yrs.

Whether to assume that a recent cluster is continuing or has ended can lead to quite different estimates of earthquake probabilities (Stein et al., 2017b). Figure 2.13b shows the different distribution of recurrence intervals corresponding to the two different chronologies and various probability density functions with parameters corresponding to the two chronologies. By far the largest difference arises from assuming either that the recent cluster continues, or that the cluster is over so the appropriate parameters are those for the entire record. Assuming that we are still in the cluster predicts higher probability than using the entire record. This effect is more important than the specific probability density function assumed. The corresponding effect appears from considering the conditional probability of a large earthquake in the next 50 years, which results from integrating the probability density functions (Figure 2.13c).

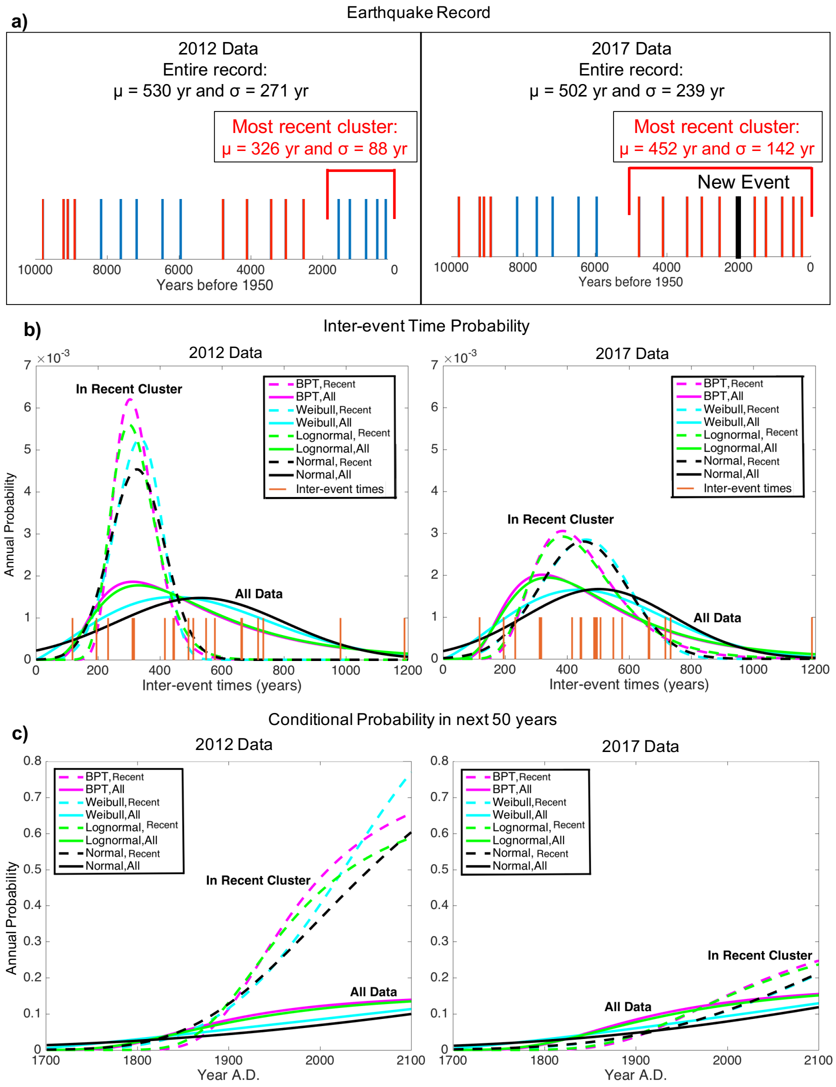

Figure 2.13. Illustration of earthquake probability issues for Cascadia due to a) differing paleoseismic records of Goldfinger et al., 2012 and Goldfinger et al., 2017, with its newly discovered event. Alternating red and blue events highlight the different clusters individual events are assigned to. b) Various probability density functions for inter-event times with parameters derived from the two chronologies in a). Orange sticks show the actual inter-event times in the corresponding records. Dashed lines use parameters of just the most recent cluster, corresponding to the assumption that the system is still in the recent cluster. Solid lines use the parameters of the entire record, corresponding to the assumption that the recent cluster has ended. c) Various conditional probabilities of an earthquake occurring in the next 50 years, using the same line designations in b). The largest difference in b) and c) arises from the recent cluster assumption, not in the specific density function assumed.

More generally, if faults have long-term memory, then individual earthquake cycles, and hence the recurrence times between successive large earthquakes, are not independent. Hence the renewal approach of modeling their probability as a function of time since the previous large earthquake can give misleading results. The problem is not that a renewal model is inappropriate, but rather that the renewal depends on release of nearly all accumulated strain, and that may occur at different times than large earthquakes. As shown in Figure 2.12, the recurrence variability due to long-term fault memory can cause short earthquake records to give a biased view of the long-term seismicity. As a result, further investigation of long-term earthquake recurrence variability is important both for understanding the nature and causes of supercycles and for improving hazard assessment.

CHAPTER 3

1952 Kern County Earthquake

3.1. Summary

Seismic intensity data based on first-hand accounts of shaking give valuable insight into historical and early instrumental earthquakes. Comparing an observed intensity distribution to intensity-prediction models based on modern calibration events allows the magnitude to be estimated for many historic earthquakes. Magnitude estimates can also potentially be refined for earthquakes for which limited instrumental data are available. However, the complicated nature of macroseismic data and the methods used to collect and interpret the data introduce significant uncertainties. In this paper, we illustrate these challenges and possible solutions using the 1952 Kern County, California, earthquake as a case study. Published estimates of its magnitude vary from M_W $7.2-7.5$, making it possibly the second largest in California during the 20th century. We considered over 1,100 first-hand reports of shaking, supplemented with other data, and inferred the magnitude in several ways using intensity prediction equations, yielding a preferred intensity magnitude M_I 7.2 ± 0.2, where the uncertainty reflects our judgement. The revised intensity distribution reveals stronger shaking on the hanging wall, south of the surface expression of the White Wolf fault, than on the footwall. Characterizing the magnitude and shaking distribution of this early instrumental earthquake can help improve estimation of the

seismic hazard of the region. Such reinterpreted intensities for historic earthquakes, combined with USGS Did You Feel It? data for more recent events, can be used to produce a uniform shaking dataset with which earthquake hazard map performance can be assessed.

3.2. Introduction

In this paper, we explore issues associated with analysis of seismic intensity data using the 21 July, 1952 Kern County, California, earthquake as a case study. Seismic intensity characterizes the level of earthquake shaking by human perceptions and effects on made-man structures and objects within them. The Modified Mercalli Intensity (MMI) scale introduced by Wood and Neumann (1931), based on earlier scales widely used previously in Europe (see Musson et al., 2010), has long been preferred in the U.S. by engineers and seismologists. Reports of earthquake shaking have been collected and assigned intensity levels by U.S. government agencies for over a century. From 1924 onward, postcard questionnaires were collected by the U.S. Coast and Geodetic Survey (USCGS), who left stacks of questionnaires with postmasters and employees of large companies to be completed by themselves and others in the community shortly after earthquakes (Byerly and Dyk, 1935). The questionnaires asked whether the respondents and those around them felt the earthquake, whether it frightened or awoke them, whether small objects on shelves shifted or overturned, whether large furniture was shifted, and whether and how buildings were damaged or destroyed. These questions correspond directly to indicators used in the MMI scale. Later, this information was collected via randomized telephone surveys and mailed postcards (Dewey et al., 1995). Presently, felt reports are collected online through the USGS Did You Feel It? (DYFI) website, making them easier to collect and interpret (Wald et al., 1999a). The DYFI system assigns numerical Community Internet Intensity (CII) values using a modification of the algorithm proposed by Dengler and Dewey (1998) to determine Community Decimal Intensities (CDI) from questionnaire responses. DYFI intensity values characterize representative shaking in an area, and hence differ systematically from conventionally interpreted MMI data, which tend to be based on the most dramatic effects (Hough, 2013).

Uncertainties in intensity values arise from several sources. Human perceptions of and reactions to shaking vary with a person's physical sensitivity and disposition. Building age and construction quality, on which information is often unknown or unspecified in first-hand reports, influence the level of damage and perceived or felt shaking. Methods of intensity data collection as well as the distribution of population relative to the epicenter can introduce sampling biases. How questionnaires are phrased can impact the responses. Information about the location of a damage report varies, sometimes giving only a city name or approximate location. These issues introduce uncertainties into the intensity values assigned from the observations. Further uncertainties arise when intensity-prediction models describing the variation of intensity with distance from the fault are used to estimate the earthquake's magnitude.

Despite these limitations, intensity values are important for earthquake hazard assessment because they directly represent damage and because the length of macroseismic records often far exceeds that of the instrumental catalog (e.g., Ambraseys, 1983). For early instrumental and even recent earthquakes, intensity distributions can reveal the shaking distribution in far greater spatial detail than is possible using instrumental data. For historical earthquakes, macroseismic information is often the primary data available.

Revisiting historic intensity data is worthwhile, because historically assigned intensities are often inflated by 0.5 - 2.0 intensity units relative to modern practice (Ambraseys 1971, 1983; Hough and Page, 2011; Hough 2013, 2014). Even following modern practice, conventional intensity assignments are inherently subjective, leading to differences of 0.5 - 1.0 intensity units for the same report by different interpreters (Hough and Page, 2011).

The Kern County earthquake was one of the largest in California during the 20th century. Some of the issues discussed in our study were noted at the time. Reports from the town of Tehachapi illustrate how historic or traditionally assigned intensities tend to be inflated. In 1952 its main street was lined with old, weakly cemented, unreinforced brick buildings that were highly susceptible to damage from shaking (Steinbrugge and Moran, 1954). Tehachapi (Figure 4.1) is close to the end of the surface rupture furthest from the epicenter and therefore in the direction of rupture which focused energy towards the town (Richter, 1958). These effects likely account for the fact that many buildings downtown were damaged, causing most of the 12 fatalities that resulted from this earthquake, and extensive damage occurred to the Southern Pacific Railroad due to tunnel collapse and track warping. The shaking and damage made Tehachapi the focus of media coverage. Richter (1958) notes, “These deaths also focused the attention of the press, the public, and even of officials who should have been better informed, on the losses at Tehachapi disproportionately to those elsewhere,” hinting at media’s tendency to report the most dramatic stories rather than more representative effects. Richter also notes, “Many structures on the main street [of Tehachapi] were of this character [outdated masonry]; they showed such damage as caused incautious observers to rate the local intensity at VIII, though it is doubtful whether more than VII was actually indicated.” In his seminal

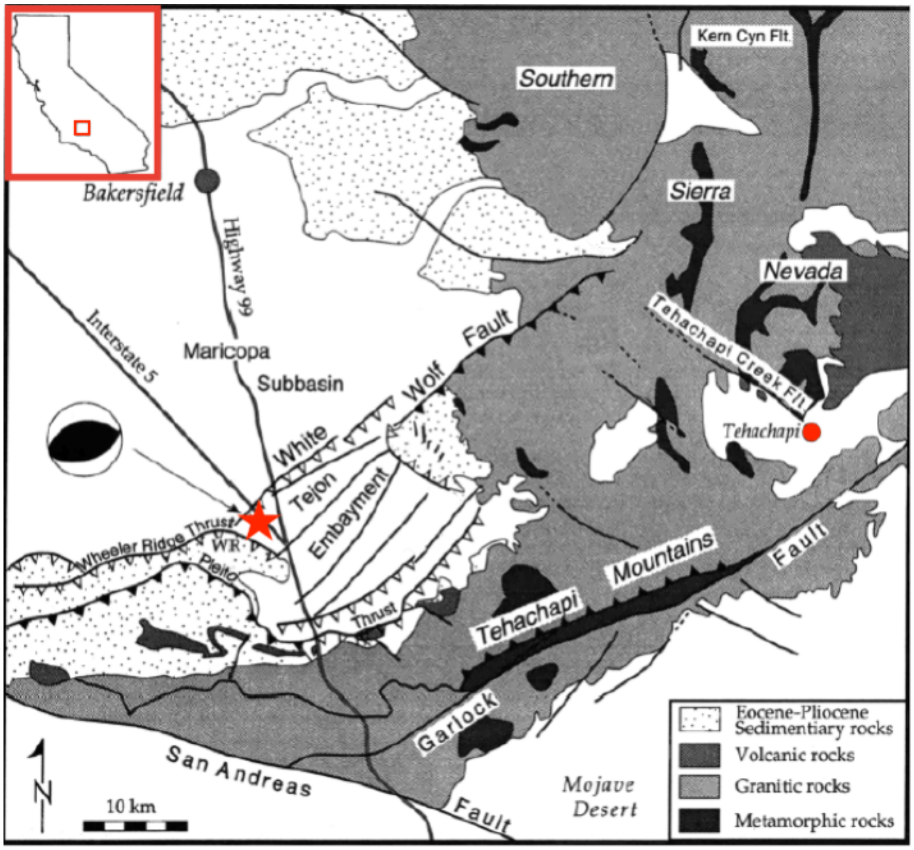

Figure 3.1. Geology of the epicentral area. Open barbs represent blind thrusts, filled barbs represent exposed thrusts. Focal mechanism is 1952 mainshock, star indicates the epicenter. (After Bawden, 1997).

volume, Richter (1958) realized that intensity data for this earthquake were susceptible to inflation and would require careful interpretation keeping these influences in mind. In this spirit, we reinterpret intensity reports for this earthquake.

We refine the magnitude estimate and explore the associated uncertainties by reinterpreting over 1,100 macroseismic intensity reports. These reports are cataloged in U.S. Coast and Geodetic Survey Abstracts of Earthquake Reports for the Pacific Coast and the Western Mountain Region, hereafter referred to as 'the primary dataset.' This refinement provides a new characterization of the shaking distribution of this early instrumental earthquake and a new estimate of its magnitude via recent intensity-prediction models.

3.3. The 1952 Kern County Earthquake

On July 21, 1952 at 4:21 AM local time (GMT - 7 hours) the White Wolf fault (WWF) ruptured in Kern County, California (Figure 1). This earthquake illustrates that large earthquakes on secondary faults trending obliquely to the San Andreas fault, such as the WWF and the nearby and similarly-oriented Garlock fault, are important contributors to seismic hazard in California (Walls et al., 1998). The WWF is a steeply-dipping (up to 75) reverse fault located east of the restraining bend in the southern portion of the San Andreas Fault (Bawden, 1997, 2001). It is thought to be slow-slipping, with long-term slip rate estimated at < 10 mm/yr, with variability throughout the Quaternary (Stein and Thatcher, 1981; Hearn et al., 2013). The Working Group for California Earthquake Probabilities (WGCEP, 2008) estimated a long-term slip rate of $0.23 - 1.55$ mm/yr for different deformation models. Cutting the southern part of the San Joaquin Valley into the northern Maricopa sub-basin and the southern Tejon Embayment, the WWF is blind from the epicenter of the 1952 mainshock at its southwest end to approximately the middle of the fault (Figure 1). From there, the 1952 earthquake ruptured the surface to its northeastern endpoint, about 60 km from the epicenter. Bawden (2001) used geodetic data to develop a two-segment, right-stepping fault model with nearly uniform reverse slip of about 1.6 - 1.9 meters and additionally up to 3.6 meters of left-lateral slip in the epicentral patch. Stein and Thatcher (1981) developed a similar rupture model, with three fault segments whose dip gradually decreases away from the epicenter.

It has been hypothesized that this earthquake was triggered or induced by stress changes associated with oil production. The epicenter was within a few kilometers of an extraction well, shortly after production started from a deep (3 km) horizon (Hough et al., 2017). Although it is impossible to prove that the earthquake was induced by anthropogenic activity, it may have been one of the largest induced earthquakes (Foulger et al., 2017).

The Kern County earthquake was the most widely felt earthquake in California over the 50 years following the 1906 San Francisco earthquake. Although the 1952 earthquake was studied in detail at the time (Steinbrugge and Moran, 1954; Oakeshott, 1955; USCGS, 1966), as well as by later studies, estimates of its magnitude vary. The current USGS estimate is $M_W 7.5$ (Hutton et al., 2010). Instrumentally derived estimates include $MS7.7$ (Gutenberg and Richter, 1954; Richter, 1958), $M_W 7.3/M_S 7.2$ (Ben-Menahem, 1977), $M_L 7.2$ (Kanamori and Jennings, 1978), and $M_S 7.8/m_B 7.3$ (Abe, 1981). Bawden (2001) geodetically determined $M_W 7.2$. Bakun (2006) used 647 reinterpreted historic intensity assignments to estimate an intensity magnitude, a magnitude derived from intensity data, of $M_I 7.3$.

3.4. Reinterpretation of Historic Macroseismic Intensity Data

Over 1,100 reports of shaking were compiled by the USCGS and assigned MMI intensities (USDOC, 1966). We compared the originally assigned intensity values with our reinterpretations of the same data. We augmented our reinterpreted intensity assignments with intensities derived from 12 strong motion data points, 67 retroactively-reported DYFI data points, 12 data points for damaged electrical transformers, and observations of precariously balanced rocks. Adding these data makes the reinterpreted dataset more robust, because many of the original intensity assignments close to the epicenter were based on geologic indicators such as rock slides, secondary ground deformation, and groundwater perturbations, which are now recognized to be unreliable indicators of shaking intensity (Ambraseys, 1983). About 7 % of the primary dataset are based on such geologic indicators, which disproportionately controlled the highest intensity assignments in the original assessment. These descriptions were not considered in our reinterpretations. The primary dataset contains 1,137 reports with historically assigned intensities that we refer to as original intensities. Our reinterpreted dataset (primary dataset minus reports describing geologic/industrial effects plus additional data) contains 1,144 reports, hereafter referred to as reinterpreted intensities.

Latitudes and longitudes for locations in the primary dataset, which mainly gives city names, come from the NOAA Earthquake Intensity Database. For locations where more detail is provided, for instance a street address or named electrical substation, latitude and longitude were found using Google Earth after confirming that street names and locations remained unchanged since 1952 by using USGS historical topographic maps (http://historicalmaps.arcgis.com/usgs/).

To explore the subjectivity of intensity assignments, we compared reinterpretations of the primary dataset made by two different analysts. This allows an evaluation of inter-rater variability and reliability. We compared intensity assignments only when both interpreters provided an assignment, i.e. excluding reports for which either interpreter was unable to assign an intensity based on the information given. 34 % of assignments were identical between interpreters, 70 % were within ± 0.5 MMI unit, and 97 % were within ± 1 MMI unit. The absolute value of the mean difference between the interpreter's intensity assignments was 0.2 MMI units.

We also considered other sources of information that can constrain ground motions, including electrical transformer damage caused by the earthquake. The toppling of transformers has been investigated by earthquake engineers and can be related to Peak Ground Acceleration (PGA) via fragility curves, which can then be used to estimate MMI (Huo and Hwang, 1995). Detailed descriptions of transformer damage (Peers, 1955; USCGS, 1966) indicate how many transformers were toppled at a given substation. To estimate PGA, we compared the estimated percentage of overturned transformers to fragility data presented by Huo and Hwang (1995). Similarly, the preservation or overturn of precariously balanced rocks has been used to infer PGA and intensity at a number of locations in the near field (Brune et al., 2004).

We also considered peak ground acceleration data from 12 strong motion recordings of the earthquake (Murphy and Cloud, 1954). These data were recorded by instruments that operated in trigger mode, with no pre-event memory. For a large earthquake, these instruments generally triggered on the P wave and captured the full S wave. We converted PGA data to instrumental MMI using the relationship determined by Worden et al. (2012). Instrumental PGAs range from 1 to 25 % g, corresponding to instrumental intensity values 3.3 to 7.2.

Table 3.1 gives examples to illustrate differences between original to reinterpreted MMI assignments. Our reinterpretation of felt reports differed from the original in the following ways:

(1) Taking the Ambraseys (1983) conservative approach to secondary geologic indicators.

(2) Using quartile decimal (e.g. 5.0, 5.25, 5.5, etc.), rather than integer values of MMI.

(3) Weighting our inferred intensities to give more weight to the disturbance of objects than to subjective human reactions. Originally, the disturbance of objects and personal reactions were given about equal weight for the lower MMI intensities (USDOC, 1966).

(4) Assigning MMI 5 only when accounts describe toppling of small objects, a key objective indicator for this intensity level (Richter, 1958).

(5) Rather than characterizing weakly felt intensities as a range of MMI 1-3, as was the earlier practice, we differentiated between MMI 2 and 3: MMI 2 for reports of "felt by few" without objective indicators, and MMI 3 for reports of "felt by many" accompanied by reports of hanging objects having swung.

(6) Assigning MMI 1 for sites "Reported not felt", following Ambraseys' (1983) recommendation.

The reinterpreted intensity values are generally although not uniformly lower than originally inferred values (Figures 4.2, 4.3). The absolute value of the mean difference between original and reinterpreted intensities is 0.6 MMI units. Some of this difference

Table 3.1. Example reports from primary dataset, comparing original and reinterpreted MMI. Groundwater and secondary geologic indicators were not considered in the reinterpretations.

Location	Report (slightly edited for space)	Orig. MMI	Reinterp. MMI
Pleasant Grove, CA	Felt by several in community. One reported electric cord swung.	1-3	3
Carmel Valley, CA	Motion slow, lasted 30 seconds. Awakened many, frightened few, felt by some outdoors. Rattled windows, doors. Hanging objects swung	5	4
Shell Beach, CA	The press reported a large bone fell off museum shelf.	6	6
Woody, CA	Motion rapid, rolling, lasted l minute. Felt by and awakened all in community, frightened many. Rattled windows, doors, dishes. Hanging objects swung. Trees, bushes shaken slightly. Shifted small objects. Overturned vases and small objects. Knickknacks, books, pictures fell. Broke dishes and vases. Water supply and springs milky.	7	5.5
Miracle Hot Springs, CA	Motion rapid, lasted 14 seconds. Felt by, awakened, and frightened all. Rattled windows, doors, and dishes; house creaked. Hanging objects swung N. Trees, bushes shaken strongly. Shifted small objects and furnishings; overturned vases, etc., small objects and furniture. Miracle Hot Springs went dry during the first shock, but resumed flow during the later shocks. Temperature of water seemed hotter. Many rock slides in canyon. Democrat Springs went completely dry.	8	6

results from our assignment of quarter units rather than integral units. Reinterpretations are consistent with the retroactively-reported DYFI, instrumental, transformer and precariously balanced rock data (Figure 4.3). Figures 4.2, 4.3, and 4.4 show locations of individual data points, with the remainder of the map interpolated through the ?surface? function of the Generic Mapping Tools, which grids data using adjustable tension continuous curvature splines (Wessel and Smith, 1991).

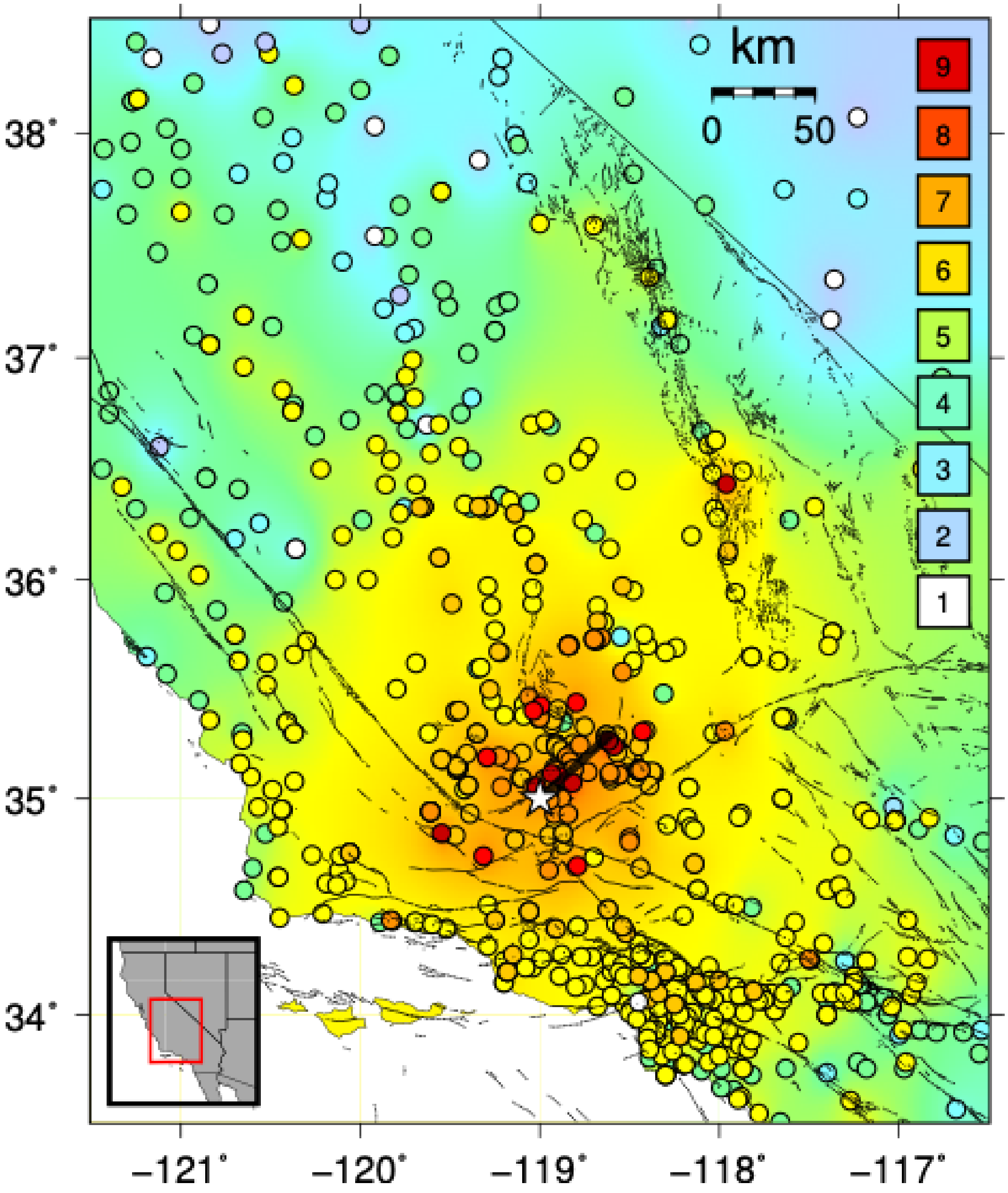

Figure 3.2. Map of original intensity assignments. Circles show individual assignments, and the map is interpolated via continuous curvature splines from those points. Color scheme follows USGS Shakemap convention (Wald et al., 1999b). Thin lines are faults, thick line is 1952 WWF rupture. Star is epicenter.

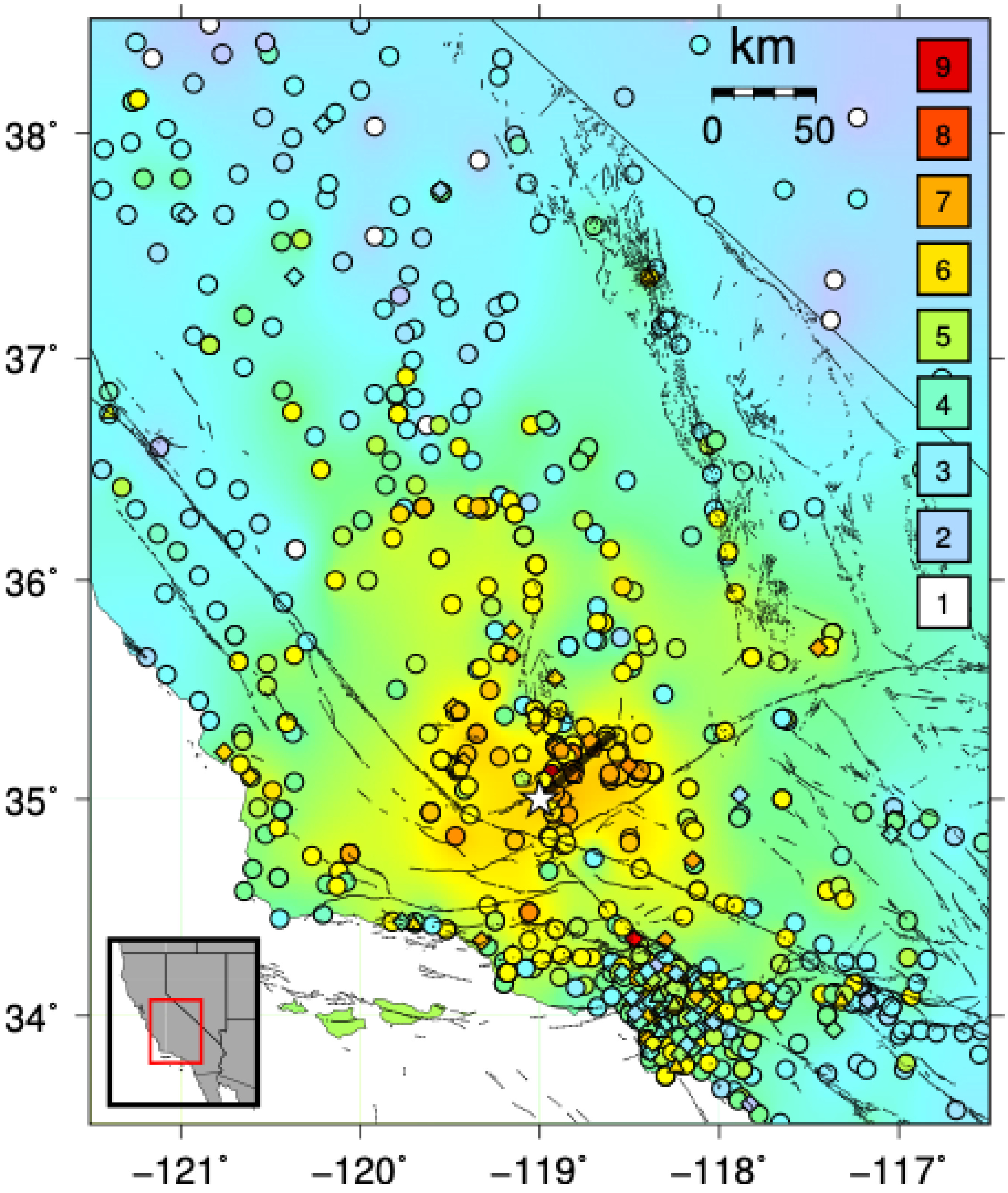

Figure 3.3. Map of intensity assignments as reinterpreted in this study (circles). Also included are intensities derived from strong motion data (triangles), transformer/precariously balanced rocks (pentagons), and retroactively-reported DYFI (diamonds) ? all consistent with reinterpreted values. Shapes show individual intensities, and the map is interpolated via continuous curvature splines from those points. Color scheme follows USGS Shakemap convention (Wald et al., 1999b). Thin lines are faults, thick line is 1952 WWF rupture. Star is epicenter.

Based on their analysis of precariously balanced rock and toppled transformer observations, Brune et al. (2004) concluded that accelerations were higher on the hanging wall (southern side) of the WWF, in particular along the northeastern half of the rupture, where the rupture reached the surface (Figure 4.1). They further concluded that along the southwestern half of the rupture, which they assume to have been blind, more significant energy was transferred to the footwall. Our reinterpreted intensity distribution also suggests that shaking was lower on the footwall and that intensities were lower along the southwest part of the foot wall as well, although this part of the intensity field is not well constrained (Figure 4.4).

To further explore the distribution of shaking, we consider the distribution of residuals relative to an appropriate baseline. As discussed at length in the following section, it is not clear what the optimal baseline is for this event. We use the intensity prediction equation from Atkinson and Wald (2007), hereafter AW07, assuming intensity magnitude MI 7.2, which is found by trial-and-error to provide a good average fit to the intensities (Figure 4.5). Because of uncertainty about the appropriate baseline, the residuals illuminate differences in relative shaking severity, but not necessarily absolute amplification/deamplification. Figure 5 reveals several first-order effects: 1) intensities are modestly elevated at some but not all sediment sites, including some sites in the Los Angeles and San Bernardino basins and in the Central Valley; 2) intensities are modestly elevated along a swath trending northwest through the Coast Ranges; 3) near field intensities are systematically low, in particular to the north of the fault trace. The amplification of shaking at sediment sites is an expected result (e.g., Borcherdt, 1970). Elevated shaking in the Coast Ranges is more enigmatic; we cannot propose an obvious explanation for

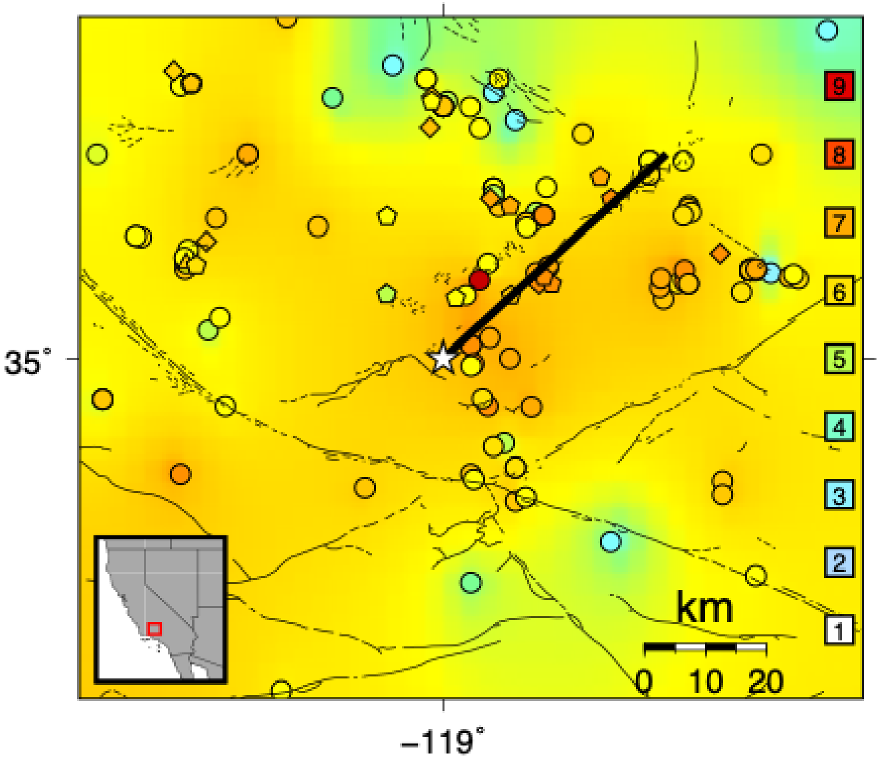

Figure 3.4. Map of near-field intensities as reinterpreted in this study (circles). Also included are intensities derived from strong motion data (triangles), transformer/precariously balanced rocks (pentagons), and retroactively-reported DYFI (diamonds). Shapes show individual intensities, and the map is interpolated via continuous curvature splines from those points. Color scheme follows USGS Shakemap convention (Wald et al., 1999b). Thin lines are faults, thick line is 1952 WWF rupture. Star is epicenter. Higher intensities are observed on the hanging (southern) wall of the fault.

this result, although it is possible that it reflects relatively low attenuation along paths to the west/southwest of the epicenter.

In the near-field, Figure 4.5 suggests deamplification of shaking, by -0.5 to as much as -1.5 intensity units. We suggest that this result reflects two effects. First, as discussed above, and proposed by Brune et al. (2004), shaking was lower on the footwall than

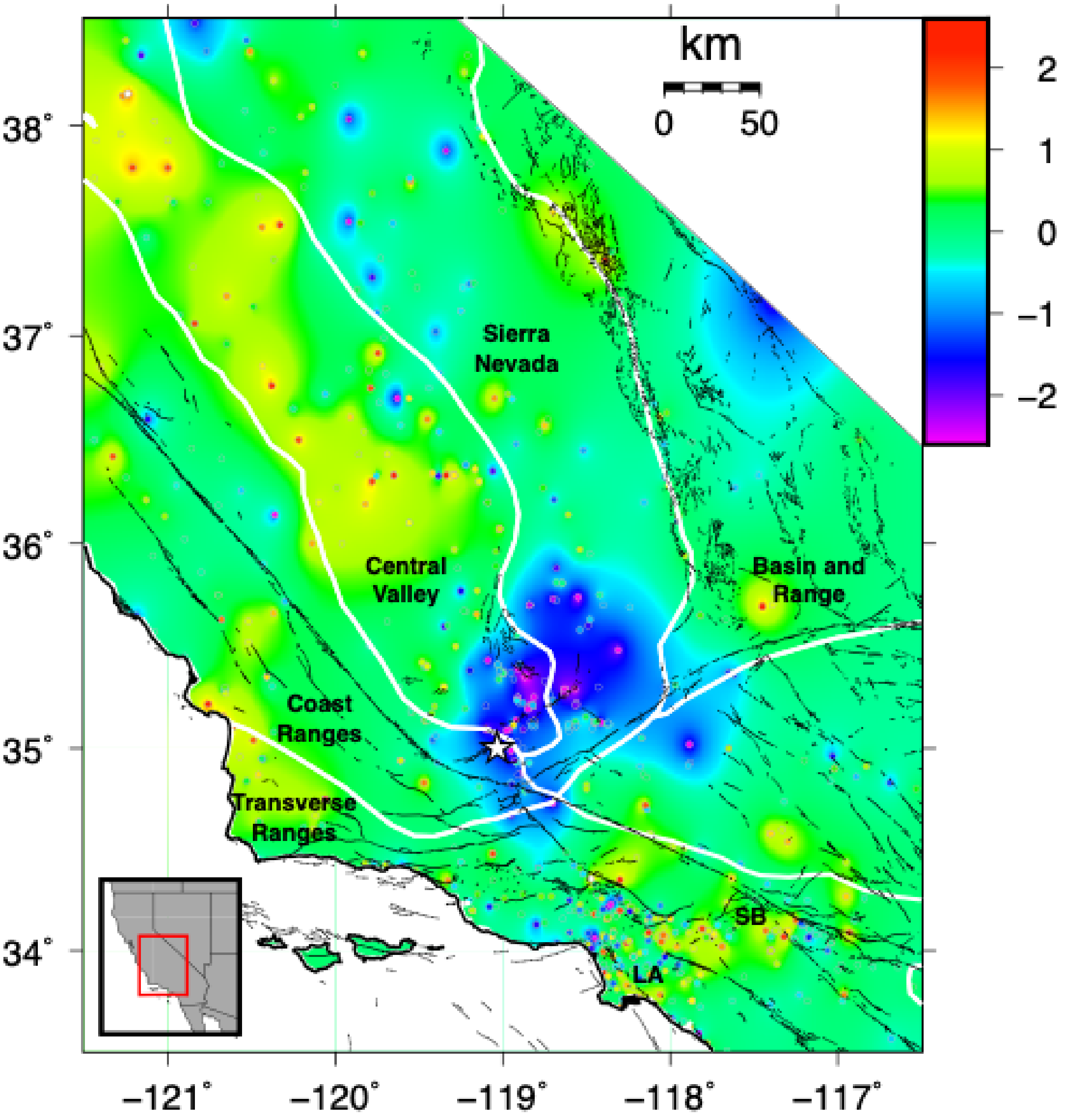

Figure 3.5. Map of intensity residuals, relative to the AW07 IPE assuming an intensity magnitude of 7.2. Circles show individual points and the map is interpolated from those points. Residuals from Nevada are masked out due to sparse data available to constrain the interpolation there. Thin black lines are faults, white lines are major physiographic boundaries, star is epicenter of 1952 earthquake. SB is San Bernardino Basin, LA is Los Angeles Basin.

the hanging wall, all the more so because we have not assigned intensities to locations on the hanging wall where only secondary geologic effects were reported. Secondly, we suggest that near-field shaking was deamplified by a pervasively non-linear response of near-surface sediments. The zone of estimated deamplification is similar to that inferred in the Kathmandu Valley during the 2015 Gorkha, Nepal, earthquake (Adhikari et al., 2017). Our conclusion is also consistent with the qualitative conclusion of Trifunac (2003), who showed that near-field damage during the 1933 Long Beach, California, earthquake was deamplified in some near-field regions by pervasively non-linear response of soft, water-saturated sediments.

3.5. Intensity Prediction Equations

Intensity Prediction Equations (IPE) predict the decrease of shaking intensity with distance for an earthquake of a given magnitude. To estimate the magnitude of the Kern County earthquake, we compare published IPEs to the intensity values. These equations are empirically determined, so uncertainties in the data used to derive them cause uncertainty in the IPEs.

We first use an IPE from Bakun (2006), hereafter B06, who built upon the IPE of Bakun and Wentworth (1997) by developing coefficients for historic earthquakes in southern California based on traditionally assigned intensity data for 20th century MW 5.0 - 7.1 earthquakes. **B06**:

$$MMI_T(M_I, D) = C_0 + C_1 M_1 - C_2 D - C_3 log(D)$$

where MMI_T is a traditionally assigned intensity, M_I is intensity magnitude, and D is distance from the epicenter in km. For southern California: $C_0 = 1.64 \pm 0.91$, $C_1 = 1.41 \pm 0.11$, $C_2 = 0.00526 \pm 0.00158$, $C_3 = 2.63 \pm 0.36$. The Kern County earthquake was not used to derive this relationship because it occurred prior to 1960 and the onset of modern and consistently calibrated instrumentation. Bakun (2006) used 647 reports from the 1952 Kern County earthquake and estimated its intensity magnitude as M_I 7.3.

We also used an IPE from Atkinson and Wald (2007). This relation is derived from modern DYFI data from California, so it might provide a better fit to our reinterpreted intensities. However, it may also not be fully appropriate for our data. Although conservatively reinterpreted, our data are still traditional intensity values determined from first-hand accounts. Hence they differ fundamentally from DYFI data in which intensities assigned from individual questionnaires are averaged within ZIP codes or geocoded cells. The **AW07** relationship is:

$$MMI_{DYFI}(M,R) = d_1 + d_2(M-6) + d_3(M-6)^2 - d_4 log(R) - d_5 R + d_6 B - d_7 M log(R)$$

where M is moment magnitude; $R = (D^2 + h^2)^{1/2}$; D is distance from fault in km; h is effective depth, equal to 14 km for California; $B = 0$ for $R \leq R_t$ and $B = log(R/R_t)$ for $R > R_t$; R_t is the transition distance in the attenuation shape, $R_t = 30.0$ for California. For California $d_1 = 12.27 \pm 0.24$, $d_2 = 2.270$, $d_3 = 0.1304$, $d_4 = 1.30$, $d_5 = 0.0007070$, $d_6 = 1.95$, $d_7 = 0.577$.

B06 predicts higher intensities than AW07 for a given magnitude at distances less than approximately 400 km, and lower intensities at distances more than 400 km (Figures

4.6,4.7) (Hough, 2013). Hough (2014) concluded that these differences likely reflect systematic bias in traditionally assigned intensities relative to DYFI intensities. Both IPEs predict negative intensities at distances greater than 400 km from the source for the lowest magnitude (M 6) tested in this study. The distance to a negative predicted intensity increases with increasing magnitude. For AW07, negative intensities occur over 200 km further from the source than B06 for the same magnitude. Following Hough (2014), sites at which B06 and AW07 predicted intensity values below 1.0 were set to 1.0, analogous to the assignment of MMI 1 for accounts of "Reported not Felt." Hough (2014) found B06 was a better fit (lower Root Mean Square misfit) than AW07 for historically assigned intensities. This is not surprising given that B06 was derived using traditionally assigned intensity data rather than the DYFI data to which AW07 was fit.

AW07 and B06 assume different earthquake source geometries and thus use different definitions for distance. AW07 defines D as distance from the closest point on the fault, whereas B06 assumes a point-source and calculates D as distance from the epicenter. For distance from closest point on fault calculations, we used a linear approximation of the WWF (Figures 4.2 - 4.4) from the epicenter ($35°$, $-119°$) (for uncertainties, see discussion in the electronic supplement to (Hough et al., 2017)), to the approximate endpoint of surface rupture ($35.3°$, $-118.6°$). The B06 point-source assumption is less appropriate for large earthquakes, such as this one with its 60 km long fault, so we also considered a point-source at the approximate midpoint of the surface rupture. We prefer to compute B06 distances from the midpoint of surface rupture rather than the epicenter because of the unilateral rupture propagation from the epicenter toward the endpoint. Considering the midpoint of surface rupture shifts the point-source toward the region of the fault where

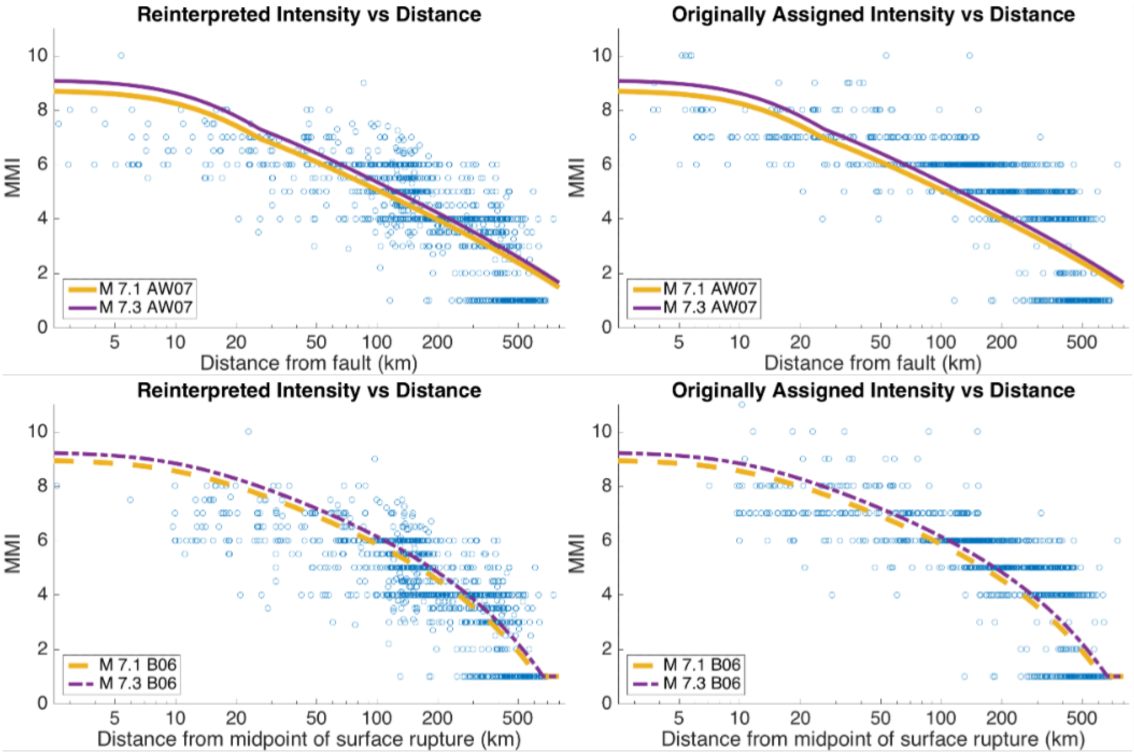

Figure 3.6. Individual intensity assginments vs. log10 distance from (Top) nearest point on fault and (Bottom) approximate midpoint of surface rupture. (Left) Reinterpreted intensity assignments. (Right) Original intensity assignments. AW07 (solid lines) and B06 (dashed lines) predictions are shown for M 7.1 (thick lines) and M 7.3 (thin lines).

radiated waves are stronger due to directivity and thus where intensities are predicted to be highest. The fault's endpoint and midpoint of surface rupture were estimated using Google Earth and the WWF trace in the USGS Quaternary fault and fold database. The effects of these different definitions for distance are explored below.

3.6. Fitting Intensity Data

To estimate the magnitude of the earthquake from the intensity data, we computed the Root Mean Square (RMS) misfit between intensity assignments, either original or

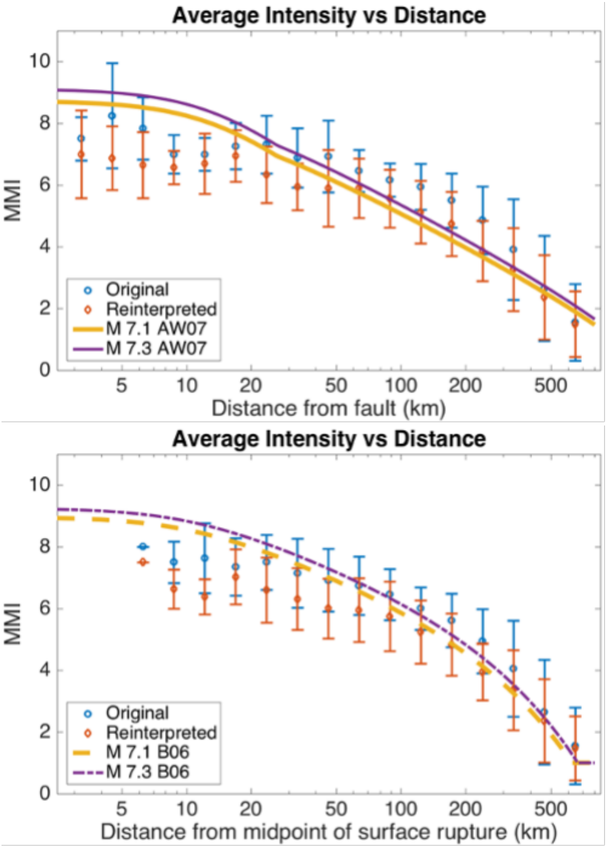

Figure 3.7. Log10 bin averaged intensities one standard deviation, with IPEs for magnitudes M 7.1 and M 7.3. AW07 (left) computed using distance from closest point on the fault. B06 (right) computed using distance from approximate midpoint of surface rupture.

reinterpreted, with those predicted by the two IPEs for a range of magnitudes (Figures 4.6 - 4.8). RMS was determined point by point: for every observation i, we predicted MMI at distance d_i, where d_i = distance in km from the source. RMS misfit is often illustrated graphically with spatial bin-averages of intensity (Figure 4.7). However, using bin-averaged data decreased the best fit magnitude by 0.2 - 0.4 magnitude units compared to the full dataset. We prefer the point by point RMS analysis because it utilizes more information per point.

Table 3.2. Magnitude, M_{best} , inferred from minimum RMS for the different data-IPE pairs. $M_{best} = M(min(RMS))$ is assessed for distances computed from the epicenter, from the approximate midpoint of surface rupture, from the closest point on the fault, and using distances calculated with respect to the IPE definition (i.e. from closest point on fault for AW07 and from approximate midpoint of surface rupture for B06).

IPE - Data pair	M_{best} from epicenter	M_{best} from midpoint	M_{best} from closest point on fault	M_{best} from IPE source
AW07 - Reinterpreted	7.3	7.2	7.1	7.1
AW07 - Original	7.7	7.7	7.5	7.5
B06 - Reinterpreted	7.0	6.9	6.8	6.9
B06 - Original	7.4	7.4	7.3	7.4
Average	7.4	7.3	7.2	7.2

RMS misfit is defined as:

$$RMS(M_j) = \sqrt{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2 / N}$$

where M_j from 6.0 to 8.3 are earthquake magnitudes

y_i is MMI of observation i, either originally assigned or reinterpreted

$\hat{y}_i$ is expected MMI at d_i based on either IPE for the assumed M_j

N = total number of observations, 1,144 for reinterpretations and 1,137 for originals

The minimum RMS over the range of magnitudes gives the best fit magnitude,M_{best} (Figure 4.8, Table 3.2). This method uses the central value of the coefficients in B06 and AW07, and thus ignores the uncertainty in the IPE coefficients.

Calculating distances from the closest point on the fault yields a slightly lower M_{best} across all IPE-data pairs compared to the two point-source distances, differing by up to 0.2 magnitude units (Table 3.2). This occurs because many of the highest intensity observations are shifted closer to the fault, where intensities are expected to be highest.

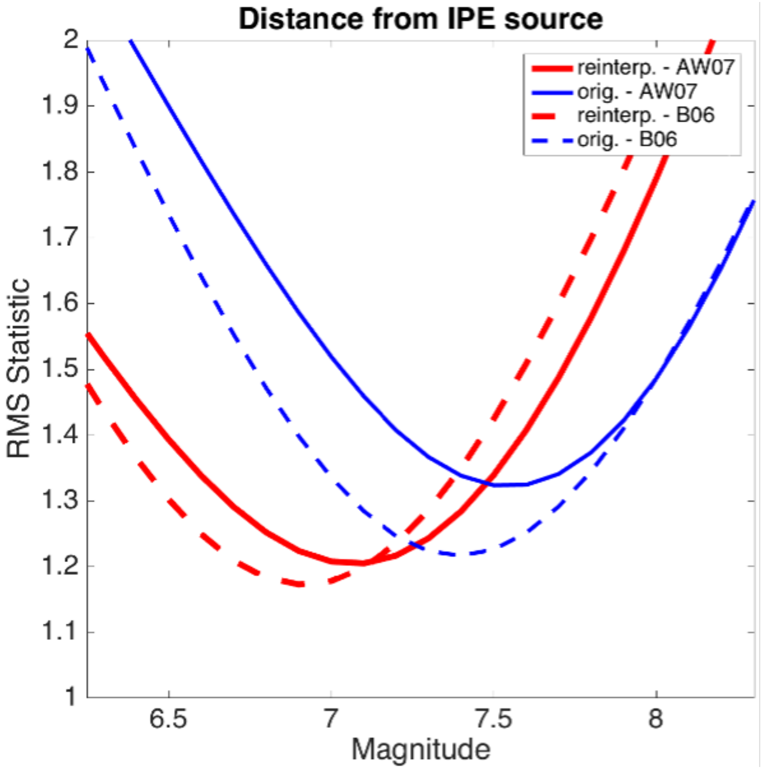

Figure 3.8. RMS misfit between IPEs and reinterpreted or original intensity assignments for a range of magnitudes, computed as distance from closest point on fault for AW07, and from approximate midpoint of surface rupture for B06. The minimum misfit gives M_{best}. Thin lines are original intensity assignments; thick lines are reinterpreted. Solid lines are for AW07; dashed are for B06.

However, both AW07 and B06 still overpredict MMI at distances less than 20 km from the source, regardless of how distance is defined (Figure 4.6). The final column of Table 3.2 considers a mixture of estimates, from closest point on fault for AW07 and from the approximate midpoint of surface rupture for B06. This mixed-source method, which we refer to as "IPE source," is our preferred combination.

Figure 4.8 illustrates the effect of the choice of IPEs and of reinterpreting the data. Reinterpretation decreased the estimated magnitude by 0.4 - 0.5 magnitude units, reflecting our more conservative assessment of the data compared to the original interpretations. Differences between the IPEs result in a difference of 0.1 - 0.3 magnitude units for either dataset, with AW07 M_{best} magnitudes always higher than B06. This occurs because AW07 predicts lower shaking than B06 for a given magnitude (Figures 4.6, 4.7). Hence, to best fit a set of data, AW07 requires a higher magnitude than B06.

The B06 relationship provides the best fit, i.e. lowest minimum RMS over all magnitudes, for both reinterpretations and originally assigned intensities. For reinterpretations, the minimum RMS over all magnitudes only improves slightly from AW07 to B06. This suggests that AW07 is almost as good a predictor as B06 for these data, and potentially for other large historical earthquakes, if the intensities are interpreted carefully, following modern practices, and if sufficient macroseismic information is available to characterize the intensity distribution in detail.

This result appears to be counter to that found by Hough (2014) when analyzing the 1868 Hayward fault intensities as reinterpreted by Boatwright and Bundock (2008a), where B06 fit the reinterpreted historic intensity data significantly better than AW07. Hough (2014) found the RMS to be nearly twice as large for AW07 for a given magnitude. We find considerably better agreement between the minimum RMS values of AW07 and B06 with reinterpreted intensity data for our preferred magnitude of the Kern County earthquake (Figure 4.8).

The difference in which IPE best fits reinterpreted intensity data may be due to the fact that the 1868 Hayward intensity dataset reinterpreted by Boatwright and Bundock

(2008a) only had 160 reports, whereas the 1952 Kern County dataset is more than seven times larger. The increased number of observations decreases spatial sampling biases. We find M_{best} is 6.9 - 7.0 using the B06 IPE with point sources and reinterpreted data, which is lower than the MI 7.3 that Bakun (2006) obtained with the same method from 647 intensities for the Kern County earthquake using distances calculated from the epicenter.

3.7. Uncertainties in M_{best}

How to address uncertainties in the IPEs and hence their suggested M_{best} is a challenge that we approached in several ways. Addressing uncertainties in IPEs requires some assumption about correlations because the correlations between pairs of coefficients in B06 and AW07 are not available. One approach is to entertain the extreme assumption that the uncertainties in the coefficients, represented by the associated ± values, have a correlation of 1, i.e. that the errors all move together in the same direction. In this perfectly correlated errors method, consideration of the errors takes the form of a simple upper/lower bound method by creating an upper and lower IPE for a given magnitude based on the upper and lower range of the coefficients. We then fit the data to the upper and lower IPEs. The M_{best} magnitudes of the minimized RMS for the upper and lower IPEs form the uncertainty ranges for the M_{best} of the central minimized RMS. AW07 provides an error range for only one of its coefficients, whereas B06 provides errors for all coefficients. Hence, the perfectly correlated errors method results in B06 having an apparently larger uncertainty than AW07 (Table 3.3, Figure 4.9). In fact, using this method results in B06's M_{best} uncertainty spanning the entire range of magnitudes tested

Table 3.3. Uncertainty ranges (± SE) of best fit magnitude from different methods for distances calculated in the same manner as the respective IPE when derived (i.e. from closest point on fault for AW07 and from approximate midpoint of surface rupture for B06). This table is summarized in Figure 10.

IPE - Data pair	Perfectly Correlated Errors	Uncorrelated Errors	Random Sample Jackknife	100 km^2 Spatial Jackknife	Standard Deviation of Residuals
AW07 - Reinterpreted	6.9 - 7.2	7.0 - 7.2	7.0 - 7.2	6.9 - 7.3	6.5 - 7.8
AW07 - Original	7.4 -7.7	7.5 - 7.6	7.3 - 7.7	7.1 - 7.9	6.9 - 8.3
B06 - Reinterpreted	6.0 - 8.3	6.4 - 7.5	6.8 - 7.0	6.8 - 7.0	6.5 - 7.4
B06 - Original	6.0 - 8.3	6.9 - 7.9	7.1 - 7.7	7.3 - 7.5	7.0 - 7.8

(6.0 - 8.3). Due to the difference in the number of coefficients with reported uncertainties between the IPEs, this method is not preferred.

Another approach is to assume that the errors of the IPE coefficients are all uncorrelated. We define δ_i as the uncertainty for coefficient $C_i, i = 0, ..., 3$. For B06, the approximate standard deviation is:

$$SD(MMI_{B06}) = \frac{1}{2}\sqrt{\delta_0^2 + (\delta_1 M)^2 + (\delta_2 D)^2 + (\delta_3 log(D))^2}$$

AW07 only provides an uncertainty for one coefficient, so if the other uncertainties are negligible the approximate standard deviation is:

$$SD(MMI_{AW07}) = \frac{\delta_1}{2}$$

(Use of the factor of $\frac{1}{2}$ is based on interpretation of the reported ± amounts as 2 standard deviations). Using these values, we again apply a simple upper/lower bound method

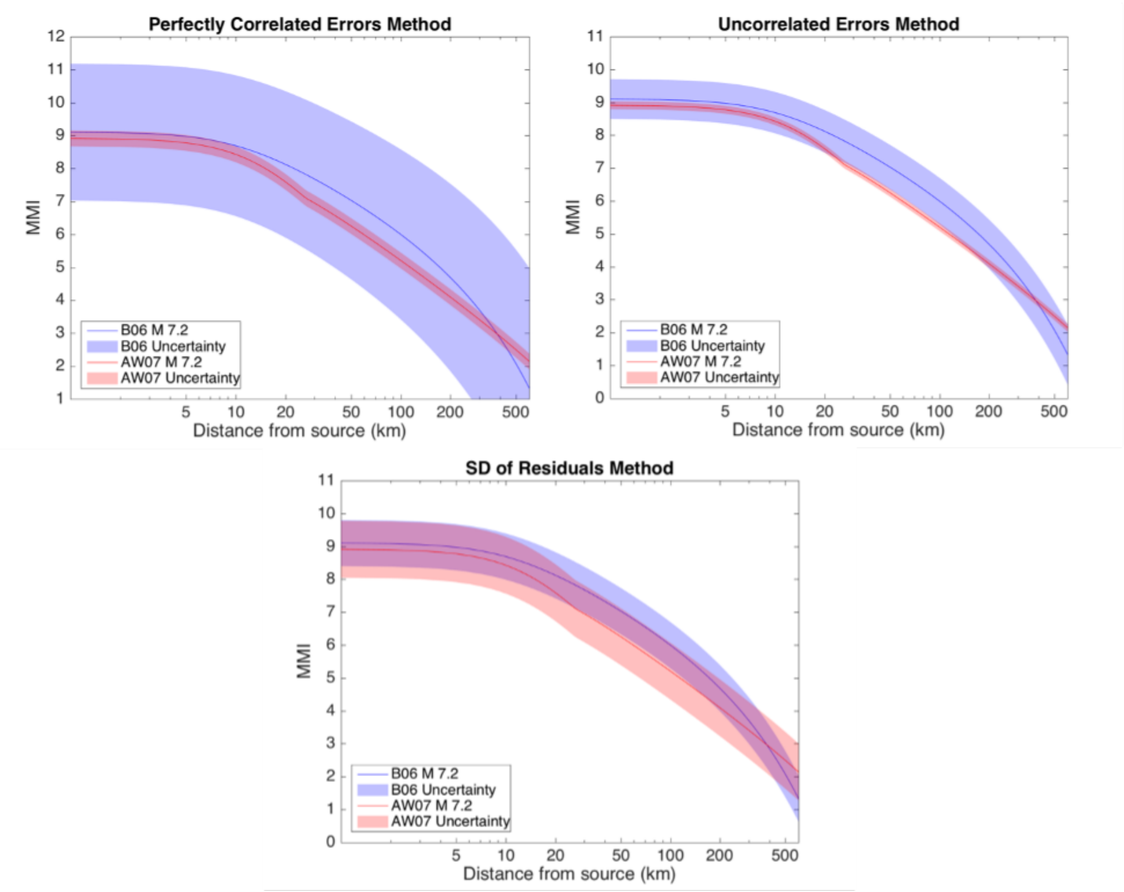

Figure 3.9. Uncertainty of AW07 and B06 MMI with distance for $M7.2$ predictions from: (top, left) perfectly correlated errors method; (top, right) uncorrelated errors method; (bottom) standard deviation (SD) of residuals method, which shows SD for the IPEs with respect to the reinterpreted intensities. Top row methods are influenced by the different number of coefficients with errors provided in the two IPEs, while bottom row is unaffected by this and thus preferred.

to the IPEs. The upper IPE for a given magnitude is the central IPE with one standard deviation added, whereas the lower IPE has one standard deviation subtracted (Figure 9). The results are summarized in Table 3.3. Like the perfectly correlated errors method, this uncorrelated errors method uses an arbitrary assumption of constant correlation and is influenced by the different numbers of coefficients with reported errors between the IPEs. Thus, the uncorrelated errors method is also not preferred.

An alternative approach involved assessing the uncertainty in M_{best} by estimating its variance with a jackknife method (Shao and Tu, 1995). A jackknife approach involves repeatedly computing the desired statistic by leaving out a different subset of the data each time. This method views the data points as coming from a random sample and takes the coefficients in the IPEs as given (and nonrandom). This method has the advantage of not being influenced by the different number of coefficients in the IPE for which uncertainties were reported. We explored two different jackknife approaches, using simple random samples and spatially clustered samples. With a simple random sample approach, each record in the dataset was assigned a random number uniformly distributed between 0 and 1. To make the two datasets evenly divisible by 100, we took a simple random sample of 1,100 observations selected without replacement from each. The data were sorted by the random numbers in ascending order, and the 1,100 observations were formed into 11 groups of 100 observations (i.e. group 1 is records 1-100, group 2 is records 101-200, etc.). Each of the 11 groups is thus a non-overlapping simple random subsample of size 100. Using smaller subsample sizes than 100 led to subsample values of M_{best} that had zero variability. We then repeatedly compute M_{best} for each jackknife replicate sample of 1,000 observations, formed by leaving out successive groups individually.

Spatial correlations are evident in the residual plot (Figure 4.5). To account for this, we also use a spatial cluster sampling jackknife. The study area was spatially gridded into 100 km2 blocks, which we treated as arising from cluster sampling. This resulted in 68 blocks with at least one observation. For the original dataset, the maximum number of observations in a block was 135, the median was 8, and the mean was 16.5. For the reinterpreted dataset, the maximum number of observations in a block was 168, the median was 8.5, and the mean was 16.8. We then computed M_{best} by excluding successive blocks of data individually. We did not include blocks with zero observations in the jackknife.

The jackknife sample excluding the jth group or block yields an estimate of M_{best}, $\hat{\theta}_{(j)}$. Combined, they yield

$$\hat{\theta}_{(.)} = \frac{1}{g}\sum_{j=1}^{g}\hat{\theta}_{(j)}$$

where g is the number of groups for the simple random sample approach, and the number of blocks that contain at least one data point for the spatial clustering approach. The jackknife estimate of variance is

$$\hat{V}_{jack} = \frac{g-1}{g}\sum_{j=1}^{g}(\hat{\theta}_{(j)} - \hat{\theta}_{(.)})^2$$

and the estimate of standard error (SE) is $SE = \sqrt{\hat{V}_{jack}}$. The resulting uncertainties are given by the range of $M_{best} \pm SE$ (Table 3.3).

A final measure of uncertainty comes from the standard deviation of residuals, defined as observed minus predicted MMI, for the different data-IPE pairs. Atkinson and Wald (2007) used this method on the mean MMI of log_{10} distance binned observations (as in Figure 4.7) and found that their model had an uncertainty of 0.4 MMI units. We followed suit and found the standard deviation of residuals for the different data-IPE pairs. Using distances from IPE source, the standard deviations of the residuals computed at the IPE-data pairs respective M_{best} ranged from 0.6 for B06 - Original to 1.0 for AW07 - Original. As done above, we created an upper and lower IPE by adding or subtracting the resulting standard deviation from the central IPE for a given magnitude (Figure 4.9). We then fit the full datasets to the upper and lower IPEs, giving a range of M_{best}. Results are summarized in Table3.3. This method, like the two jackknife methods, is not affected by the number of coefficients with errors provided in the different IPEs. Of the three preferred methods, the standard deviation of residuals yielded the largest estimate of uncertainty in M_{best}.

3.8. Results

As discussed previously, it is likely that neither the AW07 nor the B06 IPE is entirely appropriate for a reinterpreted traditional intensity dataset like that determined in this study. From the results in Tables 3.2 and 3.3, which are summarized in Figure 10, the plausible range in M_{best} values is 6.5 to 7.8. This range encompasses those for reinterpreted data from B06 and AW07 via the two jackknife and standard deviation of residuals methods. Figure 3.10 shows the uncertainty ranges from Table 3.3 and a combined sum of the number of occurrences of specific magnitudes within these ranges. The combined sum yields a well-defined peak at M 7.1 for distances computed with respect to the IPE source.

The AW07 relationship was not expected to be a good predictor of originally assigned intensities, so this pair's outlier status for minimum RMS misfit and corresponding M_{best} is not surprising. To explore the influence of this outlier, we considered the IPE-data pairs with and without AW07 - original. The variance of the RMS among the four IPE-data combinations is minimized at M 7.4 with respect to the IPE source, whereas the variance is minimized at M 7.2 when the outlier is excluded. An unweighted average of M_{best} over all four IPE-data pairs gives M 7.2 when distance is computed from IPE source (Table 3.2). An unweighted average of M_{best} using the three more consistent IPE - data pairs (i.e. excluding AW07 - original) gives M 7.1 for distances computed from IPE source.

Using inverse-variance weighting, we estimate a combined-best magnitude for each uncertainty method considering all four IPE-data pairs and considering the three consistent pairs, i.e. excluding AW07 - original. Using weights designed to capitalize on small uncertainties, the inverse-variance weighting method is:

$$\hat{M} = \frac{\sum_{i=1}^{n} \frac{M_i}{\epsilon_i^2}}{\sum_{i=1}^{n} \frac{1}{\epsilon_i^2}}$$

Where $\hat{M}$ is the combined-best magnitude estimate i is the IPE-data combination n is the number of IPE-data pairs considered, in this study 3 or 4 M_i is M_{best} for combination i, from Table 3.2 ϵ_i is the uncertainty range of M_{best} for combination i, i.e. the ranges from Table 3.3

Results of this analysis are summarized in Table 3.4. The unweighted average of all four IPE-data pairs is the same as the average inverse-variance weighted M_{best} of all four pairs over the five different uncertainty methods.

Based on these results, which give single points, and those summarized in Table 3.3 and Figure 3.10, which give ranges, our best estimate of the magnitude of the Kern County earthquake is $M_I 7.2 \pm 0.2$. Recall that intensity magnitude M_I, a magnitude obtained from intensity data, is derived in terms of moment magnitude M_W and so is designed to reflect

Table 3.4. Results of inverse-variance weighting for the different uncertainty methods, with the average over all methods, using distances calculated in the same manner as the respective IPE when derived (i.e. from closest point on fault for AW07 and from approximate midpoint of surface rupture for B06). Inverse-variance is computed using the three IPE-data pairs with the lowest minimum RMS values (inverse-variance 3) and with all four IPE-data pairs (inverse-variance 4).

IPE - Data pair	Perfectly Correlated Errors	Uncorrelated Errors	Random Sample Jackknife	100 km^2 Spatial Jackknife	Standard Deviation of Residuals	Av.
Inverse - variance 3	7.1	7.1	7.0	7.1	7.2	7.1
Inverse - variance 4	7.3	7.4	7.1	7.2	7.2	7.2

the moment magnitude. Because it is not a moment magnitude in the strictest sense, we prefer the M_I terminology. The reported uncertainty of ±0.2 reflects the collective (and consensus) judgement of the authors, based on our uncertainty analyses, that the odds are 2:1 that the true M_I is in the range 7.0 to 7.4.

In this study, AW07 combined with reinterpreted intensity data using distances computed from the closest point on the fault gives M_{best} that agrees most closely with this preferred magnitude. Hence, we conclude that AW07 best describes the reinterpreted shaking distribution of the 1952 Kern County earthquake.

3.9. Discussion

The different methods presented to infer the uncertainty in M_{best} are sensitive to different aspects of the underlying uncertainties in both the IPEs and the data. The perfectly correlated and uncorrelated error methods account only for uncertainties in the IPEs. The jackknife methods account for sampling errors under assumed hypothetical

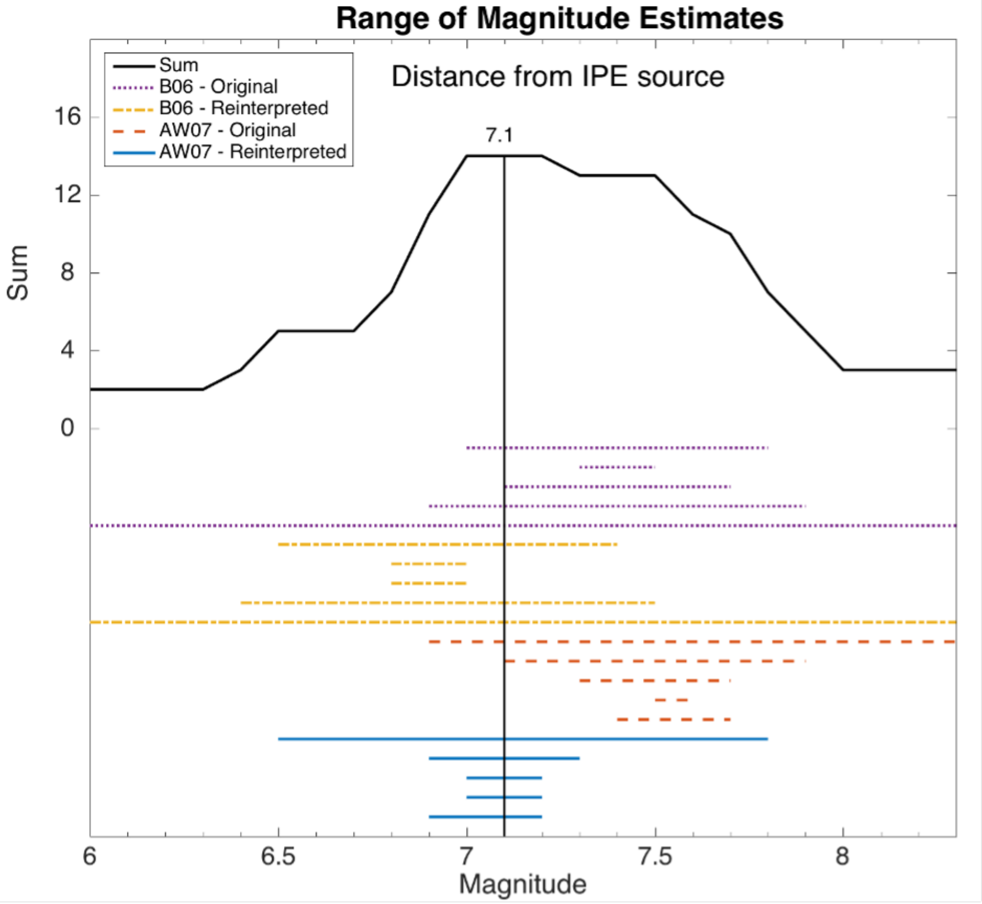

Figure 3.10. Range of M_{best} magnitude estimates from the five uncertainty methods for the four IPE-data pairs (Table 3.3), computed as distance from closest point on the fault for AW07, and from approximate midpoint of surface rupture for B06. The sum line is the sum of the occurrences of a given magnitude, showing a peak at centered on M 7.1.

sampling models. The spatial jackknife additionally takes spatial correlation of data into account. The standard deviation of residuals method is sensitive to both the IPE and, to some extent, the data sampling. The distance-binning of the residuals method takes some spatial correlation into account, although not to the extent that the spatial jackknife method does, while also considering lack of fit of the data to the IPE.

Although questions remain about the consistency of DYFI intensities and traditional intensities assigned following modern practices (Hough, 2013, 2014), we believe that our reinterpretation of historic intensity data for the Kern County earthquake provides an improved characterization of both its magnitude and the ground shaking. Our preferred magnitude estimate is slightly lower than the current catalog estimate (M_W 7.5; Hutton et al., 2010) but consistent with the geodetically determined magnitude (M_W 7.2; Bawden, 2001) and the previously determined intensity magnitude (M_I 7.3; Bakun, 2006). The earthquake generated potentially damaging ground motions (MMI 6 - 7) to distances of over 300 km, with an overall felt extent covering most of California. Our results suggest that shaking was stronger on the hanging (south) wall of the White Wolf fault than on the footwall, a fortuitous result because most of the population in the near-field region was north of the fault.

A growing body of evidence further suggests that near-field shaking in large earthquakes can be significantly tempered by pervasively non-linear response of soft, water-saturated sediments (e.g., Trifunac, 2003; Adhikari et al., 2017). Non-linear effects pose a potential challenge for the development of intensity prediction equations, in particular if they are not well constrained for large magnitudes. Hence extrapolations that are reasonable for relatively weak shaking levels might not be appropriate at the strongest near-field shaking levels. A further practical caution, however, is that the extent of deamplification will depend on the impedance and degree of water saturation of near-surface sediments. As discussed by Trifunac (2003), non-linear effects will be less pronounced on older, more consolidated sediments than on soft, water-saturated sediments.

Our reinterpreted intensity distribution provides a missing puzzle piece for the larger picture of California's maximum historically observed earthquake intensities. Other big pieces are already in place: intensities have been revisited in recent years for the large events of Fort Tejon in 1857, Hayward in 1868, Owens Valley in 1872, Laguna Salada in 1892, San Francisco in 1906, and Santa Barbara in 1925, as well as many moderate historic events (Meltzner and Wald, 1999; Hough and Elliot, 2004; Boatwright and Bundock, 2008a, 2008b; Hough and Hutton, 2008; Hough and Martin, 2018). These reinterpretations, combined with DYFI data for more recent events, will eventually produce a uniform shaking dataset with which earthquake hazard map performance can be assessed.

Probabilistic seismic hazard assessment maps show the amount of shaking a site is expected to experience in a given observation window, for an event with a given return period, with a certain probability. Maps of intensity observations have been used to assess earthquake hazard map performance in Japan, in Italy, and in the Central and Eastern United States (Stein et al., 2015; Brooks et al., 2016, 2017). However, such testing has not been possible in California due to a lack of a long record of consistently interpreted intensity data. Further work will be needed to compile such a dataset for all moderate to large historic and early instrumental earthquakes in California. Further studies, informed by the discussion of uncertainties presented in this study, can also potentially improve magnitude estimates of key events.

CHAPTER 4

California Historical Intensity Mapping Project (CHIMP)

4.1. Summary

Historical seismic intensity data are useful for myriad reasons, including assessment of the performance of Probabilistic Seismic Hazard Assessment (PSHA) models and corresponding hazard maps by comparing their predictions to a dataset of historically observed intensities in the region. To assess PSHA models for California, a long and consistently interpreted intensity record is necessary. For this purpose, the California Historical Intensity Mapping Project (CHIMP) has compiled a dataset that combines and reinterprets intensity information that has been stored in disparate and sometimes hard-to-access locations. The CHIMP dataset also includes new observations of intensity from archival research and oral history collection. Version 1 of the dataset includes 46,502 intensity observations for 62 earthquakes with estimated magnitudes ranging from 4.7 to 7.9. The 162 years of shaking data show observed shaking lower than expected from seismic hazard models. This discrepancy is reduced, but persists, if historical intensity data for the largest earthquakes are smoothed to reduce the effects of spatial under-sampling. Possible reasons for this discrepancy include other limitations of the CHIMP dataset, the hazard models, and the possibility that California seismicity throughout the historical period has been lower than the long-term average. Some of these issues may also explain similar discrepancies observed for Italy and Japan.

4.2. Introduction

Because major earthquakes and the resulting strong shaking are rare events in any one area, it is difficult to assess how well earthquake hazard models and the corresponding maps describe the actual shaking that occurs. The problem is challenging both because of limitations in the available data and because of conceptual issues in how to assess the performance of probabilistic forecasts (Gneiting and Katzfuss, 2014; Marzocchi and Jordan, 2014; Wang, 2015; Vanneste et al., 2018). Ideally, assessments should be prospective, i.e. use only shaking that occurred after a model was made. For example, Brooks et al. (2017; 2019) compared intensities from (presumed) induced earthquakes in the Central United States with predictions of a one-year hazard model. However, shaking data recorded since PSHA began typically span a time period that is short compared to the return period of a PSHA model, and hence rarely include data from the moderate and large earthquakes that control hazard. The methods referenced above have therefore been developed to allow historically observed intensities to be compared numerically to PSHA model predictions, over the duration of the historical catalog.

Retrospective assessments, or hindcasting, using compilations of historical shaking data spanning hundreds of years, can reduce this problem (Stirling and Peterson, 2006; Stirling and Gerstenberger, 2010; Nekrasova et al., 2014; Mak et al., 2014; Stein et al., 2015; Mak and Schorlemmer, 2016; Brooks et al., 2016; 2017a,b; 2019). Such assessments compare PSHA model predictions to historically observed shaking from earthquakes that occurred before the models were made. We use the term “historical” to mean the time before the modern instrumental catalog, i.e. pre-1970s and going back to the earliest available written accounts of shaking.

In this paper, we assess the United States Geological Survey (USGS) National Seismic Hazard Model (NSHM) for California and surroundings that uses fault rupture forecasts from UCERF-3 and Ground Motion Models (GMMs) from the NGA-West2 project (Bozorgnia et al., 2014; Field et al., 2014; Petersen et al., 2014). These models include information about historical earthquakes, but do not explicitly use historically observed shaking intensities, some of which had not yet been compiled when the models were made. Because the hazard model parameters were not chosen to specifically match the past intensity data, comparing the model and historic shaking data can yield insight into the models' performance and potential improvements.

Historical accounts of shaking by witnesses of an earthquake provide the basis for assignment of seismic intensity, a measure based on the effects of shaking on man-made structures and objects within them. Intensity is therefore a good tool for characterizing the distribution of shaking at many locations both near and far from the source. The U.S. government has been collecting first-hand accounts of shaking over the past century, and newspaper accounts go back even further (Byerly and Dyk, 1936; Toppozada and Branum, 2004). These reports have been collected in various government publications and assigned individual intensity levels using the Modified Mercalli Intensity (MMI) scale (Wood and Neumann, 1931; Richter, 1958). Best practices for assigning MMI values have evolved over time (Ambraseys, 1971; Hough, 2014). Compiling a consistent record of intensity data requires consistent reinterpretation of intensity assignments, a need that motivated the development of CHIMP.

This paper presents version 1 of the CHIMP (CHIMP-1) dataset of intensity values from large and moderate earthquakes expected to control the maximum shaking in

California over the past 162 years (Figure 1; see Supplemental Material). The dataset includes "Did You Feel It" (DYFI) data for recent earthquakes ($\sim$ 1990 onward), intensities reinterpreted from felt reports since about 1924 when the U.S. government began collecting them systematically, intensities inferred from historical accounts, and for two earthquakes, intensities from oral history collected by the authors of this study. CHIMP contains consistently reinterpreted intensity assignments for each individual earthquake and a map of historically observed maximum shaking intensities. We also present CHIMP version 1A (CHIMP-1A), which includes smoothed data for the three largest and oldest earthquakes in the dataset.

We compare the CHIMP datasets to the 2018 USGS NSHM (Rukstales and Petersen, 2019) to explore various aspects of how the models perform and possible approaches to improving them. We use different metrics to explore how model performance varies in space and time. The shaking dataset has been developed specifically for comparison with hazard models, so features intrinsic to the historical data are identified and addressed in the performance assessment.

4.3. Historical Macroseismic Intensities

CHIMP builds on past studies of historical and instrumental earthquakes in California, including seminal work by Toppozada et al. (1981) and Boatwright and Bundock (2005; 2008), long-running postcard questionnaire programs by government agencies (e.g., Byerly and Dyk, 1936; Dewey et al., 1995), and, most recently, the USGS “Did You Feel It?” (DYFI) system. DYFI collects macroseismic information over the web and assigns Community Internet Intensity (CII) values using an algorithm (Wald et al., 1999)

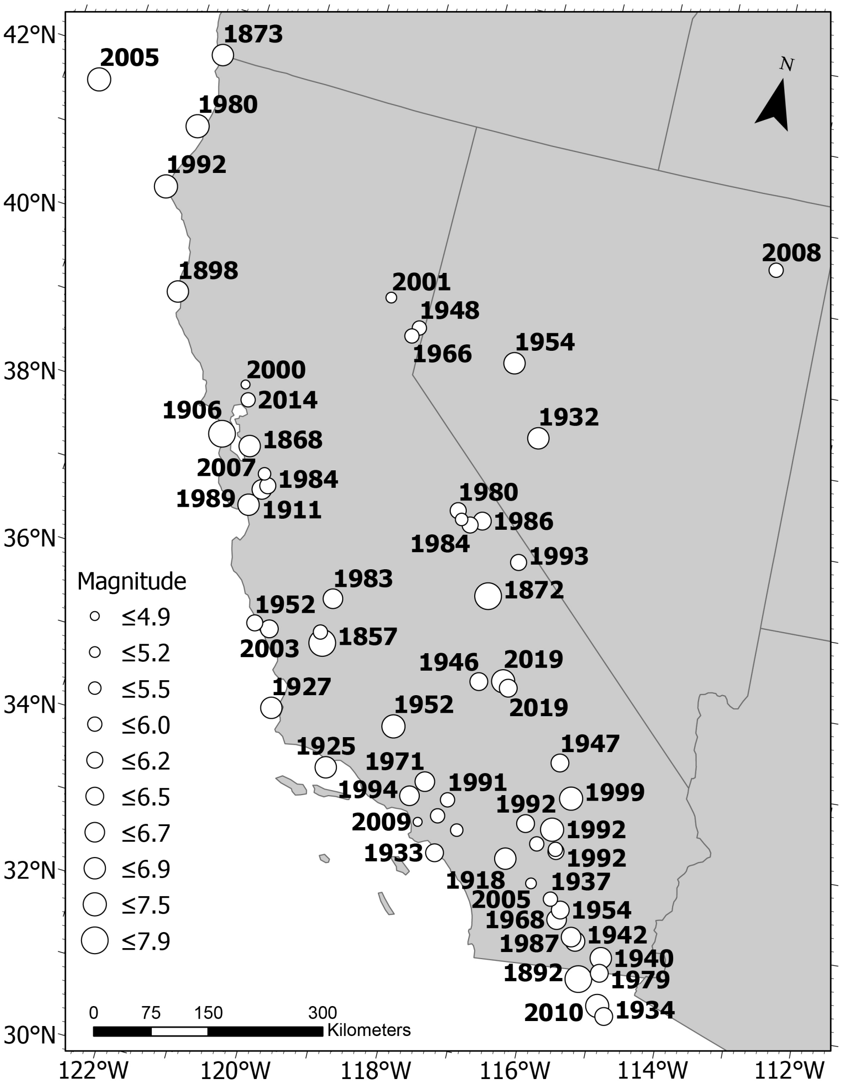

Figure 4.1. Map of earthquakes included in the maximum shaking dataset, labeled by year and scaled by moment magnitude.

adapted from the Community Decimal Intensity (CDI) algorithm developed by Dengler and Dewey (1998). DYFI initially reported average CII values within postal ZIP codes, and increasingly reports intensities for geocoded cells. DYFI intensities have also been collected retrospectively for some large earthquakes. For example, although DYFI debuted in 1999, DYFI intensities from the 1992 Landers and 1994 Northridge earthquakes are available from over 300 and over 700 ZIP codes, respectively. Efforts like CHIMP and DYFI are part of a growing effort worldwide to collect, disseminate, and analyze regional and global seismic intensity information (Stucchi et al., 2004; Locati and Cassera, 2010; LaMontagne and Burke, 2019), although methods and intensity scales still vary widely (De Rubeis et al., 2019). Some intensity data for past earthquakes worldwide (1638 - 1985) are available via the National Oceanic and Atmospheric Administration (NOAA) Earthquake Intensity Database (See Data and Resources). For U.S. earthquakes, this database includes intensities from the United States Earthquakes report series, with some modification. The California Geological Survey's Historic Earthquake Online Database (see Data and Resources) contains felt report summaries and intensity assignments for earthquakes in California and adjoining regions, drawing from earlier work by Toppozada et al. (1981) and others.

Combining and consistently interpreting intensity data from different sources and times requires care, for several reasons:

(1) A large body of work, starting from Ambraseys (1971, 1983), shows that, although environmental effects like landslides and liquefaction potentially provide useful information about ground motions, without careful consideration they are not reliable indicators of peak ground acceleration (PGA), as was assumed in the initial formulation of the MMI scale. For this reason, we exclude reports that mention only environmental effects in CHIMP.

(2) The assignment of intensity based on a written account is inherently subjective, and the degree of conservatism varies among researchers.

(3) The NOAA database is especially problematic, with originally assigned intensities of 1-3 (or 1-4) generally listed as MMI 3 (or MMI 4).

(4) Assessment of intensity values evolved significantly from the early 20th century onward. Early assessments gave more weight to subjective human response, e.g. if people reported being frightened by shaking. Moreover, accepted practice was to assign a higher intensity if a single indicator corresponded to intensity level N, even if overall effects suggested N-1 (or even N-2) (Dewey, personal communication, 2018).

The issue of subjectivity in factor 2) has been addressed in recent work showing that differences in intensity estimates decrease as the level of expertise increases and that assigning intensities through numeric forms, like DYFI and equivalent systems, reduces the difference in estimates from historical documents (Sira et al., 2019). MMI assignments tend to vary by ±1 units between researchers (Hough and Page, 2011; Salditch et al., 2018). Reinterpretations in CHIMP are given as the mean of two independent assignments to help reduce uncertainty. For the 1857 Ft. Tejon earthquake, the oldest earthquake in CHIMP and hence the most uncertain, four independent intensity assignments were made and averaged. Although it remains difficult if not impossible to formally assign uncertainties to intensity assignments, ±1 unit is a reasonable estimate of uncertainty for an individual assignment (see Hough and Page, 2011).

Factor 4) is the most potentially problematic because the assessment of numerical intensities evolved over time from assigning intensities based on the most dramatic reported effects, to assigning representative values for an area. The DYFI algorithm determines representative intensities by averaging responses to each individual question on the questionnaire and calculating a weighted average intensity within a ZIP code or geocoded cell (Wald et al., 1999).

Intensity values have also been colored by a reporting bias whereby the media focus on the most dramatic damage rather than the overall level of damage (Richter, 1958; Hough, 2013; 2014). This bias tends to be strongest when written reports are short, so more extensive accounts typically provide a better sense of overall effects. Thus, the bias tends to be more significant for earlier earthquakes, for which newspaper and other accounts are often especially fragmentary, than for more recent events. A reporting bias still commonly exists in coverage of even recent earthquakes, such that care is needed if intensities are determined from available media accounts.

Hough (2014) proposed a correction curve approach to convert conventional intensities to DYFI intensities. This approach was developed based on older published conventional intensity values. Over time, as noted, the assignment of traditional intensities has evolved toward a more conservative approach, such that subjective assignments better align with DYFI values. The reinterpretations undertaken as part of the CHIMP effort were largely motivated by the need to address this issue for earthquakes for which a modern reinterpretation had not been done by studies such as Boatwright and Bundock (2005, 2008). The CHIMP dataset itself provides an argument against applying a correction curve approach; for example, the extensive set of intensities for the 1906 earthquake (Boatwright and Bundock, 2005) includes 23 MMI 1s (not felt) and 93 2.0-2.5 values, out of a total of 684 locations. The correction curve would reduce these low values, in many cases to intensities below 2, which would be inconsistent with DYFI assignments. We recognize that subjectively determined intensities will never be entirely consistent with those determined by an algorithm such as that used by the DYFI system, a fundamental limitation of the CHIMP dataset. Nonetheless, the value of creating a complete dataset outweighs the unavoidable limitations associated with use of different intensity types.

The above limitations of early intensity data, some of which persist with modern studies, caused some researchers to denigrate the value of such “unscientific” data (Hough, 2000). Characterization of shaking severity by a single number has limitations, because the character of ground motions depends on duration and frequency content as well as peak velocity or peak acceleration. Overwhelmingly, shaking intensities are controlled by frequencies between 1-8 Hz (Trifunac and Brady, 1975; Sokolev and Chernov, 1998). With rare exception, intensity data thus cannot constrain long-period (< 1 Hz) shaking effects (Hanks and Johnston, 1992; Hough, 2014). Accordingly, intensity magnitudes based on macroseismic data alone provide an estimate of energy magnitude, but only limited constraint on moment magnitude (Hough, 2014). By the same token, historically observed intensities are not expected to provide much constraint on the levels of long-period ground motions that will potentially affect large modern structures.

Intensity data do, however, provide an integrated measure of shaking over the main frequency range of engineering concern, and are increasingly recognized to be of great value, if they are interpreted carefully with an appreciation of limitations. To cover the range of perceptible earthquake ground motions with a 10-step intensity scale, each step in intensity must correspond robustly to a factor of approximately 2 in PGA (Hough, 2000). DYFI and other internet-based systems demonstrate that consistently interpreted intensity data provide surprisingly reliable indicators of ground motions (e.g., Atkinson and Wald, 2007; Worden et al., 2012) and can provide important insights about earthquake source parameters, site, and path effects. USGS ShakeMaps - which depict intensity values calculated from instrumental recordings - rely on DYFI data to flesh out shaking distributions, in particular in areas where instrumentation is sparse. As we will discuss shortly, consistently interpreted intensity data can be used for assessing the performance of PSHA models, in particular where PSHA models are cast in terms of intensity (e.g., Brooks et al., 2016; 2017a,b; 2019).

4.4. CHIMP Dataset

The CHIMP dataset encompasses 62 earthquakes occurring between 1857 and 2019 (Figure 1). Because reconsideration of all macroseismic data for California earthquakes would be prohibitively time-consuming, we focus primarily on the set of $M_W > 6$ earthquakes since 1857 that we expect will control maximum observed shaking throughout the state (Table 4.1). We make exceptions on the magnitude cutoff for some recent, smaller events that may control the maximum shaking in their epicentral area. Hence the data go back 162 years but are not uniformly complete - the coverage and number of contributing reports is heavily weighted toward the more recent DYFI end of the dataset. Earlier earthquakes included are fewer and larger in magnitude. Including moderate earthquakes is only possible for the recent DYFI-era events because of the better spatial coverage overall and sensitivity to smaller events. The 1987 Whittier Narrows earthquake is the earliest in the dataset with magnitude lower than M_W 6, followed by Sierra Madre in 1991. The final six earthquakes smaller than M_W 6 occur during 2000 - 2009. Our selection of events for the CHIMP dataset was subjective, so we may have missed other moderate events that controlled intensities in their local area.

Table 4.1. Example reports from primary dataset, comparing original and reinterpreted MMI. Groundwater and secondary geologic indicators were not considered in the reinterpretations.

CHIMP Label	Date (m/d/yr)	Location	Moment Magnitude	Epicenter Longitude	Epicenter Latitude	No. IDP
1857	1/9/1857	Fort Tejon*	7.9	-120.300	35.700	71
1868	10/21/1868	Hayward	6.8	-122.100	37.700	162
1872	3/26/1872	Owens Valley	7.8	-118.100	36.700	147
1873	11/23/1873	Cali/Oregon border	6.9	-124.200	42.000	120
1892	2/23/1892	Laguna Salada	7.8	-115.630	32.550	37
1898	4/15/1898	Mendocino*	6.9	-123.800	39.200	33
1906	4/18/1906	San Francisco	7.9	-122.550	37.750	684
1911	7/1/1911	South Bay*	6.6	-121.750	37.250	38
1918	4/21/1918	San Jacinto*	6.8	-117.000	33.750	141
1925	6/29/1925	Santa Barbara	6.8	-119.800	34.300	237
1927a	11/18/1927	Bishop*	5.5	-118.750	37.500	8
1927b	11/4/1927	Lompoc*	6.9	-120.774	34.813	160
1932	12/20/1932	Nevada*	6.8	-117.910	38.631	3
1933	3/10/1933	Long Beach	6.4	-118.000	33.631	223

1934	12/31/1934	Colorado River*	6.4	-115.176	32.180	39
1937	3/25/1937	Buck Ridge*	6.0	-116.250	33.400	65
1940	5/18/1940	Imperial Valley*	6.9	-115.381	32.844	201
1942	10/21/1942	Fish Creek*	6.6	-115.785	32.975	143
1946	3/15/1946	Walker Pass*	6.3	-117.944	35.702	187
1947	4/10/1947	Manix*	6.5	-116.532	34.983	220
1948a	12/4/1948	Desert Hot Springs*	6.0	-116.331	33.983	275
1948b	12/29/1948	NE California*	6.0	-120.080	39.550	81
1952a	7/21/1952	Kern County	7.5	-118.998	34.958	1,062
1952b	11/22/1952	Bryson*	6.2	-121.328	35.723	86
1954a	3/19/1954	San Jacinto	6.4	-116.081	33.299	149
1954b	7/6/1954	Nevada*	6.8	-118.530	39.420	170
1966	9/12/1966	Northern CA/Truckee*	5.9	-120.160	39.438	62
1968	4/8/1968	Borrego*	6.6	-116.103	33.180	306
1971	2/9/1971	Sylmar	6.6	-118.370	34.416	115
1979	10/15/1979	Imperial Valley	6.4	-115.359	32.667	19
1980a	5/25/1980	Mammoth*	6.1	-118.831	37.590	23
1980b	11/8/1980	Eureka*	7.2	-124.253	41.117	18
1983	5/2/1983	Coalinga	6.7	-120.312	36.232	50
1984a	4/24/1984	Morgan Hill*	6.2	-121.679	37.310	194

1984b	11/23/1984	Round Valley*	6.1	-118.607	37.460	33
1986a	7/8/1986	N. Palm Springs	6.0	-116.608	33.999	40
1986b	7/21/1986	Chalfant Valley*	6.4	-118.443	37.538	32
1987a	10/1/1987	Whittier Narrows*	5.9	-118.079	34.061	127
1987b	11/24/1987	Superstition Hills*	6.6	-115.852	33.015	14
1989	10/17/1989	Loma Prieta	6.9	-121.880	37.036	27
1991	6/28/1991	Sierra Madre	5.8	-117.993	34.270	31
1992a	4/22/1992	Joshua Tree*	6.1	-116.317	33.960	59
1992b	4/25/1992	Rio Dell	7.2	-124.449	40.335	1
1992c	6/28/1992	Landers	7.3	-116.437	34.200	157
1992d	6/28/1992	Big Bear	6.3	-116.827	34.203	310
1993	5/17/1993	Big Pine*	6.1	-117.774	37.165	5
1994	1/17/1994	Northridge	6.7	-118.537	34.213	351
1999	10/16/1999	Hector Mine	7.1	-116.265	34.603	721
2000	9/3/2000	Northern CA	4.9	-122.413	38.379	213
2001	8/10/2001	Northern CA	5.2	-120.617	39.811	295
2003	12/22/2003	San Simeon	6.5	-121.100	35.700	1,070
2004	9/28/2004	Shandon	6.0	-120.366	35.818	565
2005a	6/12/2005	Anza	5.2	-116.567	33.532	489
2005b	6/14/2005	Mendocino	7.2	-125.953	41.292	388

2007	10/30/2007	San Francisco Bay	5.5	-121.774	37.434	5,775
2008a	2/21/2008	Wells, Nevada	5.9	-114.872	41.144	997
2008b	7/29/2008	Chino Hills	5.4	-117.766	33.949	7,289
2009	5/17/2009	Lennox	4.7	-118.336	33.938	589
2010	4/4/2010	Baja	7.2	-115.295	32.286	4,170
2014	8/24/2014	Napa	6.0	-122.312	38.215	3,938
2019a	7/4/2019	Ridgecrest	6.4	-117.504	35.705	1,242
2019b	7/6/2019	Ridgecrest	7.1	-117.599	35.770	12,045

CHIMP includes the 15 April 1898 Mendocino earthquake, the only pre-1900 earthquake in California for which an instrumental magnitude has been determined (Abe, 1994), and several large historical earthquakes: 1857 Fort Tejon, 1868 Hayward, 1872 Owens Valley, 1873 California/Oregon border, and 1892 Laguna Salada. Intensity values have been reinterpreted for these earthquakes (Hough and Elliot, 2004; Boatwright and Bundock, 2008; Hough and Hutton, 2008; Brocher, personal communication 2019 (unpublished USGS OFR)). Intensity distributions have been revisited for several of the largest 20th century earthquakes, including 1906 San Francisco (Boatwright and Bundock, 2008), 1925 Santa Barbara (Hough and Martin, 2018), 1933 Long Beach (Hough and Graves, 2020), and 1952 Kern County (Salditch et al., 2018). DYFI data are available for 25 earthquakes (Table 4.1), including the 1989 Loma Prieta earthquake.

We reconsidered the intensity distributions for 29 earthquakes. For most of these earthquakes, summaries of postcard questionnaires are available in reports published by the U.S. Weather Bureau, Coast and Geodetic Survey, and later by the USGS. The primary contributors are "United States Earthquakes" and "Abstracts of Earthquake Reports." The latter represents an intermediate phase between the primary sources, the original questionnaires/press reports, and the secondary summaries given in "United States Earthquakes." The Abstracts document effects not commonly reported in the former, such as effects observed in low-intensity communities for earthquakes that produced high intensities elsewhere (Dewey, personal communication 2019). The Abstracts from 1929 - 1973 are available as scanned copies online through resources such as the Hathi Trust Digital Library (see Data and Resources) and most exist as paper or digital copies in the archives of the National Earthquake Information Center in Golden, Colorado. The summaries in these reports required transcription and reinterpretation of the originally assigned intensities. These data sources were supplemented by newspaper accounts (e.g., Toppozada et al., 1981), which in some cases were augmented with accounts gleaned from searchable online newspaper repositories and other sources, which again required transcription and intensity assignments.

Full characterization of the shaking distribution of the 29 historical and early instrumental earthquakes would be valuable for myriad reasons, but prohibitively time consuming. Because our work focuses on maximum intensities observed throughout California, and the handful of largest earthquakes (1857, 1872, 1906, etc.) effectively establishes MMI 4 as a lower bound on maximum observed intensities anywhere in the state, we reviewed available sources to compile information for earthquakes where MMI 5 or greater shaking was observed, while still including lower intensities values where described by full felt reports. If a report contained just a location and MMI assignment without a report, we did not include it in CHIMP. For some earthquakes, the few reports of moderately high intensities are straightforward to compile. For more widely felt earthquakes, however, such as the 1918 San Jacinto and 1927 Lompoc earthquakes, macroseismic information is more plentiful, and required more time to review.

Reinterpretation methods for CHIMP differ from the original MMI classification by Wood and Neumann (1931), updated by Richter (1958), in the following ways:

(1) Taking Ambraseys' (1983) conservative approach to secondary geologic indicators.

(2) Using quartile decimal (e.g., 5.25) rather than integer values of MMI when some but not all indicators for a certain shaking level are present.

(3) Giving more weight to the disturbance of objects than to subjective human reactions. Originally, the disturbance of objects and personal reactions were given about equal weight for the lower intensities (USDOC, 1966).

(4) Assigning MMI 5 only when accounts describe toppling of small objects, a key objective indicator for this intensity level (Richter, 1958).

(5) Differentiating between MMI 2 (felt/felt by few) and 3 (felt by many/hanging objects swing), rather than characterizing weakly felt intensities as MMI 1-3, as was the earlier practice in the US Earthquake report series.

(6) Assigning MMI 1 for sites "Reported not felt", following Ambraseys' (1983) recommendation and generally in line with current practice (including DYFI).

CHIMP-1 includes 46,502 intensity data points (IDPs). The final dataset reports MMI to one decimal place, rounded up, in keeping with DYFI standards. The vast majority (88% or 40,887 IDPs) of those points are from DYFI, although they account for just 12% of the time period covered and 48% of the earthquakes. Historical IDPs, of which 5,615 were included in this study, represent the remaining 12%. Of those 5,615 individual historical felt reports, 2,794 were reinterpreted by two authors of this study. We give the mean value of the two assignments in CHIMP-1. These assignments matched 38% of the time, 72% were within ±0.5 MMI units, and 94% were within ±1 unit. Figure 4.2 shows the maximum intensities observed in 10- by 10-km grid cells for which one or more intensity value is included in the CHIMP-1 dataset.

CHIMP can give different levels of detail depending on user needs. So that this information can be reinterpreted later, we provide in the 'repository' (see Data and Resources) the full felt report for each reinterpreted intensity assignment, and give the two assignments that go into the mean. The 'dataset' provides CSV/txt files of longitude, latitude, and MMI/CDI (see Supplemental Material).

4.5. Oral History Collection

In the 1990s, during the transition from postcards to the DYFI system, which went online at the turn of the millennium, the USGS collected intensity information for only the most damaging events, e.g. the M_W 7+ 1992 Landers and 1994 Northridge earthquakes. As a result, two minimally damaging M_W 6.1 earthquakes, those of 1992 in Joshua Tree and 1993 near Big Pine, did not have intensity data compiled in the usual systematic way. Because these earthquakes pre-date the onset of the DYFI system but post-date the

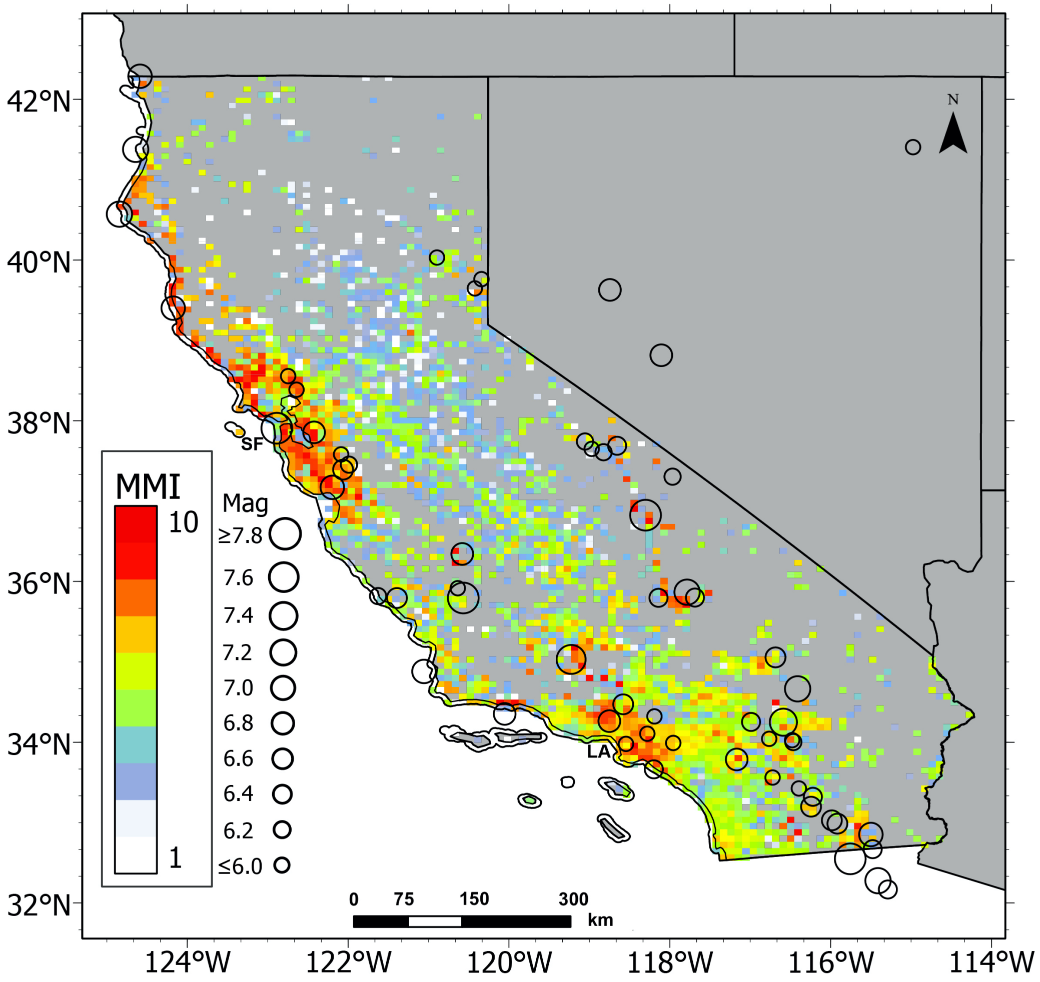

Figure 4.2. CHIMP-1 map of maximum observed intensities in California. Open circles represent the causative epicenters, scaled by moment magnitude. SF is San Francisco, LA is Los Angeles.

collection of postal questionnaires, there is a lack of publicly available intensity information for them.

To correct for this data gap two authors of this study, Salditch and Gallahue, gathered intensity information by collecting oral history reports of shaking by local witnesses and newspaper accounts. This field work resulted in 13 in-person interviews with residents, several more questionnaire responses (designed after DYFI surveys), and dozens of newspaper reports giving 61 new IDPs for Joshua Tree and 5 for Big Pine. The latter location is much more sparsely populated and so provided fewer new IDPs. The new data are especially useful for the Owens Valley portion of the CHIMP map, which has areas without reports in the Sierra Nevada and Basin and Range regions.

Most interviewees learned of the study from advertisements. Local media, including radio station Z 107.7 and The Hi-Desert Star newspaper, covering the high desert around Joshua Tree, as well as The Inyo Register and 100.7 KIBS, covering the Owens Valley/Eastern Sierra Nevada region around Big Pine, ran announcements and articles about the project. The Yucca Valley, Palm Springs, and Joshua Tree Public Libraries generously hosted drop-in interviews. Newspaper accounts came from the archives of the Desert Sun on microfilm at the Palm Springs Public Library, and bound volumes of the Hi-Desert Star available at their offices in Yucca Valley.

This oral history collection could be replicated for other moderate earthquakes occurring during the lifetime of current generations. Memories are the most evanescent source of intensity data available - and therefore most at risk of being lost.

4.6. Retrospective Assessment of Hazard Maps with Historical Data

We assessed the performance of two of the 2018 USGS NSHM models following approaches we have used elsewhere (Brooks et al., 2016; 2017a,b; 2019). These involve comparing the hazard models' forecasts to the historic shaking data using different performance metrics to assess various aspects of how the models performed. Assessing a model's performance can be used to improve future models and the models used to generate them, in that the factors contributing to the model's performance (source locations, maximum magnitude, GMM, etc.) can be evaluated. These comparisons can give insight into current models and possible approaches to improving them.

At any point on a PSHA map, the probability p that during t years of observations shaking will exceed the value shown in a model with a T year return period is assumed to be described by a negative exponential distribution, $p = 1 - \exp(-t/T)$ (Cornell, 1968; Field, 2010). We consider two of the 2018 USGS models, both with $t = 50$ years, but differing in return periods. For one, $p = 0.1$ or 10%, giving T = 475 years. For the other, $p = 0.02$ or 2%, giving $T = 2475$ years. Equivalently, during t years on average 10% and 2% of the sites should experience shaking greater than shown on the map with return period $T = 475$ years and $T = 2475$ years, respectively. Utilizing the ergodic assumption, which states that a system has the same behavior over time as it does over space, we assume that p also reflects the fraction of sites where observed shaking exceeds the modeled value. This approach, introduced by Ward (1995) and used in many subsequent analyses (e.g., Albarello and D'Amico, 2008; Fujiwara et al., 2009; Miyazawa and Mori, 2009; Stirling and Gerstenberger, 2010; Tasan et al., 2014; Nekrasova et al., 2014; Mak and Schorlemmer, 2016) considers many sites to avoid the difficulty that large motions at any given site are infrequent.

Comparison between the predicted and observed maximum shaking allows the calculation of performance metrics (Stein et al., 2015). Assuming that the frequency sample estimates correspond to the probabilities, p, the fractional exceedance metric, $M0 = |f - p|$ is the absolute value of the difference between the observed fraction of points f above the diagonal line in Figure 4.3a - sites where the largest observed shaking exceeds prediction - to the fraction p expected. The remaining sites plot below the line, because the model predicted shaking higher than observed. As the ratio of the observation time to the return period (t/T) increases, p should also increase because an increasing fraction of the area will have experienced larger earthquakes and thus higher shaking.

Hence the shaking shown in a model with $T = 500$ years should be exceeded at 22% of the sites in 125 years ($t/T = 0.25$), 39% of the sites in 250 years ($t/T = 0.5$), and 63% of the sites in 500 years ($t/T = 1$) (Figure 4.3b). $M0$, which is based on the definition of PSHA, has the limitation that it counts exceedances as binary - shaking at a site either exceeded the modeled value or did not. Hence a model with exceedances at exactly as many sites as predicted ($M0 = 0$) could significantly overpredict or underpredict the magnitude of shaking (Stein et al, 2015). Because hazard model assessment is a relatively new enterprise and only a few cases have so far been assessed, there is currently no threshold defined for a "good" score on the $M0$ metric.

The related question of how well a model could realistically be expected to describe observations, given the uncertainties in model parameters and variability in earthquake occurrence, has been investigated via numerical simulations of ground motion by assuming earthquakes occur randomly within a study area (Vanneste et al., 2018). As shown by Vanneste et al. (2018) (Figure 4.3c), some places experience shaking higher than on the hazard map, while others experience shaking lower than shown on the map. For example, after only 50 years some sites experienced shaking stronger than shown on the $T = 2500$ year map. When large earthquakes happen, shaking often exceeds that shown on the hazard maps. As the observation time increases, the fraction of sites exceeding the

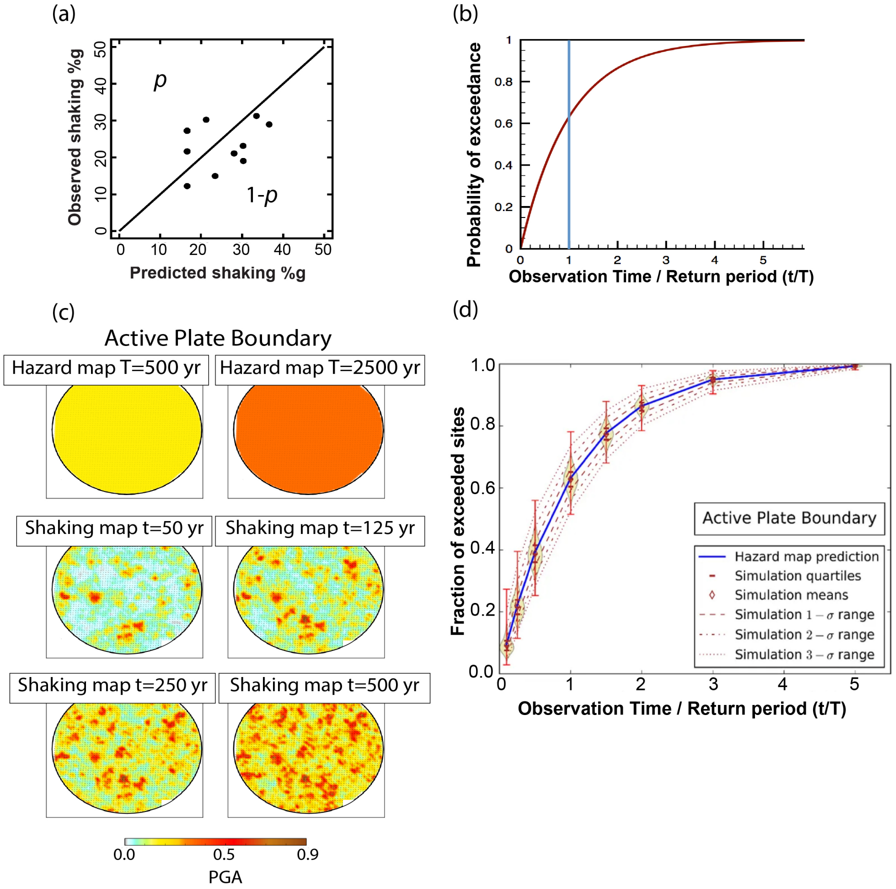

Figure 4.3. a) Schematic observed-vs.-predicted plot showing p, the fraction of sites at which the largest shaking exceeded the mapped values. b) p should increase as the ratio of the observation time t to the model's return period T increases. c) Numerical simulation of maximum shaking over time assuming earthquakes occur randomly (with a spatially uniform distribution) within the model area with seismicity comparable to active plate boundaries. Top row: Hazard models for return periods of 500- and 2500- years. The models and associated maps are uniform across the area, because the expected level of shaking is the same. Middle and bottom rows: Maps of maximum shaking at each point after observation times of 50, 125, 250 and 500 years, for one simulation. d) Comparison between expected and "observed" fractions of sites where maximum shaking exceeds that predicted by the 500-year hazard model as a function of observation time. Scatter decreases for longer simulations. (Panels c) and d) modified from Vanneste et al., 2018).

mapped value increases. An ensemble of simulations yields shaking distributions whose mean is consistent with the model, but individual shaking histories show large scatter (Figure 4.3d). The scatter decreases for longer simulations (increasing t/T), because as observation time increases, the largest earthquakes and resulting shaking are increasingly likely to have occurred.

We also define a squared-misfit metric $M1 = \frac{1}{N} \sum s_i - {x_i}^2$ where x_i and s_i are the maximum observed shaking and predicted shaking at each of the N sites (Stein et al., 2015). Graphically, $M0$ reflects the fraction of sites plotting above the diagonal line in Figure 4.3a, whereas $M1$ reflects how close to the line sites plot. $M0$ is a metric based on the definition of PSHA, describing how well a PSHA model predicts the fractional exceedance that occurs. $M1$, a metric which is not based on the definition of PSHA, describes how spatially similar the observed shaking and hazard model are. $M1$ quantifies the comparison of maps of predicted and observed shaking, similar to a visual comparison.

The two metrics characterize different aspects of map performance. Hence, together they give a fuller picture of map performance than one measure could (Stein et al., 2015; Brooks et al., 2017a,b; 2019). PSHA models do not predict specific shaking levels but rather probabilities of shaking exceedance. Thus a lower $M1$ score does not necessarily mean that the map has performed better - as measured by $M0$ - than one with a higher $M1$ score. Situations may arise where decreasing $M1$ may produce larger $M0$ scores. When comparing metric scores, it is important to be aware of potential tradeoffs in these metrics.

Because hazard model assessment is a relatively new enterprise and only a few cases have so far been assessed, there is currently no threshold defined for a “good” score on either the $M0$ or $M1$ metrics. These metrics, and others that can be used (Stein et al., 2015), are thus most useful as a tool to compare maps to observations and assess the performance of different maps. Future consideration of many maps and numerical simulations should provide improved understanding of the meaning of high- and low-metric scores (Vanneste et al., 2018; Brooks et al., 2019).

4.7. Comparison of Maps and Dataset for California

Using the performance metrics described above, we compare the CHIMP-1 dataset with the 2018 USGS time-independent seismic hazard models for California (Rukstales and Petersen, 2019). We limit our assessment to the state of California because the rupture forecast model used there (Field et al., 2014) differs from the models used in surrounding states like Nevada and Arizona (WGCEP, 2013). Figure 4.2 shows the maximum shaking values in CHIMP-1, which are sorted into 10 x 10 km grid cells, giving 38% spatial coverage of California. For the comparison, we use two of the 2018 hazard models which assume a reference site condition to be National Earthquake Hazards Reduction Program (NEHRP) site class boundary B/C, corresponding to firm rock/very dense soil and soft rock, and a $V_{s30} = 760 m/s$ (Petersen et al., 2020). One has $T = 475$ years and hence 10% probability of exceedance in 50 years, and one has $T = 2,475$years corresponding to a 2% probability of exceedance in 50 years. USGS PSHA models include exceedance forecasts for a number of ground motion parameters. Models of hazard in terms of MMI are available, calculated using Worden et al.'s (2012) ground motion intensity conversion equations (GMICE) to convert Peak Ground Acceleration (PGA) to MMI intensity. Mapped hazard values (Figure 4.4) are shown for points at which CHIMP-1 has values.

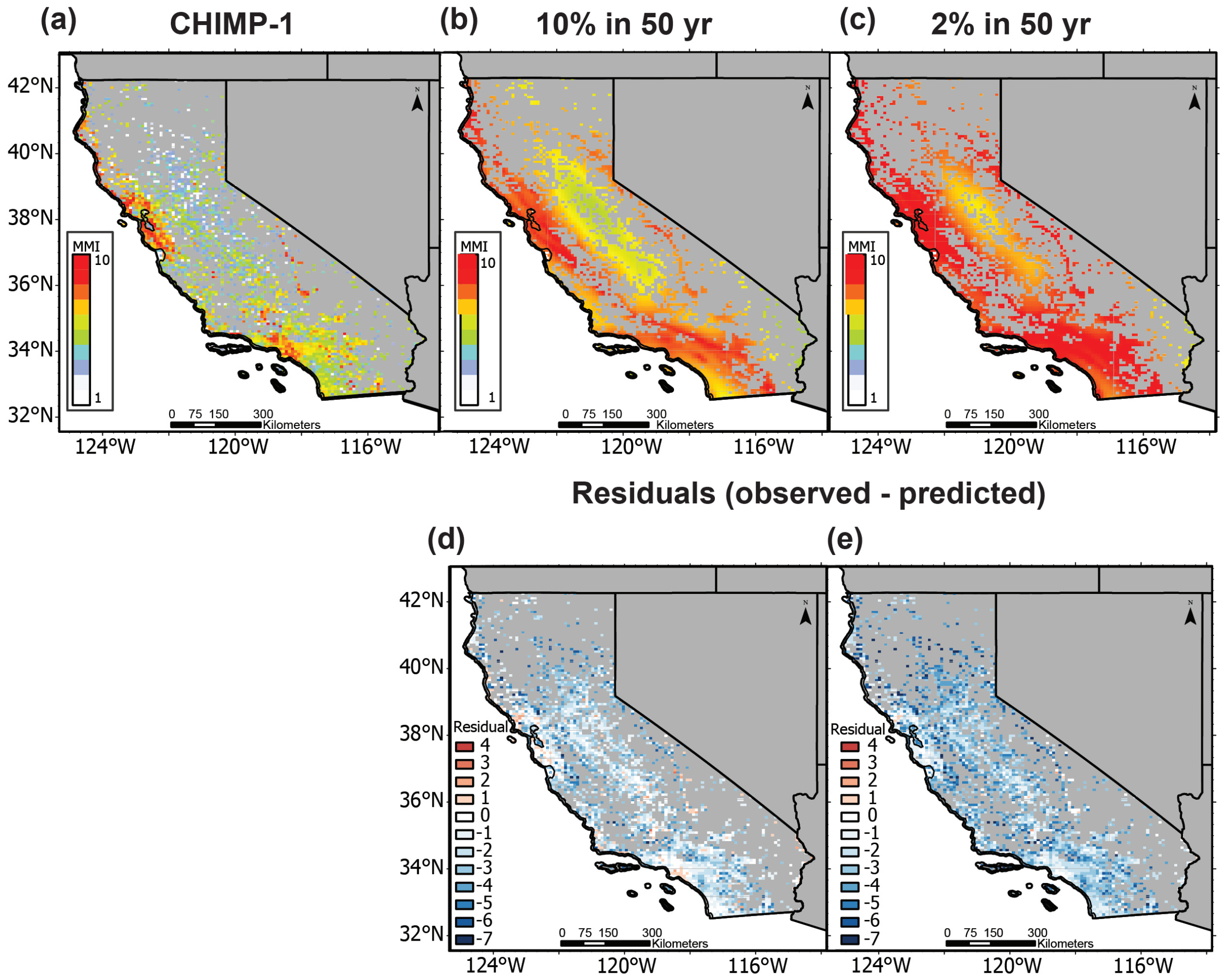

Figure 4.4. Top: Comparison of (a) CHIMP-1 maximum shaking dataset to the 2018 USGS hazard models, for (b) 475-year return period model with 10% chance of exceedance in 50 yrs and (c) 2,475-year return period model with a 2% chance of exceedance in 50 yrs. (d), (e) Residuals (dataset - model values) relative to both models.

In general, the shaking data are similar spatially in trends to the models. Some fault segments have not experienced a large earthquake since 1857, notably the southern

San Andreas. Such effects should ideally be accounted for in the model via the PSHA algorithm. For the 162 years of observations, $p = 1 - \exp(-162/475) = 29\%$ of the sites should have experienced maximum shaking greater than shown in the 475-year model. A smaller fraction, $p = 1 - \exp(-162/2475) = 6\%$, should have experienced maximum shaking greater than the values in the 2475-year model, because 162 years is a much smaller fraction (6% vs. 34%) of the model return period. The longer return period model predicts higher shaking because the largest earthquakes and shaking are more likely to occur during the longer return period. However, comparison of the largest observed shaking at sites to predictions of the hazard models shows observed maximum shaking (Figure 4.4) and fractional exceedances (Figure 4.5) less than predicted. The residual plots of (observed - predicted) intensity in Figure 4.4 show the distribution of discrepancies. The 475-year return period model has a mean residual of -2.3 MMI, and the 2475-year return period model has a mean residual of -3.4 MMI. Thus, on average, the observations are approximately 3 MMI units lower than the hazard model values. The $M0$ score for the 475-year model, where f is more than four times smaller than p, is 0.2248 (Table 4.2). $M0$ for the 2,475-year model, where f is an order of magnitude smaller than p, is 0.057 (again assuming that frequency sample estimates correspond to probabilities, p). The $M1$ score for the 475-year model is 8.800, and 15.048 for the 2,475-year model.

4.7.1. Possible Causes of Discrepancy

This apparent overprediction of intensities by the PSHA model may arise due to biases in the dataset, the hazard model, or chance. We discuss each of these possible biases in this section.

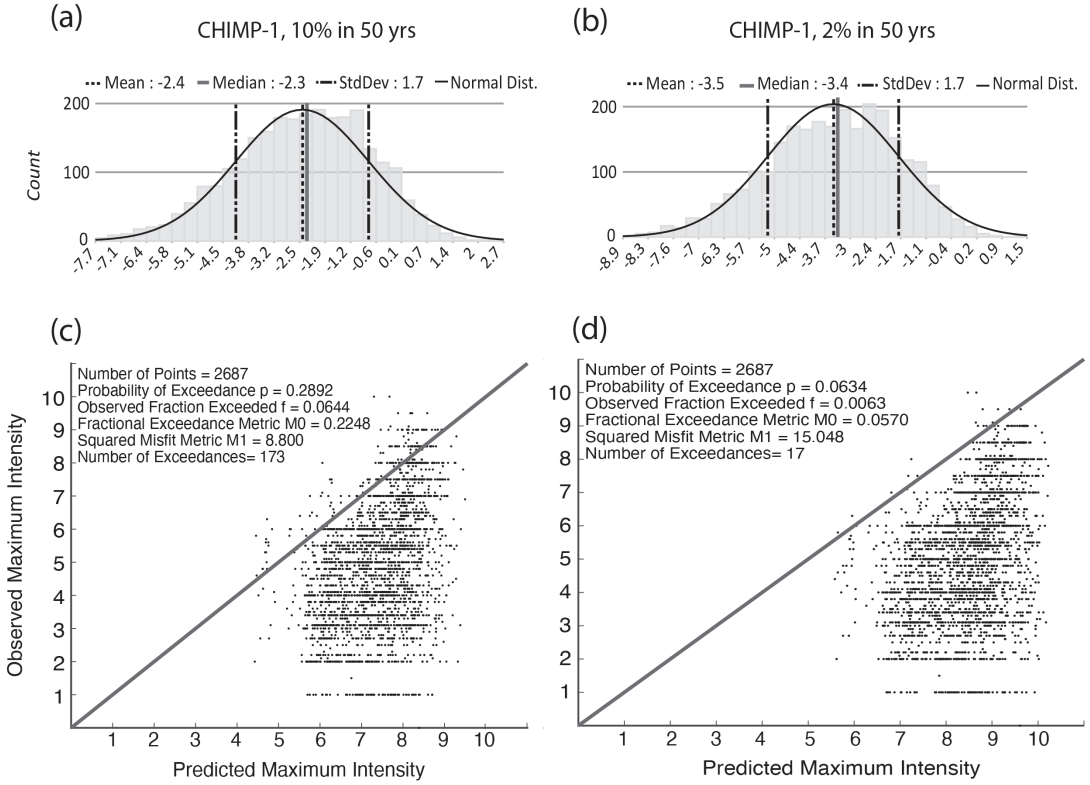

Figure 4.5. Histograms of residuals for CHIMP-1 points in California relative to a) 2018 USGS hazard model for 10% in 50 yrs and b) 2018 USGS hazard model for 2% in 50 yrs. Corresponding predicted-vs.-observed plots for CHIMP-1 relative to c) 10% in 50 yrs model (c) and d) 2% in 50 yrs model.

Table 4.2. Comparison of performance metrics between CHIMP-1 (maximum observations only) and CHIMP-1A (maximum observations plus smoothed data for 1906, 1857, and 1872).

Return period	Model Probability Exceedance	Dataset Version	p	f	$M0$	$M1$
475 yrs	10% in 50 yrs	CHIMP-1	0.2892	0.0644	0.2248	8.800
		CHIMP-1A	0.2892	0.0989	0.1903	5.656
2475 yrs	2% in 50 yrs	CHIMP-1	0.0634	0.0063	0.0570	15.048
		CHIMP-1A	0.0634	0.0048	0.0585	11.152

4.7.2. Data Underestimation

The intensity data may be biased low. In particular, the shaking data do not capture the full extent of shaking from large historic earthquakes in the dataset, due to lack of population. Historical accounts, and hence historical intensity assignments, are biased by the locations and growth of population centers through time. Other potential issues may result from the non-uniform spatial sampling that becomes sparser the further back in time one goes: DYFI data are denser than traditional intensity data estimated from written or archival accounts, and values inferred from historical accounts are sparser still. Hence, some cells do not have a reported intensity value from the large historical events. Thus, if shaking from a more recent smaller earthquake is reported there, the maximum observed intensity in the CHIMP-1 dataset will be too low.

To address this incompleteness due to historical population distribution, we developed CHIMP-1A by adding intensity values interpolated from the existing historical IDPs for the large 1857, 1872, and 1906 events. We used the Natural Neighbor interpolation method in ArcGIS to produce a smoothed intensity map for each earthquake, which we then used to augment the CHIMP-1 maximum shaking dataset. This algorithm assigns a value to a query point by finding the closest subset of input samples and weighting the value proportionally to the area of overlapping Voronoi polygons (Sibson, 1981). This method produces an objective estimate because it locally interpolates values depending only upon known data points and their spatial distribution. In contrast to traditional hand-drawn isoseismal maps, it is simple and reproducible (Sirovich et al., 2002). However, because hand-drawn isoseismal maps were created using expert judgement and knowledge of local geology, we compared our smoothing results to them (Stover and Coffman, 1993) to confirm that results are reasonable. We added the smoothed data for each of the three historical earthquakes to the CHIMP-1 maximum observed dataset to study their effects on performance metrics (Figure 4.6). Addition of the smoothed data more than doubles the total number of grid cells with observations and increases the total number of exceedances by a factor of 2 to 3 (Figure 4.7). The residual plots now show slight underprediction in an area reaching from the Owens Valley to the Transverse Ranges, and slight overprediction in northernmost California.

The overall overprediction remains for CHIMP-1A but is slightly reduced. The mean residual for the $T = 475$-year model improves from CHIMP-1 to -1A by 0.4 MMI units to -1.9 MMI, and by 0.5 units to -2.9 MMI for the 2475-year model. Interestingly, $M0$ remains largely unchanged, improving by 0.03 for the 475-year model and by just 0.001 for the 2475-year model (Table 4.2). However, $M1$ improves for both models, becoming 5.656 for the 475-year model and 11.152 for the 2,475-year model.

A further possible data bias stems from the incompleteness of the CHIMP-1A dataset, via our assumption about which events will control maximum observed intensities. Earthquakes smaller than M_W6 can generate locally high intensities. Because they are more common than larger earthquakes, they may control the maximum historically observed intensity in any one location. The DYFI database itself illustrates this point. Since the introduction of the DYFI system in 1999, 18 earthquakes, 11 of which are between $M_W4.7 - 5.9$, generated CDI values of 6.0 or higher. (We do not consider earthquakes smaller than M_W 4.7). Of 48 earthquakes since 1999 that have generated CDI values of 5.0 or higher, 39 are smaller than M_W6. It would be possible to assess the extent to which

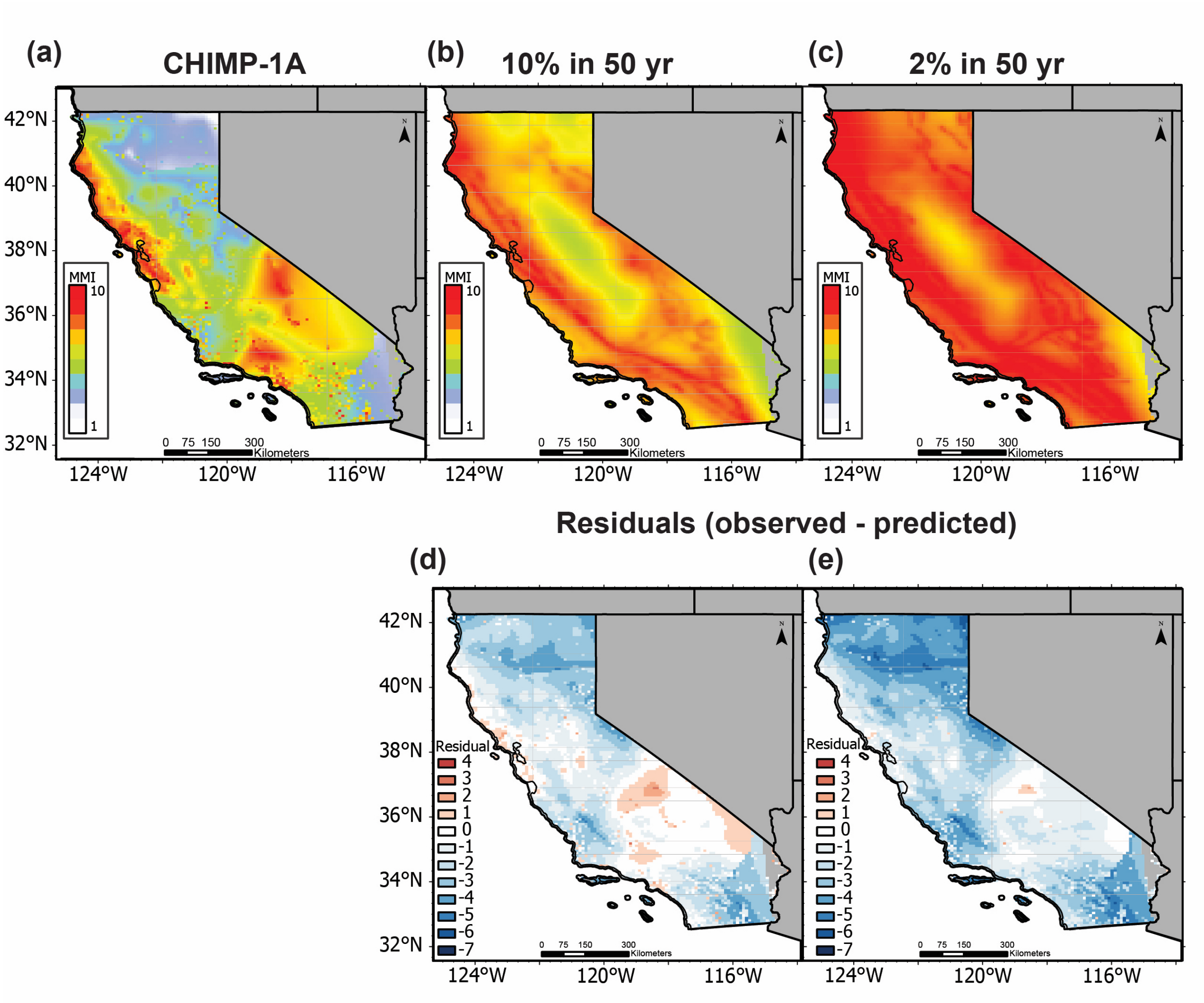

Figure 4.6. Comparison of CHIMP-1A maximum shaking dataset, which includes smoothed data for 1857, 1872, and 1906 earthquakes (a), to the 2018 USGS hazard models with (b) 475-year return period with 10% chance of exceedance in 50 yrs and (c) 2,475 year return period with a 2% chance of exceedance in 50 yrs. (d), (e) Residuals (dataset - model values) relative to both models.

earthquakes smaller than M_W 6.0 contribute to the PSHA map by recalculating the map using only $M_W \geq 6.0$ sources.

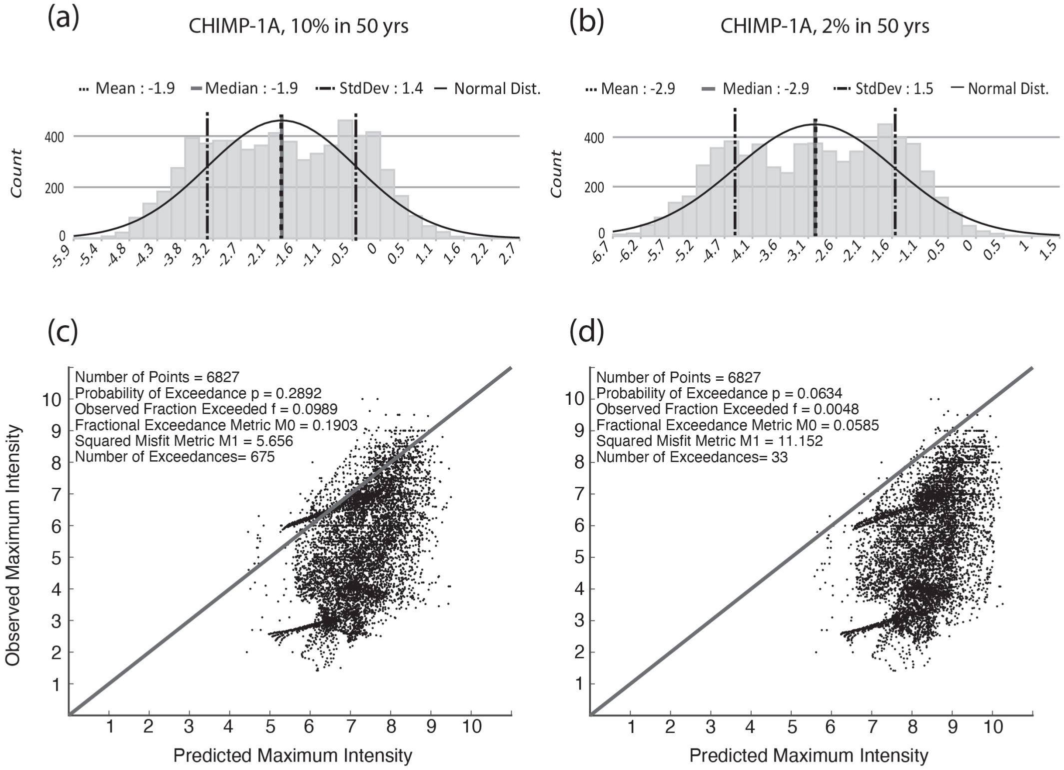

Figure 4.7. Comparison of CHIMP-1A, which includes smoothed data for the 1906, 1857, and 1872 earthquakes, to hazard maps. Histograms of residuals for CHIMP-1A points in California relative to a) 2018 USGS hazard model for 10% in 50 yrs and b) 2018 USGS hazard model for 2% in 50 yrs. Corresponding predicted-vs.-observed plots for CHIMP-1A relative to c) 10% in 50 yrs model and d) 2% in 50 yrs model.

4.7.3. Model Overprediction

The model may be biased toward higher levels of shaking by various effects. One possible cause is that California seismicity during the historical period may have been lower than the long-term average, either due to random variability of earthquake occurrence or a stress-shadow due to heightened activity from 1800 - 1918 that released much of the accumulated stress, though the latter is debated (Harris and Simpson, 1998; Felzer and

Brodsky, 2005; Biasi and Scharer, 2019). There is compelling evidence that California has been in a state-wide lull in seismic moment release since 1910. Biasi and Scharer, (2019) show that, effectively, the UCERF3 map overpredicts the rate of surface-rupturing earthquakes on the San Andreas/San Jacinto/Hayward fault systems. The most recent earthquake rupture forecast model for California (Field et al., 2014) in fact assumes that long-term state-wide earthquake rates are slightly higher than rates through the entire historic period. A number of other factors could bias the models to too-high values, including the normal tendency toward conservatism in engineering design (Vick, 2002).

4.7.4. Chance

Some of the misfit may arise purely by chance. Figure 4.3 illustrates this effect for an ideal hazard model. Even if all parameters are perfectly known, the variability of earthquake recurrence can give rise to a range of values. In a real case, where parameters are not known a priori and are unlikely to be exactly estimated, the variability could be larger. If the ideal case (Vanneste et al, 2018) is representative, then these simulation studies indicate that the misfit is large enough that it is unlikely to have arisen purely by chance due to variability in earthquake recurrence. Hence it likely represents, at least in part, biases in the hazard model, data, or both. Increasing the observed shaking at all sites in CHIMP-1 and CHIMP-1A by a constant shift (Figure 4.8), which is possible given the uncertainty in historical intensity assignments, improves $M0$ and $M1$. Figure 4.8 shows the uncertainty range of the metrics for each dataset-model pair given the inherent MMI uncertainty of ± 1. The minima of $M0$ are within the uncertainty range of the intensities for all but the CHIMP-1 475-year model, while the minima of $M1$ are outside that range

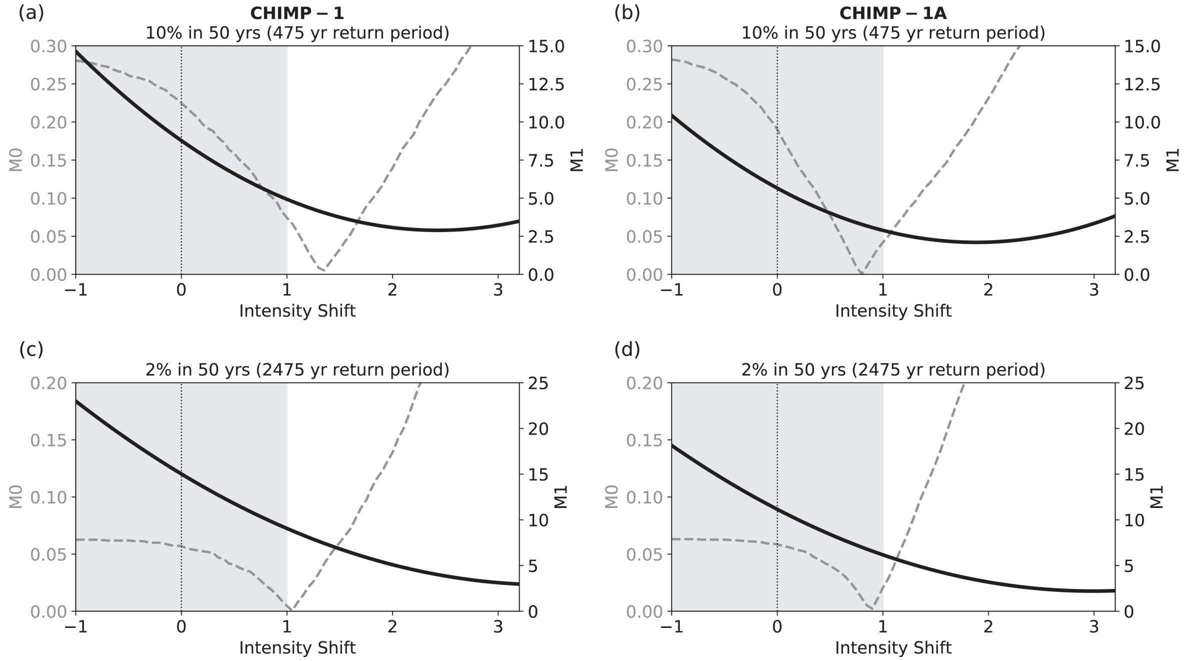

Figure 4.8. Effect of applying a uniform shift to either the map's predictions or the CHIMP-1 observations on performance metrics M0 (dashed grey line) and M1 (solid black line). The uncertainty for intensity values typically ranges ±1 unit (grey shaded region). Lines indicate how the metrics would change given a bulk shift. Positive intensity shift values correspond to an increase in CHIMP observations or a decrease in the map's predictions. Negative intensity shifts reflect a decrease in CHIMP observations or an increase in the map's predictions. (a) CHIMP-1 intensities compared to the 2018 USGS 10% in 50 yr model. (b) CHIMP-1A intensities compared to the 2018 USGS 10% in 50 yr model. (c) CHIMP-1 intensities compared to the 2018 USGS 2% in 50 yr model. (d) CHIMP-1A intensities compared to the 2018 USGS 2% in 50 yr model.

for all models. A similar change could result from decreasing the predicted shaking, or a combination of both effects.

4.7.5. Chance

Some of the misfit may arise purely by chance. Figure 4.3 illustrates this effect for an ideal hazard model. Even if all parameters are perfectly known, the variability of earthquake recurrence can give rise to a range of values. In a real case, where parameters are not known a priori and are unlikely to be exactly estimated, the variability could be larger. If the ideal case (Vanneste et al, 2018) is representative, then these simulation studies indicate that the misfit is large enough that it is unlikely to have arisen purely by chance due to variability in earthquake recurrence. Hence it likely represents, at least in part, biases in the hazard model, data, or both. Increasing the observed shaking at all sites in CHIMP-1 and CHIMP-1A by a constant shift (Figure 4.8), which is possible given the uncertainty in historical intensity assignments, improves $M0$ and $M1$. Figure 4.8 shows the uncertainty range of the metrics for each dataset-model pair given the inherent MMI uncertainty of ±1. The minima of $M0$ are within the uncertainty range of the intensities for all but the CHIMP-1 475-year model, while the minima of $M1$ are outside that range for all models. A similar change could result from decreasing the predicted shaking, or a combination of both effects.

4.8. Reconsideration of Significant Earthquakes

Although retrospective assessment of PSHA models motivated our creation of CHIMP-1, a dataset of observed intensities is potentially useful for many other purposes. Intensity datasets can be used to revisit magnitudes and locations of historic and early instrumental earthquakes, and to further explore the distribution of ground motions generated by recent as well as historic events. As an example, Figure 4.9 presents a shaking intensity map for the 1987 M_W5.9 Whittier Narrows earthquake. To generate this distribution, we augmented the CHIMP-1 values with retroactively contributed DYFI (reported well after the time of the earthquake) intensities at all shaking levels (1-km geocoded values) and intensities estimated from available PGA data (see Data and Resources) using the Worden et al. (2012) GMICE. The intensity distribution is constrained by intensity values at 816 locations, revealing variability of shaking across much of the greater Los Angeles metropolitan region. There is a suggestion, for example, of elevated intensities inland of the Newport-Inglewood Fault, where the Los Angeles Basin deepens considerably. Similar amplification was observed in the 2008 M_W5.4 Chino Hills earthquake (Hauksson et al., 2008). The CHIMP-1 dataset will provide an opportunity to explore key questions regarding ground motions, including the variability of site response across geologically complex regions.

4.9. Discussion and Future Work

CHIMP illustrates the value of developing consistently-interpreted shaking datasets and comparing them to hazard models. We see similar discrepancies, with historical intensity data much lower than hazard model predictions, for Italy and Japan (Stein et al., 2015; Brooks et al, 2016). This could be coincidence, or could indicate a common bias. For example, the hazard models' assumed seismicity and fault slip rates may be too high, the ground motion models may predict too-high shaking, or site effects could result in localized deamplification. We plan to investigate how well the ground motion models fit the historical intensity data, as we did for the 1952 Kern County earthquake (Salditch et al., 2018). Alternatively, the data may be biased low due to spatial sampling bias.

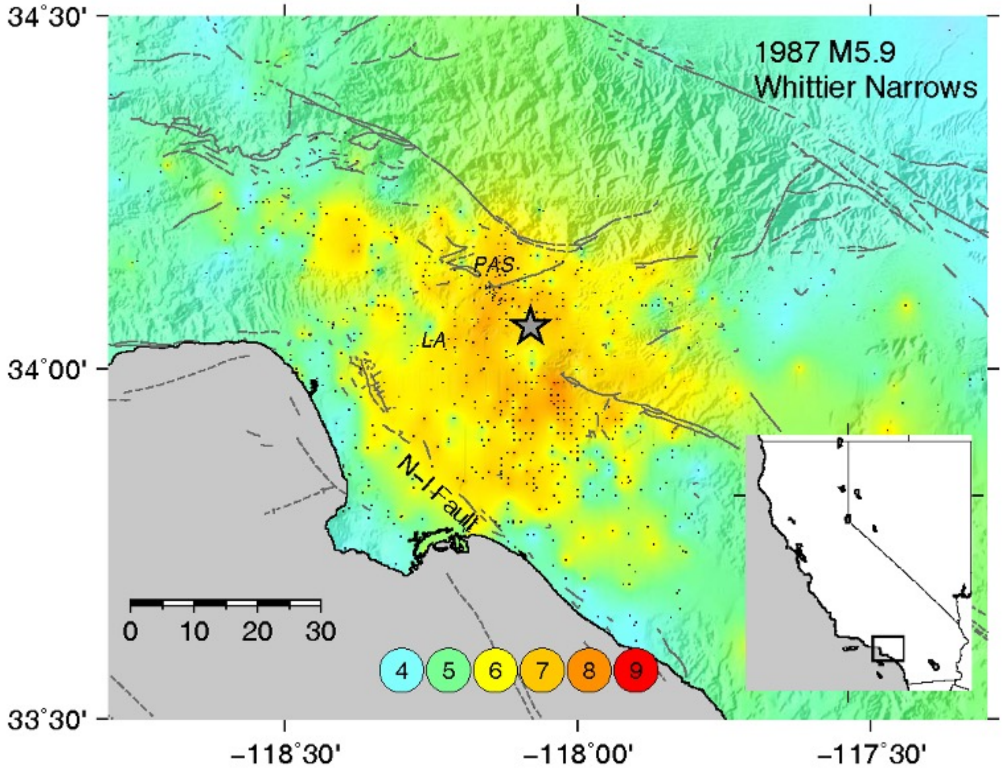

Figure 4.9. Shaking intensity map for the 1987 $M_W5.9$ Whittier Narrows earthquake, generated using data from the CHIMP-1 dataset, augmented by retroactively contributed DYFI data at all distances and intensity values calculated from instrumentally recorded PGA data, converted to intensity using the Worden et al. (2012) GMICE. Locations of Pasadena (PAS), central Los Angeles (LA), and the Newport-Inglewood fault zone (N-I Fault) are indicated. Inset shows location of map within California.

Even after including smoothed data for the largest events, CHIMP-1A may biased low by locally high intensities from moderate ($M_W 4.5-6$) earthquakes that are not in the CHIMP dataset. A statistical modeling approach might address this by estimating how many M_W 5s and M_W 6s are missing and calculating their expected MMI distribution using Atkinson et al.'s (2014) intensity prediction equation. Another way might be to start with the residual data, hypothesize that all of the misfit was due to small events, infer how many are needed, and see if that makes sense relative to the b-value curve. The DYFI dataset collected to date suggests that moderate earthquakes, i.e., $M_W 4.7-5.9$, may be important in controlling hazard due to their prevalence.

To address possible bias introduced by the different intensity data sources, detailed archival accounts of earthquake effects could be entered into the DYFI numerically coded questionnaire and then averaged according to DYFI procedures. This would provide a direct comparison of the consistency of subjective CHIMP intensities and DYFI values. Preliminary experiments with the 1987 Whittier earthquake data do demonstrate consistency between retroactively reported DYFI, historically assigned MMI, and instrumental data (PGA recordings converted to MMI via Worden et al.'s (2012) method. Converting CHIMP MMI to DYFI CDI, may not, as noted, be warranted, but if further analysis suggests discrepancies, the task would generally be tractable in the U.S. given the relatively small size of the historical catalog. The conversion would be less tractable in Europe, for example, where the historical record is thousands of years longer than the electronic.

Discussion of the maximum observed shaking for an area leads to questions of the second highest shaking, the third highest shaking, etc. This question usually stems from desire to investigate the effects of removing a single earthquake from the record. Does the performance of the model change with the removal of a single event? By how much? There is also a benefit in seeing how model performance changes over time. Our current assessments involve static datasets, using the maximum shaking over an interval. It will be useful to explore the performance of the model as time progresses and compare this to that expected as the ratio of observation time to model return period increases. We have a number of sites at which the modeled shaking has been exceeded several times. These will let us explore how well examining model exceedances at many sites compares to examining multiple exceedances at individual sites over time. In addition, we plan to conduct simulation studies to explore the variability in the expected shaking likely to have arisen due to variability in earthquake recurrence, and its consequences for how well the models' predictions should match the observed shaking.

4.10. Data and Resources

The CHIMP dataset is available `https://github.com/salditch/CHIMP`. Historic earthquakes in the 'dataset' are .csv files formatted as LON, LAT, MMI. DYFI earthquakes are .txt files in their original format. The 'repository' contains full felt reports and MMI assignments for earthquakes that were reinterpreted for this study.

Intensity data sources are listed in Table 4.3.

2018 USGS hazard model data are available for download from Science Base at `https://www.sciencebase.gov/catalog/item/5cbf47c4e4b0c3b00664fdef` (last accessed 30 December 2019).

NOAA intensity dataset can be accessed at `https://www.ngdc.noaa.gov/hazard/intintro.shtml`, (last accessed 12 February 2020).

Most of the Abstracts of Earthquakes reports can be accessed online at `http://www.hathitrust.org` (last accessed 12 February 2020). Strong motion data, including recorded peak ground acceleration values, for the 1987 Whittier Narrows earthquake are available from (`https://strongmotioncenter.org`).

Table 4.3. References for earthquake data sources. * indicates events reinterpreted in this study.

CHIMP Label	Date (m/d/y)	Location	Reference
1857	1/9/1857	Fort Tejon*	S.E. Hough, L. Salditch, M.G. Gallahue, 2020. This study.
1868	10/21/1868	Hayward	Boatwright, J., Bundock, H., 2008. Modified Mercalli Intensity Maps for the 1868 Hayward Earthquake Plotted in ShakeMap Format: Spreadsheet of 1868 Intensity Sites. Open-File Report 2008-1121.
1872	3/26/1872	Owens Valley	Hough, S.E., Hutton, K., 2008. Revisiting the 1872 Owens Valley, California. Earthquake Bull. Seismol. Soc. Am. 98 (2), 931-949.
1873	11/23/1873	CA/OR border	Tom Brocher, 2019. Unpublished OFR, personal communication.
1892	2/23/1892	Laguna Salada	Hough, S.E., Elliot, A., 2004. Revisiting the 23 February 1892 Laguna Salada earthquake. Bull. Seismol. Soc. Am. 94 (4), 1571?1578.
1898	4/15/1898	Mendocino*	Toppozada, T.R., Real, C.R., Parke, D.L., 1981. Preparation of isoseismal maps and summaries of reported effects for pre-1900 California earthquakes, Calif. Div. Mines Geol. Open-File Rep. 81-11 SAC, Appendix D.

1906	4/18/1906	San Francisco	Boatwright, J., Bundock, H., 2005. Modified Mercalli Intensity Maps for the 1906 San Francisco Earthquake Plotted in ShakeMap Format. U.S. Geological Survey. Open-File Report 2005-1135. Version 1.0. US Geological Survey, 345 Middlefield Road, MS 977, Menlo Park, CA 94025.
1911	7/1/1911	South Bay*	Toppozada, T.R. 1984. History of Earthquake Damage in Santa Clara County and Comparison of 1911 and 1984 Earthquakes. In Special publications of the California Division of Mines and Geology, Section 1, Earthquake Damage.
1918	4/21/1918	San Jacinto*	Madeleine C. Lucas, 2019. This study.
1925	6/29/1925	Santa Barbara	Hough, S.E., Martin, S.S., 2018. A proposed rupture scenario for the 1925 M_W 6.5 Santa Barbara, California, earthquake, Tectonophysics, Volumes 747?748. P. 211-224. `https://doi.org/10.1016/j.tecto.2018.09.012`.
1927a	11/18/1927	Bishop*	Frank Nuemann, U.S. Coast and Geodetic Survey. Seismological Report July, August, September 1927. Serial number 495. Washington: G.P.O.
1927b	11/4/1927	Lompoc*	Madeleine C. Lucas, 2019. This study.
1932	12/20/1932	Nevada*	Coast and Geodetic Survey, 1984. United States Earthquakes, 1928-1935. Department of the Interior, U.S. Geological Survey. OFR 84-928.
1933	3/10/1933	Long Beach	Hough, S.E. and Graves, R., (2020). The 1933 Long Beach, California, earthquake: Ground motions and rupture scenario, in review, Scientific Reports.
1934	12/31/1934	Colorado River*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region July 1, 1934 to September 30, 1934. Dept. of Commerce, U.S. Coast and Geodetic Survey Field Station 510 Custom House San Francisco, California.

1937	3/25/1937	Buck Ridge*	Coast and Geodetic Survey, United States Earthquakes, 1928-1935. Department of the Interior, U.S. Geological Survey. OFR 84-928.
1940	5/18/1940	Imperial Valley*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region April 1, 1940 to June 30, 1940. MSA-26. Dept. of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1942	10/21/1942	Fish Creek*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region October 1, 1943 to December 31, 1947. MSA-40. Dept. of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1946	3/15/1946	Walker Pass*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region Jan 1, 1946 to March 31, 1946. MSA-49. Dept. of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1947	4/10/1947	Manix*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region April 1, 1947 to June 30, 1947. MSA-54. Dept. of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1948a	12/4/1948	Desert Hot Springs*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region October 1, 1948 to December 31, 1948. MSA-60. Dept. of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.

1948b	12/29/1948	NE California*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region October 1, 1948 to December 31, 1948. MSA-60. Dept. of Commerce, U.S. Coast and Geodetic Survey, Seismological Field Survey 214, Old Mint Building San Francisco 3, California.
1952a	7/21/1952	Kern County	Salditch, L., Hough, S.E., Stein, S., Spencer, B.D., Brooks, E.M., Neely, J.S. and Lucas, M.C. (2018). The 1952 Kern County, California earthquake: A case study of issues in the analysis of historical intensity data for estimation of source parameters. Physics of the Earth and Planetary Interiors, Vol. 283, p. 140-151, `https://doi.org/10.1016/j.pepi.2018.08.007`.
1952b	11/22/1952	Bryson*	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region, 1 July 1951 to 30 September, 1951. MSA-71. 214 Old Mint Building, San Francisco 3, California.
1954a	3/19/1954	San Jacinto	Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region January 1, 1954 to March 31, 1954. MSA-81. Dept. of Commerce, U.S. Coast and Geodetic Survey Seismological Field Station 14 Old Mint Building San Francisco 3, California. MSA-83.
1954b	7/6/1954	Nevada	* Abstracts of Earthquake Reports for the Pacific Coast and The Eastern Mountain Region July 1, 1954 to September 30, 1954. MSA-83. Dept. of Commerce, U.S. Coast and Geodetic Survey Seismological Field Station 14 Old Mint Building San Francisco 3, California. MSA-83.
1966	9/12/1966	Northern CA/Truckee*	von Hake, C.A., Cloud, M.K. 1984. United States Earthquakes, 1966. Department of Interior, U.S. Geological Survey. OFR 84-966.

1968	4/8/1968	Borrego*	Abstracts of Earthquake Reports for the United States April 1, 1968 to June 30, 1968. MSA-137. Dept. of Commerce, Seismological Field Survey, San Francisco, California.
1971	2/9/1971	Sylmar	DYFI, DYFI intensity summary geocoded UTM aggregated (10km spacing).txt. accessed 1/21/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci3347678/executive`
1979	10/15/1979	Imperial Valley	DYFI, DYFI intensity summary geocoded UTM aggregated (1km spacing).txt, accessed 1/21/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci3352060/executive`
1980a	5/25/1980	Mammoth*	Stover, C. W. and von Hake, C.A. United States Earthquakes, 1980. OFR 84-980. Dept. of the Interior, U.S. Geological Survey.
1980b	11/8/1980	Eureka*	Stover, C. W. and von Hake, C.A. United States Earthquakes, 1980. OFR 84-980. Dept. of the Interior, U.S. Geological Survey.
1983	5/2/1983	Coalinga	DYFI, DYFI intensity summary geocoded UTM aggregated (1km spacing).txt, accessed 1/21/19. `https://earthquake.usgs.gov/earthquakes/eventpage/nc1091100/executive`
1984a	4/24/1984	Morgan Hill*	Stover, C.W., 1988. United States earthquakes, 1984. Bulletin 1862, p.26. U.S. G.P.O. `https://doi.org/10.3133/b1862.`
1984b	11/23/1984	Round Valley*	Stover, C.W., 1988. United States earthquakes, 1984. Bulletin 1862, p.47. U.S. G.P.O. `https://doi.org/10.3133/b1862`.
1986a	7/8/1986	N. Palm Springs	DYFI, DYFI intensity summary (city or zip code aggregated).txt, accessed 1/21/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci700917/executive`
1986b	7/21/1986	Chalfant Valley*	Carl W. Stover and Lindie R. Brewer. U.S. Earthquake reports, 1986.

1987a	10/1/1987	Whittier Narrows*	Jim Dewey, 2019. US Earthquakes, 1987, unpublished. Personal communication.
1987b	11/24/1987	Superstition Hills*	Jim Dewey, 2019. US Earthquakes, 1987, unpublished. Personal communication.
1989	10/17/1989	Loma Prieta	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 1/18/19. `https://earthquake.usgs.gov/earthquakes/eventpage/nc216859/executive`
1991	6/28/1991	Sierra Madre	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 3/18/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci2021449/executive`
1992a	4/22/1992	Joshua Tree*	Salditch and Gallahue, 2019. This study.
1992b	4/25/1992	Rio Dell	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 3/18/19. `https://earthquake.usgs.gov/earthquakes/eventpage/nc269151/executive`
1992c	6/28/1992	Landers	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 1/18/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci3031111/executive`
1992d	6/28/1992	Big Bear	DYFI, DYFI intensity summary aggregated by city or zip code.txt, accessed 1/18/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci3031425/executive`
1993	5/17/1993	Big Pine*	Salditch and Gallahue, 2019. This study.
1994	1/17/1994	Northridge	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 1/17/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci3144585/executive`

1999	10/16/1999	Hector Mine	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 1/17/19. https://earthquake.usgs.gov/earthquakes/eventpage/ci9108652/executive
2000	9/3/2000	Northern CA	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 3/18/19. https://earthquake.usgs.gov/earthquakes/eventpage/nc21123384/executive
2001	8/10/2001	Northern CA	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 3/18/19. https://earthquake.usgs.gov/earthquakes/eventpage/nc21188442/executive
2003	12/22/2003	San Simeon	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 1/17/19. https://earthquake.usgs.gov/earthquakes/eventpage/nc21323712/executive
2004	9/28/2004	Shandon	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed3/18/19. https://earthquake.usgs.gov/earthquakes/eventpage/ci14095628/executive
2005a	6/12/2005	Anza	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 3/18/19. https://earthquake.usgs.gov/earthquakes/eventpage/ci14151344/executive
2005b	6/14/2005	Mendocino	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 1/17/19. https://earthquake.usgs.gov/earthquakes/eventpage/usp000dt25/executive
2007	10/30/2007	San Francisco Bay	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 3/18/19. https://earthquake.usgs.gov/earthquakes/eventpage/nc40204628/executive

2008a	2/21/2008	Wells, NV	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 1/17/19. `https://earthquake.usgs.gov/earthquakes/eventpage/nn00234425/executive`
2008b	7/29/2008	Chino Hills	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 3/18/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci14383980/executive`
2009	5/17/2009	Lennox	DYFI, DYFI intensity summary geocoded (10km spacing).txt, accessed 3/18/19. `https://earthquake.usgs.gov/earthquakes/eventpage/ci10410337/executive`
2010	4/4/2010	Baja	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 1/17/19. `https://earthquake.usgs.gov/earthquakes/eventpage/usp000jhr6/executive`
2014	8/24/2014	Napa	DYFI, DYFI intensity summary geocoded (1km spacing).txt, accessed 1/13/2019. `https://earthquake.usgs.gov/earthquakes/eventpage/nc72282711/executive`
2019a	7/4/2019	Ridgecrest	DYFI, `cdi_ geo_1km.txt`, accessed 8/25/2019. `https://earthquake.usgs.gov/earthquakes/eventpage/ci38443183/executive`
2019b	7/6/2019	Ridgecrest	DYFI, `cdi_geo_1km.txt`, accessed 8/25/2019. `https://earthquake.usgs.gov/earthquakes/eventpage/ci38457511/executive`

CHAPTER 5

Map Comparisons: Can metrics indicate if one PSHA model is better than another?

5.1. Introduction

Because PSHA maps depend on assumptions about future earthquakes and how hazard is characterized, plausible alternative input parameters for an area can yield quite different results (Stein et al., 2018). For example, the precise locations of historical earthquakes are not well known and may not fully show the locations of future earthquakes. Similarly, because the area of a fault that ruptured controls the resulting earthquake's magnitude and distribution of shaking intensity, there are substantial uncertainties regarding past earthquake source parameters and shaking. Hence, there are significant uncertainties in the maps.

In the hazard curve calculations (Figure 5.1), this is accounted for via complex interplay between the return periods, the earthquake rates and the probabilities of exceedance (of some ground motion), and their respective epistemic and aleatory uncertainties. The earthquake rates are a function of the Seismic Source Characterization (SSC) which also determines the sizes and locations of earthquakes in the model. The probabilities of exceedance are a function of the GMMs. The aleatory uncertainty is the effect of random variability of the system, as opposed to the epistemic uncertainty which represents the limitations of our knowledge of the system.

For shorter return periods, like the 475 yr models investigated in this chapter, the rate of earthquakes tends to control the hazard. In order to create a longer return period model, a low probability of exceeding the ground motion is set. Hence at longer return periods, the aleatory variability of the GMMs tends to control the hazard, because it influences that probability of exceedance.

In high seismicity areas, like Nepal, we have better constraints on the rates of earthquakes because we have more data. Because the rate is high and the probability of exceedance low, the result is a steep slope of the hazard curve. With a steep slope, a change in the rates of earthquakes (corresponding to a shift in y-axis in figure 5.1) does not make much of a change in the ground motion (corresponding to the x-axis in figure 5.1). In contrast, a small shift in the x-axis value leads to a large change in the y-axis value (the rate). In this case, the epistemic uncertainty in the hazard is controlled by the epistemic uncertainty in the GMM.

For low-to-moderate seismicity areas, like France, there is larger uncertainty in the rates of earthquakes and the slope of the resulting hazard curve is flatter. In this case, a change in the earthquake rate can have a similar effect as a change in the GMM. So, the epistemic uncertainty in the hazard can at times be controlled by the SSC model (rates, sizes, and locations of earthquakes).

These are important considerations when creating and evaluating hazard maps, and leads to the question; can our metrics indicate if one map is better than another? It would be useful to know if changes to a model constituted actual improvement.

I explore this issue using two case studies. First, for Nepal, I consider two maps that differ in basic assumptions and parameter choices for GMMs, seismic sources, and site effects. Both maps characterize the hazard level using the mean hazard, a weighted average of all probabilities of exceedance and earthquake rate combinations. However, there is reason to believe that this representation may not be the best approach. The median hazard (50th percentile in Figure 5.1), typically smaller than the mean, may be a better characterization for assessing a map's performance (Abrahamson and Bommer,

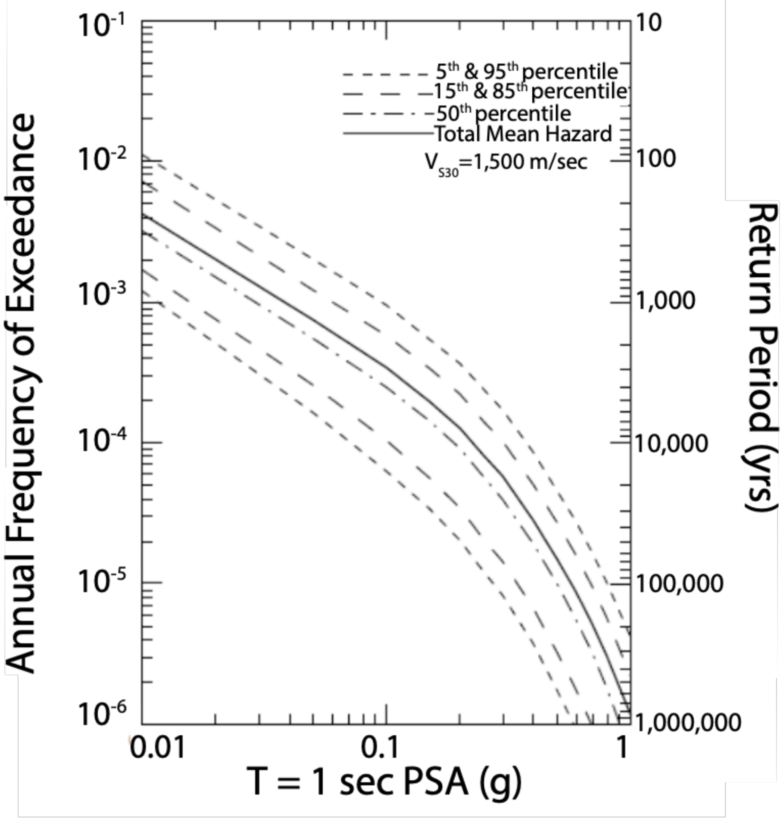

Figure 5.1. Hazard curve at a particular site with large uncertainties. The quantiles represent different combinations of the earthquake rates and the GMMs. Mean hazard, traditionally used in maps, is the weighted average of combinations of GMMs and earthquake rates. Note here that the mean is higher than median (50th percentile). Difference between mean and median is often smaller in higher seismicity regions.

2005). This is because the median, the central estimate, is less affected by outliers and long tails on the probability distribution which cause the mean to predict higher hazard because of rare events. While the median may prove most useful for assessing map performance, there is still value in using the mean hazard, particularly for engineering concerns, because the mean penalizes uncertainty. This suggests that the intended use of a hazard map should be considered in performance evaluation, and that what is best for testing and what is best for engineering applications may not always be the same thing.

Because median hazard maps for Nepal are not readily available, I turn to the second case study, France, a region of more moderate seismicity, to assess the difference between mean and median hazard models. The results indicate an improvement in model performance going from mean to median hazard. For any given map, it will likely be a combination of better hazard characterization and parameter assumptions that improves a models' performance.

5.2. Nepal

Nepal is an area of high hazard because the collision between the Eurasian and Indian plates can create M 9+ earthquakes. Nepal is located directly atop the Main Himalayan Thrust (MHT), the primary decollement (basal detachment fault) locking the interface between the Indian and Eurasian plates, which are converging at 15 ± 2 mm/yr (Wesnousky, 2020). $M_W > 8.5$ earthquakes occurred in the past (most recently in 1200-1400 CE) and may recur in the near future based on recurrence estimates of 500-1,000 yrs (Wesnousky, 2020). Such earthquakes would be catastrophic, possibly resulting in upwards of 300,000 casualties in Nepal and India (Wyss et al., 2018). However, the most recent damaging event in Nepal, the 2015 Gorkha M_W 7.8 earthquake that killed 9,000 people, produced less severe ground motions than expected (Martin et al, 2015). Understanding the seismic hazard of the Himalayas is therefore extremely important to mitigate the consequences of future great earthquakes.

Some hazard maps proposed for Nepal (e.g. Ram and Guoxin, 2013; Chaulagain et al., 2015) show lower hazard for central Nepal. Other maps (e.g. Stevens et al. 2018 and GEM 2020 (a.k.a. Nath and Thingbaijam, 2012) to a lesser degree) show a more homogeneous or constant hazard along the strike of the MHT front. The latter is consistent with the fact that GPS data show no significant variation in plate coupling between areas of recent large earthquakes (Avouac et al., 2015). Hence with present knowledge, the entire zone could be regarded as equally hazardous and perhaps vulnerable to much larger earthquakes than those currently known, with long recurrence times (Stein et al., 2018). However, there are indications that changes in dip along strike of the MHT controlled the rupture length of the 2015 Gorkha earthquake, hence our understanding of the MHT remains uncertain (Bai et al., 2019). A challenge for hazard maps of Nepal is that although there is a long record of earthquakes going back to $\sim$ 800 CE (Bilham, 2019), there are large uncertainties regarding their magnitudes and locations of the the ruptures associated with these events.

Nepal has a long record of earthquake intensity data, going back to 1636 CE (Martin and Szeliga, 2010), and more detailed data for the largest recent earthquakes, the 2015 Gorkha M_W 7.8 and the 1833 M_W 7.5 events (Martin et al., 2015) (figure 5.2). However, despite the long record, the spatial coverage of the data is not consistent along strike of the MHT. The region surrounding Kathmandu is densely populated in Nepal and most intensity observations are clustered there. This makes it extremely difficult to evaluate the differences in hazard maps in low population and uninhabited places, where there is most disagreement between maps.

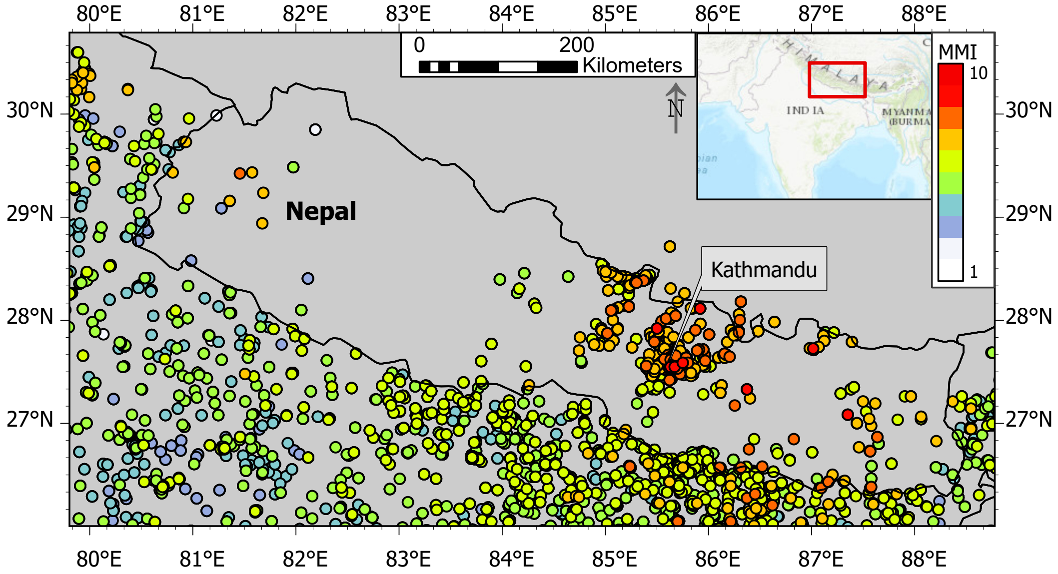

Figure 5.2. Nepal historic intensity data, in Modified Mercalli Intensity (MMI) units. Data are from Martin and Szeliga (2010) and Martin et al. (2015). Population and observations are clustered around the capital Kathmandu.

To evaluate the maps, I grid the intensity data into 10 by 10 km cells, and find the maximum in each cell (figure 5.3). I compare this maximum shaking map to two hazard maps, those of GEM 2020 and Stevens et al. (2018), hereafter GEM2020 and Stevens18 (figure 5.4). Both hazard maps show the estimated level of shaking with a 10% probability of exceedance in 50 yrs time corresponding to a return period of 475 yrs. These maps differ due to alternative parameter choices, as illustrated by subtracting the GEM2020 map from the Stevens18 map (figure 5.5).

The GEM2020 map (Nath and Thingbaijam, 2012) makes the following assumptions and parameter choices:

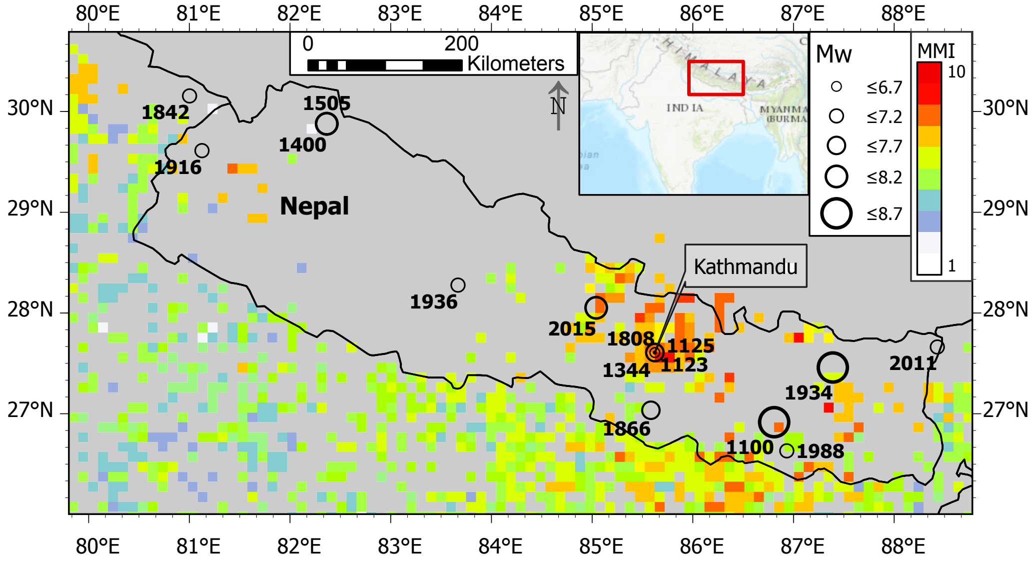

Figure 5.3. Data from Figure 1 gridded in 10 x 10 km cells to show maximum observed shaking in each cell. Locations and magnitudes of historic earthquakes g back to 800 CE are plotted as open circles, with years of occurrence labelled (Bilham, 2019).

(1) It uses four different tectonic domain-specific Ground Motion Prediction Equations (GMPEs): active shallow crust, subduction interface/intraslab, intraplate, and intraplate margin. None of these GMPEs are region-specific, that is, they were not made specially for use in Nepal. It is common practice to use globally derived GMPEs, but it is unknown whether these are appropriate (Delavaud et al., 2012; Barani et al., 2017).

(2) Areal seismogenic sources are divided into four hypocentral depth ranges (< 25 km, 25-70 km, 70-180 km, and 180-300 km).

(3) The MHT is a segmented source.

(4) Uniform site effects from assumption of non-varying Vs30. That is, the local geology and soil rigidity, which can amplify or deamplify ground motions, are ignored in this model.

In contrast, the Stevens18 map:

(1) Uses two globally derived tectonic-domain specific GMPEs (subduction and active shallow crust). That is, the GMPEs were created by combining data from a global selection of a specific type of tectonic domain.

(2) Uses a mix of fault, area, and background seismicity sources. The MHT is modeled as a planar fault surface because it is well defined; the northern grabens

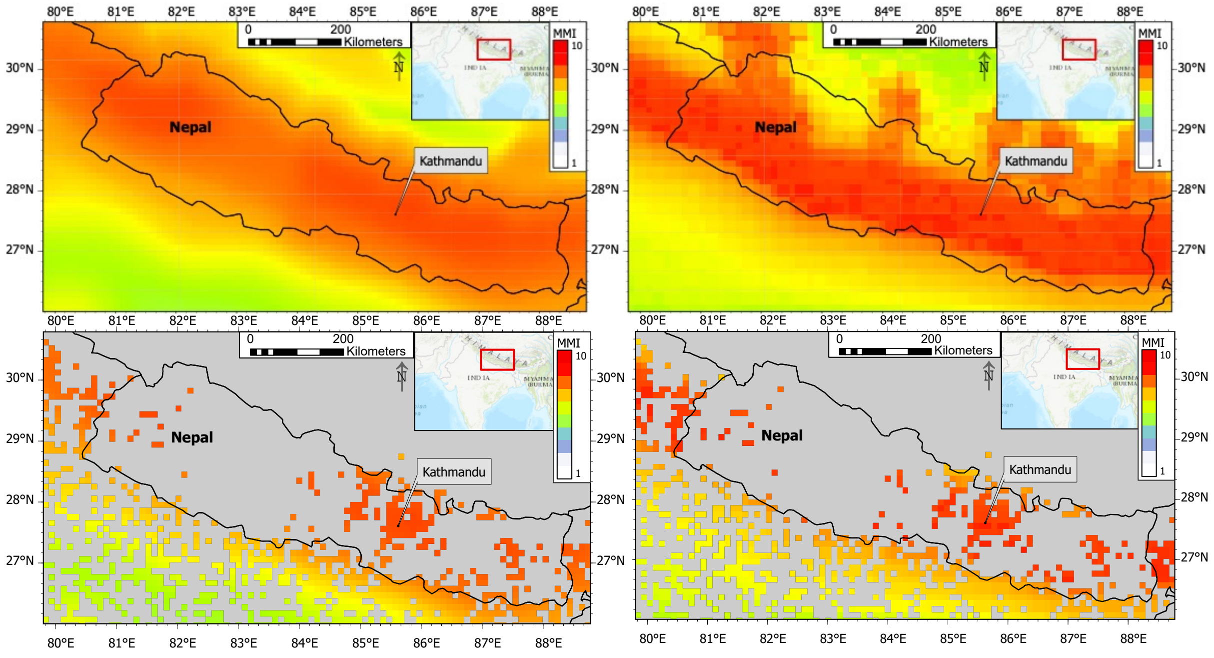

Figure 5.4. Top left: GEM 2020 hazard map, showing shaking levels expected to be exceeded at 10% of sites in the next 50 yrs. Top right: Stevens et al., 2018 10% in 50 yr map. Bottom left: GEM 2020 map showing grid cells only where historical observations exist. Bottom right: Stevens2018 map showing grid cells only where historical observations exist.

are modeled as areas where earthquakes can occur anywhere within their bounds because the faults are not well identified, but deformation is known to occur; background seismicity allows earthquakes to occur anywhere on the map regardless of fault or area source.

(3) Treats the MHT as a single source that could produce M 9+ earthquakes anywhere along its length.

(4) Includes varying site effects via estimates of Vs30 inferred from topographic slope. The slope of the topography is assumed to be related to the thickness of sediments that have accumulated. Hence flatter slopes indicate more sediments, which can amplify ground motions.

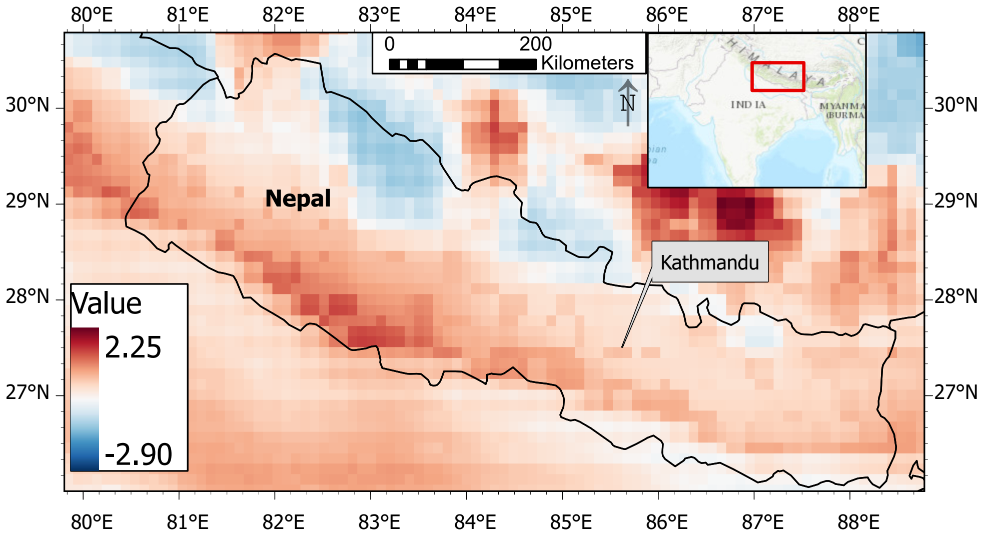

Figure 5.5. Subtracting Stevens18 from GEM2020 reveals the differences (in MMI units) between the two maps.

I converted the hazard map prediction from peak ground acceleration (PGA) to MMI via equation 3 of Worden et al. (2012), the median Ground Motion Intensity Conversion Equation (GMICE), excluding magnitude and distance corrections as well as aleatory variability of MMI. It is possible that the Worden method, which was derived using data from California, is not entirely appropriate for this setting, and future work may improve the conversion of PGA to MMI in Nepal using a regionally derived relationship, which may affect the results of the performance evaluations.

To elucidate the subtle differences between the observations and the map predictions, I examined residual plots, defined here as (observed - predicted) (figure 5.6). The residual maps show that the maximum observed shaking is generally lower than predicted by the hazard maps, indicated by the negative values of the residuals. The exceedances (red areas on the residual map) tend to cluster in the Kathmandu area, as well as to the north and south there. The histograms of residuals (figure 5.7) show that the GEM2020 map is on average closer to the observed values than the Stevens2018 map by 0.2 MMI units, well within the limits of the uncertainty of MMI. Consequently, the maps are practically indistinguishable in areas where observations exist. The long negative tails on the residual distributions indicate that the maps are more likely to overpredict shaking when compared to observations. The performance metrics are likewise quite similar (figure 5.8). Based on the ratio of the observation time (379 yrs) to the return period of the map (475 yrs), 55% of sites should have shaking exceeding predictions The actual percentage is 7% and 5%, a factor of 10 lower than expected, giving an M0 =0.48 and 0.5, respectively, for the GEM2020 and Stevens2018 maps.

From figure 5.8, the M0 scores show that the GEM2020 map is on average marginally closer to the observed values than the Stevens2018 map, indicated by the M0 being 0.0177 units lower for GEM2020, or a difference of 4%. These results indicate that current hazard maps are consistent with one another in regions where observations currently exist. That is, despite the differences in GMMs, seismic sources, and site effects, the assessments do not indicate that there is a substantial difference between the resulting maps. Both maps overpredict ground shaking compared to historical observations by a factor of ∼ 10.

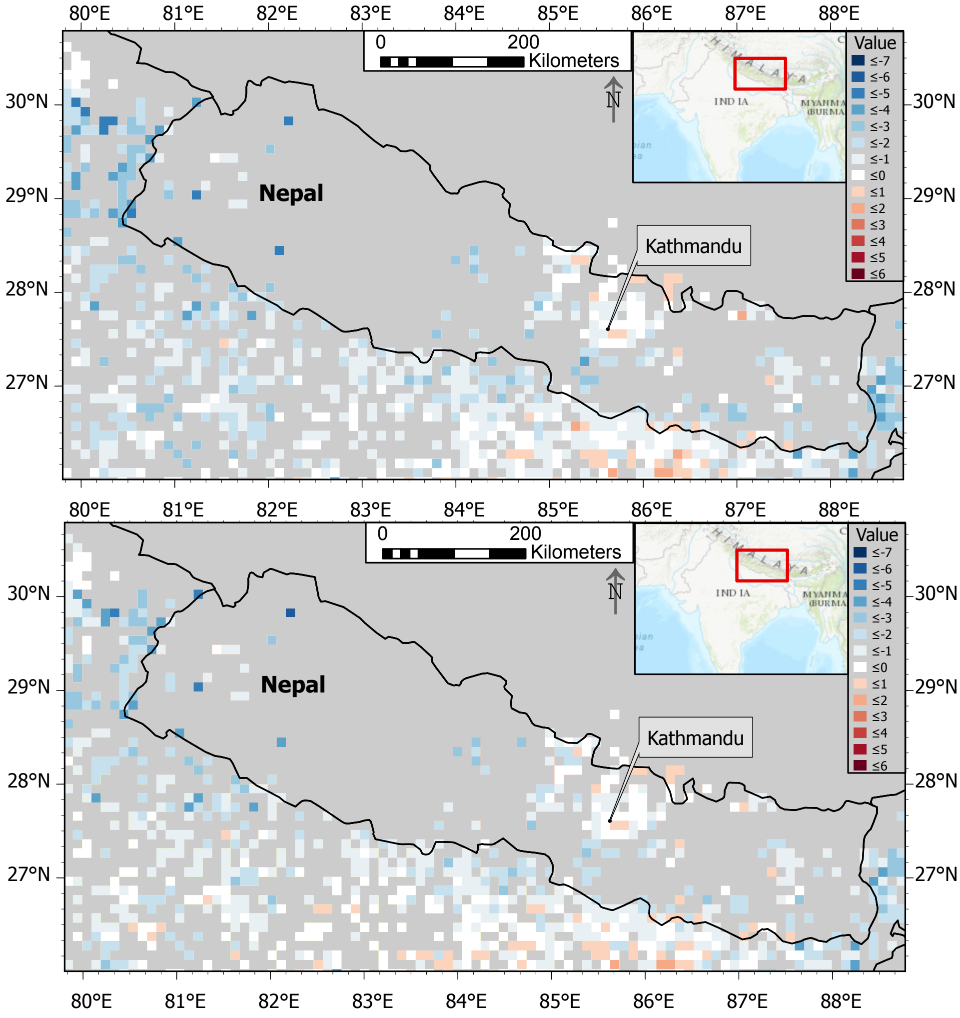

Figure 5.6. Residuals (observed - predicted) for GEM 2020 (bottom), and Stevens2018 (top). Blue and white colors indicate observed shaking lower than predicted by the hazard map. Red colors indicate observed shaking higher than predicted by the map.

To investigate the effect of bias from a too-high map or too-low data, and considering that historical MMI has an uncertainty of ± 1, I explore if a bulk shift of MMI values would improve the M0 metric (Brooks et al., 2016). Figure 5.9 shows the effect on the M0 performance metric of adding or subtracting the same amount from all observed intensities, using the Nepal hazard map from GEM2020. Increasing the observed intensities by

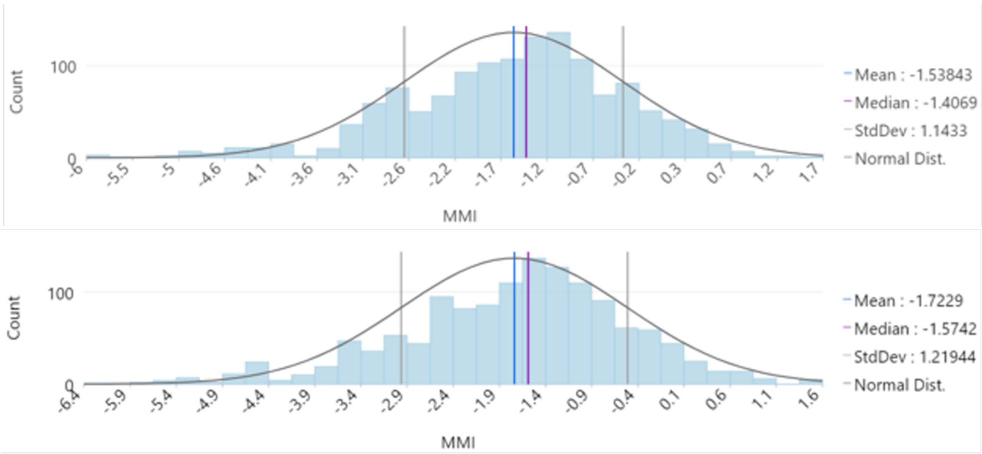

Figure 5.7. Histograms of residuals (observed - predicted) for GEM2020 (top) and Stevens2018 (bottom).

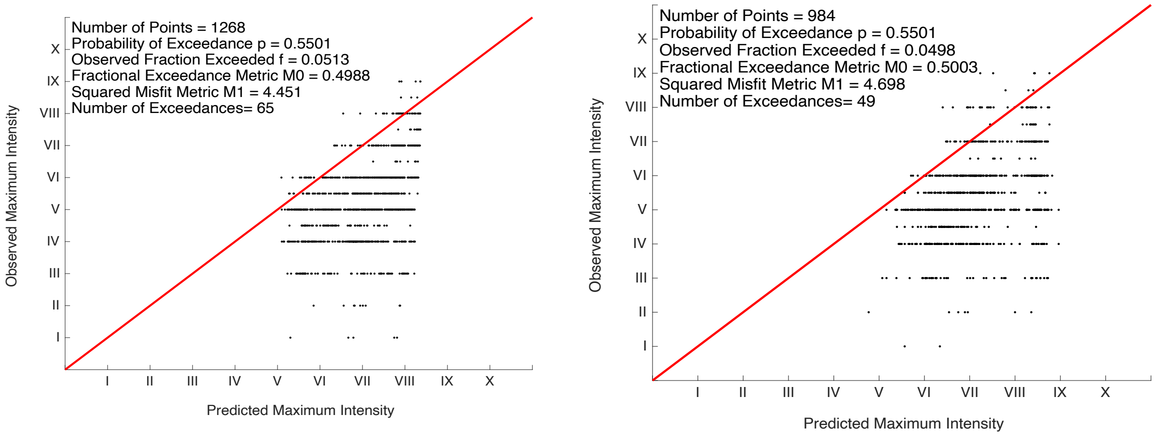

Figure 5.8. Performance metrics for GEM2020 (left) and Stevens2018 (right).

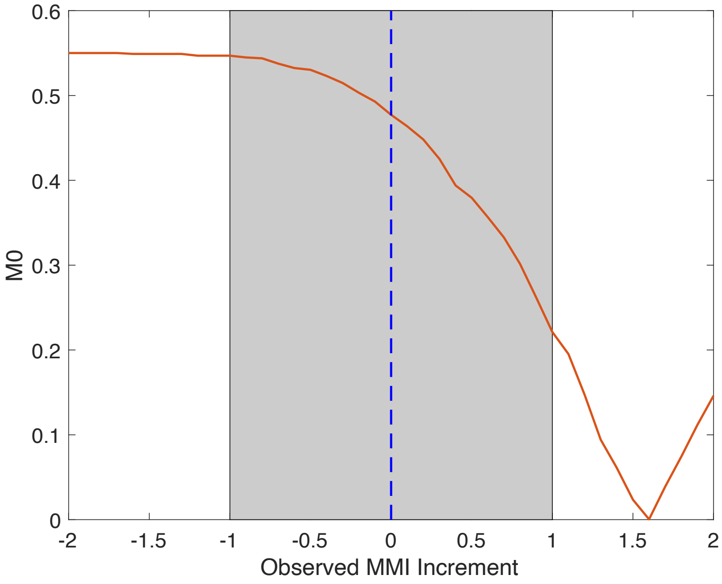

Figure 5.9. Effect on the M0 performance metric of adding or subtracting the same amount from all observed intensities, using the Nepal hazard map from GEM2020. A bulk shift of all observed intensities of +1.5 MMI would give a perfect M0 score of zero, however, this lies outside the uncertainty range (grey shaded area) of MMI assignments.

+ 1.5 MMI would give a perfect M0 score of zero. However, such a large shift exceeds the expected uncertainty range (grey shaded area) of MMI assignments.

5.3. France

France is an intraplate environment not typically associated with damaging earthquakes, but it is flanked by the Pyrenees and the Alps which contribute considerable deformation. At present, only 475 yr return period maps (10% in 50 yrs) are available. However, they are provided for different levels of hazard characterization including the

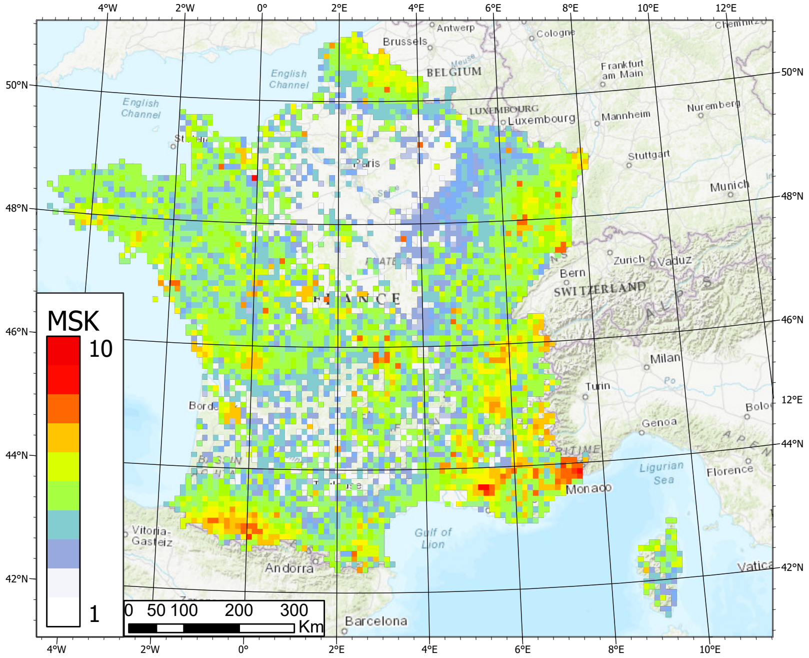

Figure 5.10. Map of maximum observed intensity in France. MSK is the Medvedev-Sponheuer-Karnik intensity scale.

mean and the median (Stephane Drouet, Personal Communication). This gives an opportunity to test the hypothesis that the median may predict map performance better than the mean (Abrahamson and Bommer, 2005).

The SisFrance seismic intensity database contains over 60,000 IDPs (data provided by and used with permission from EDF, the French electricity utility) for earthquakes back to 217 BCE, although the dataset is only complete starting around 1500 CE (Sisfrance, https://sisfrance.irsn.fr/; Drouet et al., 2020). The most recent event in the catalog is from 2007, making the time covered by the reliable data equal to 507 yrs. The IDPs are given in the MSK-64 scale (Medvedev et al., 1967), and are associated with M_W 3.3 - 6.3 earthquakes. Figure 5.10 shows the maximum observed intensity. Musson et al., 2006 determined that although there are subtle variations between different seismic

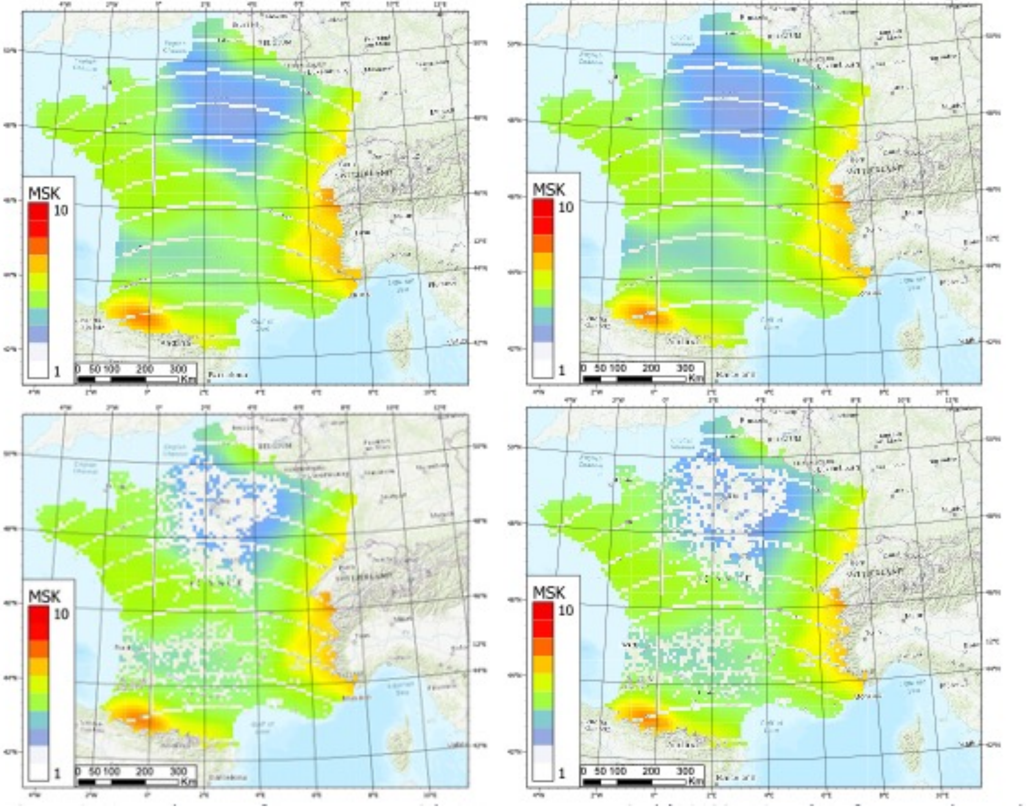

Figure 5.11. Hazard maps of France with 475 yr return period (10% in 50 yrs). Left: mean hazard. Right: median hazard. Top row: full maps. Bottom: maps show values only where observations exist in figure 5.9.

intensity scales, their differences are not so large as to necessitate conversions. Hence any 12-step or 10-step (the common practice today) intensity scales should be approximately equivalent. Thus because no conversion between PGA and MSK is readily available, I use the GMICE of Gomez-Capera et al. (2020), which was developed for conversion of Mercalli-Cancani-Seiburg (MCS) intensity data to MMI.

I compare these observations to the mean and median 475 yr return period hazard maps (figure 5.11). Figure 5.12 shows the difference between the maps, by subtracting the median map from the mean map. The largest difference at any point is about half

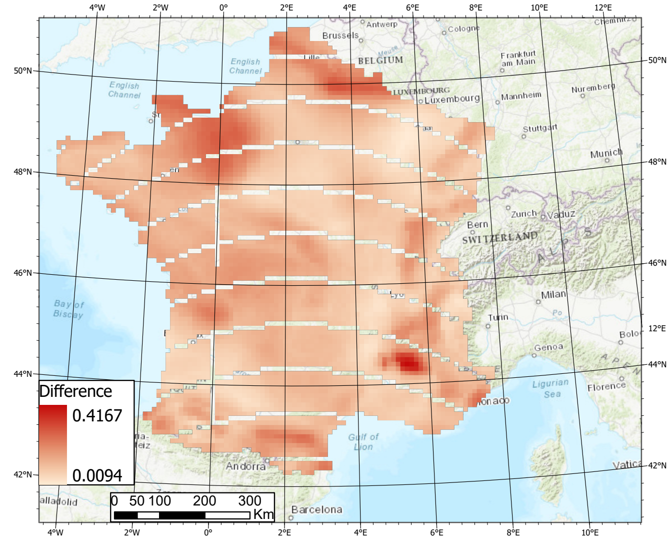

Figure 5.12. Mean hazard minus median hazard. Differences between the maps are relatively small.

an intensity unit. The residual maps (figure 5.13) reveal exceedances (red areas) where observed shaking was larger than predicted. White and blue parts of the map indicate where the map over-predicted shaking. Histograms of the residual maps (figure 5.14) show that the distributions for both maps are skewed slightly negative, indicating both maps have a net over-prediction. Based on the time covered by the observations (507 yrs) and the return period of the map (475 yrs), we expect 66% of sites to have observations exceeding predictions. However, only 34% and 39% of sites exceed predictions for the mean and median maps, respectively (figure 5.15). While both maps for France over-predict, there is a 14% difference in their M0 metrics (0.3 for the mean and 0.26 for the median).

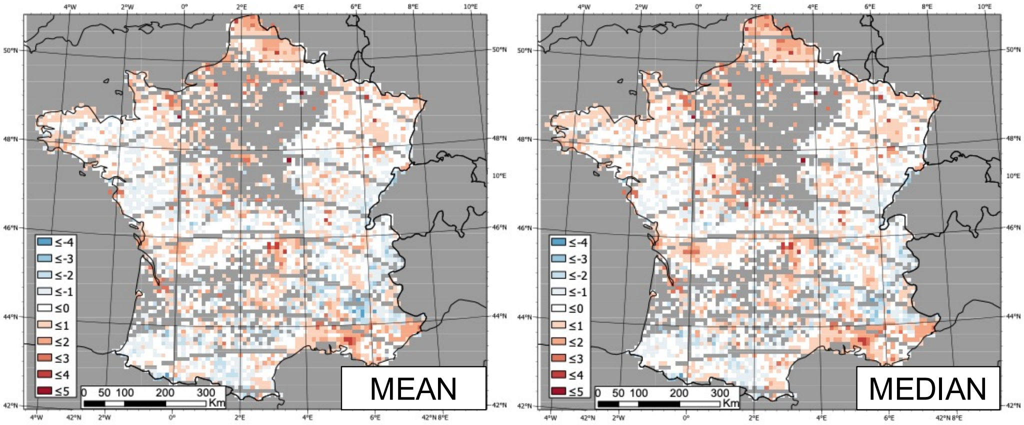

Figure 5.13. Residuals (observed - predicted) for mean and median hazard maps of France (10% in 50 yrs). Blue colors indicate observed shaking lower than predicted by the map. Red colors indicate observed shaking higher than predicted by the map.

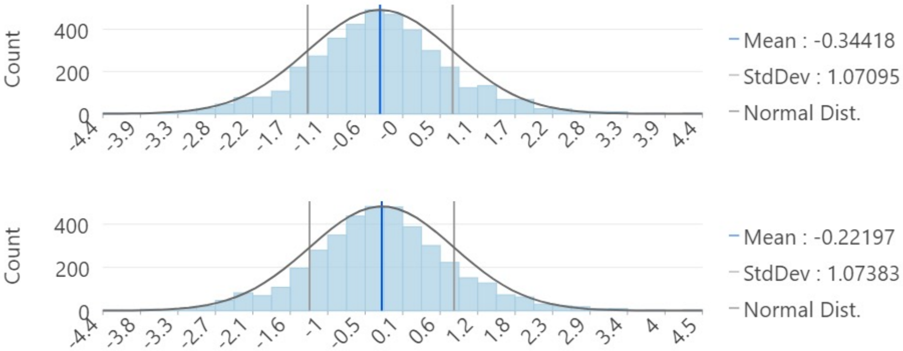

Figure 5.14. Top: Mean hazard residual histogram. Bottom: Median hazard residual histogram

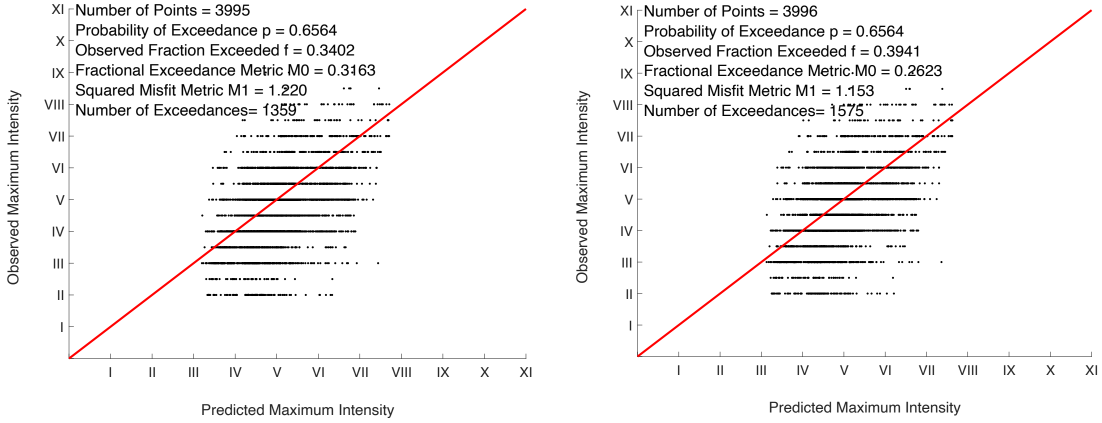

Figure 5.15. Performance metrics for French 475 yr hazard maps. Mean hazard (left) and median hazard (right).

The results that in France $f/p \sim 0.5$, can be better understood by exploring the implications of the uncertainties in Figure 1, the hazard curve for a low-to-moderate seismicity region. In this hazard curve, the total uncertainty in the model is about 10, based on the distance between the 5th and 95th percentiles. Considering the mean predicted shaking, p, and its ratio with the observed, f, being $p/f \sim 2$, it can be seen from Figure 5.1 that a reduction from the mean hazard, p, of a factor of 2 is about at the 30th fractile for any given ground motion level. Hence, 30% of the models would produce hazard that is at or below the observed rate, and conversely 70% would produce hazard that is above the observed rate. This line of analysis is the current focus of the International Atomic Energy Agency (IAEA), which focuses on the very long-term safety (tens to hundreds of thousands of years) of nuclear reactors and waste repositories. Hence it is important to understand how hazard maps with longer return periods (lower probabilities of exceedance) perform.

5.4. Discussion

Using metrics to compare alternative hazard maps is a useful way to explore the component of a hazard model that is introducing differences. The two maps I looked at in Nepal both showed the mean hazard for a 475 yr return period but differed in basic geologic assumptions. Adding background seismicity, treating the MHT as a single source, and varying Vs30 in the Stevens2018 map did not improve map performance over the GEM2020 map. M0 for Stevens2018 was a marginal 4% smaller than GEM2020. Thus these differences in model assumptions did not produce appreciably different maps.

In France, I compared two different characterizations of the hazard - the mean and the median - for a 475 yr hazard model. In this low-to-moderate-seismicity region, the difference between the M0 metrics of the mean and median models is substantial, with the median (M0=0.26) performing 14% better than the mean (M0=0.3). This suggests that hazard characterization may have more impact than SSCs and GMMs.

In future work I will investigate these longer return period maps of France to see if it holds that they are better assessed using median hazard. If the median continues to improve map performance, then this may be an important implication for hazard assessment. Furthermore, I will identify which GMMs or assumptions used in in the logic trees are the primary sources of misfits. In this way, I will continue to use metrics to refine models and improve map performance.

CHAPTER 6

Conclusion

The M0 fractional exceedance performance metric has been applied to hazard maps across the globe (Stein et al., 2015; Brooks et al., 2018; 2019; Salditch et al., 2020). By this metric, nearly all of these maps over-predict shaking when compared to maximum observed historical intensities. M0 can be visualized as the distance between the predicted (black line) and observed (triangles, circles, diamonds) fractional exceedances, where the predicted line is the exponential CDF, $p = 1 - exp(-t/T)$, t is the observation time in years, and T is the map's return period in years (Figure 6.1). Numerical experiments (Vanneste et al., 2018) indicate for ideal (unbiased) maps, observed shaking is equally likely to be above and below the prediction. However, none of the maps studied under-predict shaking, i.e. none of the observed fractional exceedances were larger than model predictions.

These results indicate a possible systematic bias in PSHA models. Analysis of French hazard maps implies that the median hazard may better indicate map performance than the mean hazard. Because all other maps studied to date characterize the hazard using the mean, this difference could be an important source of bias. If $p = f$ (M0 = 0, a perfect score) for the median hazard, then the GMM and seismic source models would be "centered" on the MMI data. In practice, weights are not adjusted on the models to give a hazard centered on the MMI data because of the question of whether the available MMI data adequately represent the long-term shaking. It will still be uesful to look at which models predict shaking much higher or much lower than that observed and determine if there are good seismological reasons to reject (or down-weight) those models.

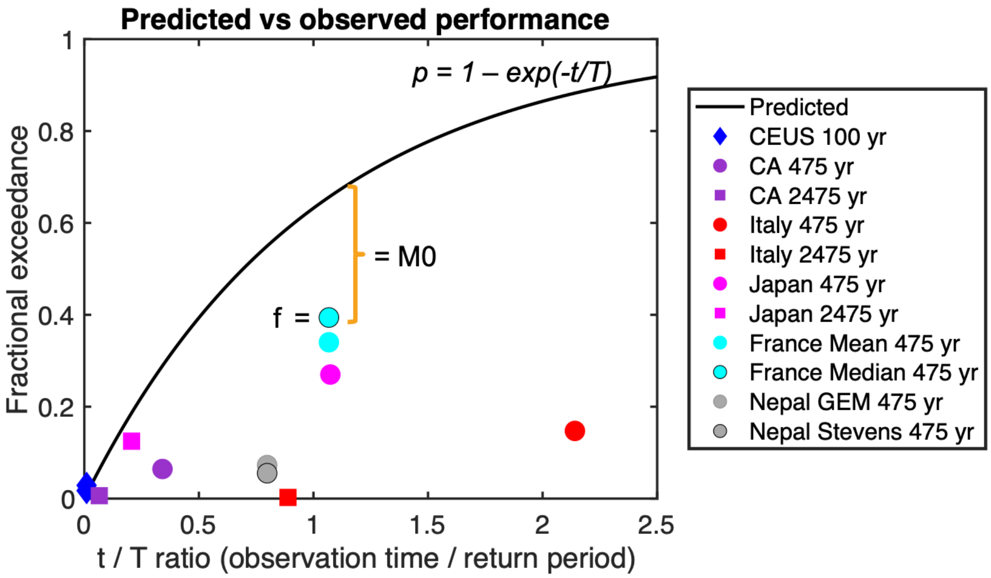

Figure 6.1. PSHA map performance. Predictions follow the formula $p = 1 - exp(-t/T)$ where t is the observation time, and T is the return period of the map. Symbols represent the observed fractional exceedances, f. The M0 metric can be visualized as the distance between f and p. CEUS represents the 2016 and 2017 USGS models for natural and induced seismicity. CA represents the 2018 USGS seismic hazard map for California compared with CHIMP data. Numbers following regions in legend are map return periods.

There are also issues with the completeness and length of the observed catalogs and the influence of seismicity rate changes. The ratio of t/T is a crucial parameter for the M0 metric. Most map-catalog pairs studied here have $t/T \leq 1$, which may be too short to reliably capture the seismicity rate of a region (Page and Felzer, 2015; Vanneste et al., 2018). However, if there were a map-catalog pair with, for example, $t/T \sim 3$, p would be 95%. That is, the expected fractional exceedance at that t/T ratio is 0.95, so exceedances are expected in all but 5% of locations. The usefulness of a map with 95% predicted exceedances is questionable.

6.1. Smallest M0 Score

The best performing map is the 2016 USGS 1-yr model (100 yr return period) for induced and natural seismicity in the Central and Eastern United States (CEUS), with an M0=0.0073. This and one made in 2017 (M0=0.0188) are the top performing maps studied to date. These maps differ in many ways from the others investigated in this thesis. The most important difference is that these maps focus on induced seismicity associated with injecting wastewater from oil and gas production deep into the ground (Petersen et al., 2016). These maps were made several years after the beginning of the induced seismicity increase that plagued the otherwise aseismic CEUS during the hydraulic fracturing boom of the early 2000s (Ellsworth, 2013). This allowed map makers to include parameters like well injection volumes into the model.

The CEUS maps are the shortest-term (only one year) maps, so it was feasible to wait that time and compare the model with observations made after the model?s creation. These 1-yr maps are therefore the only forecasts that I considered, with the others being hindcasts due to their 50 yr time window.

6.2. Largest M0 Score

The Italian maps studied here have the highest M0 scores, or the poorest fit to model predictions (Stein et al., 2015). Italy is geologically active due to the complex tectonics of the Mediterranean region. The geological backbone of Italy is the Apennine mountains, with normal faulting due to divergence between the Adria microplate and Eurasia (Anderson and Jackson, 1987; Calais et al., 2002; Stein and Sella, 2005). Western extension is the result of back-arc spreading of the migrating subduction zone in the Adriatic and

Ionian seas to the east (Loreto et al., 2021). To the north, the Alps formed as the result of the northward collision of the African plate into the European plate, and Adria is converging on Eurasia there today. The result is a seismically and volcanically active country with a history of devastating earthquakes and eruptions.

Italy has a long historical record of written accounts describing the effects of earthquakes as well as an abundance of sites from which earthquake effects can be deciphered from the damage to historic architecture. DBMI15v.2 is an Italian database of standardized macroseismic intensities covering the period 1000 - 2018 C.E. (Locati et al., 2019). I compared these data to the 475 yr return period European Seismic Hazard Model (ESHM13) (Woessner et al., 2015). Even with an observation time more than twice the length of the return period of the map, it performs the least well of all models studied to date with an M0 = 0.7.

6.3. Discussion

It is expected that the larger the ratio of the observation time, t, to the return period, T, the closer the observed fractional exceedance f will be to the model prediction p (Vanneste et al., 2018). However, in our studies this effect is not observed (Figure 6.1). Observed exceedances do not systematically get closer to predictions with increasing t/T. In fact, the best performing models (CEUS and the 2475 yr models of California and Japan) have the smallest t/T ratio. Map performance does not appear to have a straightforward relationship with return period, as the best performing maps have the shortest (100 yr) and longest (2475 yr) return periods. However, when a 475 yr map and a 2475 yr map are both available for the same region, the 475 yr always performs less well. This is likely due to the large influence of the aleatory uncertainty at shorter return periods (Abrahamson and Bommer, 2005). The worst performing maps tend to be 475 yr models. More map-catalog pairs with larger t/T ratios are needed to make a robust conclusion about any kind of systematic bias.

There is another reason to pursue catalogs and maps with large t/T ratios. Page and Felzer (2015) showed that catalogs need to be many times the length of the return period of the largest magnitude event to accurately estimate the seismicity of a region. It should be noted that in our calculations T is the return period of the map, which is not the same as the return period or interevent time of the largest magnitude event. T is simply the inverse of the annual probability of exceedance, or the average number of years it takes to get an exceedance (USGS Earthquake Hazards: `https://www.usgs.gov/natural-hazards/earthquake-hazards/science/earthquake-hazards-201-technical-qa`). The average interevent time of the largest magnitude earthquake may be longer than the return period of the map. That the observed catalogs are not long enough to be representative of the average activity level is an important source of uncertainty in the results presented here. The inherent variability of earthquake recurrence times, part of the aleatory uncertainty, and the propensity of large earthquakes to cluster in time suggests that a given observation period may not be indicative of the overall seismicity of a region (Salditch et al., 2020). Most of the map-catalog pairs studied here have $t/T \leq 1$, so the maximum observed intensities during that time may not be representative and likely underestimate the long-term average (Naylor et al., 2009).

Some of the issues in PSHA addressed by the performance metrics indicate that improvements are feasible. My next steps are to test individual branches of the logic trees, corresponding to specific GMMs and seismic source models, rather than the mean or median of the logic tree (a weighted combination of branches and hence multiple GMMs and source models). In this way I can identify the logic tree branches that lead to largest discrepancies between the observed and predicted shaking. This will provide information needed to update hazard models and allow us to investigate whether misfits are driven by inappropriate GMMs or inappropriate seismic sources models. I will evaluate the assumptions underlying the branches of the logic tree that are inconsistent and investigate if there is a good reason why these branches are inconsistent with the MMI data. Models that are not consistent with observations will be identified and inappropriate aspects of the model will be examined.

References

[1] Abe, K., 1981. Magnitudes of large shallow earthquakes from 1904 to 1980, Phys. Earth and Plan. Int. 27(1), 72-92.

[2] Abe, K., 1994. Instrumental magnitudes of historical earthquakes, 1892 to 1898, Bull. Seism. Soc. Am. 84, 415-425.

[3] Abrahamson, N. A. and J. J. Bommer, 2005. Probability and uncertainty in seismic hazard analysis, Earthq. Spectra 21(2), 603-607. doi: 10.1193/1.1899158.

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

[5] Adhikari, S.R., Baysal, G., Dixit, A., Martin, S.S., Landes, M., Bossu, R., Hough, S.E., 2017. Towards a unified near-field intensity map of the 2015 Gorkha, Nepal, earthquake, Earthquake Spectra, 33:S1, S21-S34.

[6] Agnon, A., 2014. Pre-instrumental earthquakes along the Dead Sea Rift, In: Z. Garfunkel et al. (eds.), Dead Sea transform fault system: Reviews, Modern Approaches in Solid Earth Sciences 6.

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

[8] Albarello, D., and V. D'Amico, 2008. Testing probabilistic seismic hazard estimates by comparison with observations: An example in Italy, Geophys. J. Int. 175(3), 1088-1094.

[9] Ambraseys, N.N., 1971. Value of historical records of earthquakes. Nature 232, 375-379.

[10] Ambraseys, N.N., et al., 1983. Notes on historical seismicity. Bull. Seismol. Soc. Am. 73(6), 1917-1920.

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

[12] Anderson H.A., Jackson J.A., 1987 Active tectonics of the Adriatic region. Geophys. J. Roy. Astron. Soc. 91, 937-983.

[13] Atkinson, G.M., Wald, D.J., 2007. "Did You Feel It?" intensity data: A surprisingly good measure of earthquake ground motion. Seismol. Res. Lett. 78(3), 362-368.

[14] Atkinson, G.M., C.B. Worden, and D.J. Wald, 2014. Intensity prediction equations for North America, Bull. Seism. Soc. Am., 104(6), 3084-3093.

[15] Avouac, J., Meng, L., Wei, S., Wang, T., Ampuero, J., 2015. Lower edge of locked Main Himalayan Thrust unzipped by the 2015 Gorkha earthquake. Nature Geoscience. doi: 10.1038/NGEO2518.

[16] Bai, L., Klemperer, S. L., Mori, J., Karplus, M. S., Ding, L., Liu, H., Li, G., Song, B., Dhakal, S. 2019. Lateral variation of the Main Himalayan Thrust controls the rupture length of the 2015 Gorkhaearthquake in Nepal. Sci. Adv. (5).

[17] Bakun, W.H., 2006. Estimating locations and magnitudes of earthquakes in Southern California. Bull. Seismol. Soc. Am. 96(4A), 1278-1295.

[18] Bakun, W.H., Wentworth, C.M., 1997. Estimating earthquake location and magnitude from seismic intensity data. Bull. Seismol. Soc. Am. 87(6), 1502-1521.

[19] Barani, S., Albarello, D., Massa, M., Spallarossa, D., 2017. Influence of twenty years of research on Ground-Motion Prediction Equations on probabilistic seismic hazard in Italy. Bulletin of the Seismological Society of America, 107(1), 240-255. doi: 10.1785/0120150276.

[20] Bawden, G.W., 2001. Source parameters for the 1952 Kern County earthquake, California: A joint inversion of leveling and triangulation observations. J. Geophys. Res. 106(B1), 771-785.

[21] Ben-Menahem, A., 1977. Renormalization of the magnitude scale. Phys. Earth Planet. Inter. 15, 315?340.

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

[23] Beran, J., Feng, Y., Ghosh, S., Kulik, R., 2013. Long-memory processes, Springer-Verlag Berlin Heidelberg.

[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(7), 2761-2781.

[25] Biasi, G.P. and K.M. Scharer, 2019. The current unlikely earthquake hiatus at California's transform boundary paleoseismic sites, Seism. Res. Lett. 90(3), 1168-1176.

[26] Bilham, R., 2019. Himalayan earthquakes: a review of historical seismicity and early 21st century slip potential. From: Treloar, P. J. and Searle, M. P. (eds). Himalayan Tectonics: A Modern Synthesis. Geological Society, London, Special Publications, 483, 423?482. doi.org:10.1144/SP483.16.

[27] Boatwright, J. and H. Bundock, 2005. Modified Mercalli Intensity maps for the 1906 San Francisco earthquake plotted in ShakeMap format. U.S. Geological Survey. Open-File Report 2005-1135. Version 1.0. US Geological Survey, 345 Middlefield Road, MS 977, Menlo Park, CA 94025.

[28] Boatwright, J., Bundock, H., 2008. Modified Mercalli intensity maps for the 1868 Hayward earthquake plotted in ShakeMap formt. USGS Open-File Report 2008-1121.

[29] Boatwright, J., Bundock, H., 2008. The distribution of modified Mercalli intensity in the 18 April 1906 San Francisco earthquake. Bull. Seismol. Soc. Am. 98(2), 890-900.

[30] Borcherdt, R. D., 1970. Effects of local geology on ground motion near San Francisco Bay. Bull. Seismol. Soc. Am. 60(1), 29-61.

[31] Bozorgnia et al., 2014. NGA-West2 research project, Earthq. Spectra 30(3), 973?987.

[32] Brooks, E. M., Stein, S., Spencer, B. D., 2016. Comparing the performance of Japan's earthquake hazard maps to uniform and randomized maps. Seismol. Res. Lett. 87(1), 90-102.

[33] Brooks, E. M., Stein, S., Spencer, B. D., Salditch, L., Petersen, M. D., McNamara, D. E., 2017. Assessing earthquake hazard map performance for natural and induced seismicity in the central and eastern United States. Seismol. Res. Lett. 89(1), 118-126.

[34] Brooks, E., J. Neely, S. Stein, B. Spencer, and L. Salditch, 2019. Assessments of the performance of the 2017 one-year seismic hazard forecast for the Central and Eastern United States via simulated earthquake shaking data, Seism. Res. Lett., 90, 1155-1167.

[35] Brune, J.N., Anooshehepoor, A., Shi, B., Zeng, Y., 2004. Precarious rock and overturned transformer evidence for ground shaking in the Ms 7.7 Kern County earthquake: An analog for disastrous shaking from a major thrust fault in the Los Angeles basin. Bull. Seismol. Soc. Am. 94(6), 1993-2003.

[36] Byerly, P., Dyk, H., 1936. The questionnaire program for collecting earthquake data, Earthquake Investigations in California, 1934-1935, Special Publication No. 201, U.S. Department of Commerce Coast and Geodetic Survey, Washington, DC, 43-48.

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

[38] Calais E., Nocqet J.-M., Jouanne F., Tardy M., 2002. Current strain regime in the western Alps from continuous Global Positioning System measurements, 1996-2001. Geology 30, 651- 654.

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

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

[41] Chaulagain, H., Rodrigues, H., Silva, V., Spacone, E., Varum, H., 2015. Seismic risk assessment and hazard mapping in Nepal. Nat. Haz., 78,583?602. doi: 10.1007/s11069-015-1734-6.

[42] Cinlar, E., 1975. Introduction to stochastic processes, Prentice Hall, Englewood Cliffs, N.J.

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

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

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

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

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

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

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

[50] Delavaud, E., Scherbuam, F., Kuehn, N., and Allen, T., 2012. Testing the global applicability of Ground?Motion Prediction Equations for active shallow crustal regions, Bull. Seismol. Soc. Am. 102(2), 707-721. doi: 10.1785/0120110113.

[51] Dengler, L. A., and J. W. Dewey, 1998. An intensity survey of households affected by the Northridge, California, earthquake of 17 January 1994, Bull. Seismol. Soc. Am. 88(2), 441-462.

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

[53] De Rubeis, V., K. Van Noten, T. Patrizia, and P. Sbarra, 2019. 2019 Survey report on the Internet Macroseismic Practice in Europe. 7th International Colloquium on Historical Earthquakes and Paleoseismicity Studies, Barcelona, Spain.

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

[55] Dewey, J.W., Reagor, B.G., Dengler, L., Moley, K., 1995. Intensity distribution and isoseismal maps for the Northridge, California earthquake of January 17, 1994. USGS Open-File Report 95-92.

[56] Drouet, S., Ameri, G., Le Dortz, K., Secanell, R., Senfaute, G., 2018. A probabilistic seismic hazard map for the metropolitan France. Bull. Earthq. Eng. 18, 1865-1898. https://doi.org/10.1007/s10518-020-00790-7.

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

[58] Ellsworth, W.L., 2013 Injection-induced earthquakes. Science, 341. doi: 10.1126/science.1225942.

[59] Felzer, K.R., and E.E. Brodsky, 2005. Testing the stress shadow hypothesis, J. Geophys. Res., 110(5).

[60] Field, E., 2010. Probabilistic seismic hazard analysis: A primer, `http://www.opensha.org/`, last accessed 1/10/2020.

[61] Field, E.H., R.J. Arrowsmith, G.P. Biasi, P. Bird, T.E. Dawson, K.R. Felzer, D.D. Jackson, K.M. Johnson, T.H. Jordan, C. Madden and A.J. Michael, 2014. Uniform California earthquake rupture forecast, version 3 (UCERF3): The time-independent model. Bull. Seism. Soc. Am., 104(3), pp.1122-1180.

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

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

[64] Foulger GR, Wilson MP, Gluyas JG, Julian BR, Davies RJ (2018). Global review of human-induced earthquakes. Earth-Sci. Rev. 1(178), 438-514.

[65] Fujiwara, H., N. Morikawa, Y. Ishikawa, T. Okumura, J. Miyakoshi, N. Nojima, Y. Fukushima (2009). Statistical comparison of national probabilistic seismic hazard maps and frequency of recorded JMA seismic intensities from the K-NET strong-motion observation network in Japan during 1997-2006, Seism. Res. Lett. 80(3), 458-464.

[66] Gneiting, T. and M. Katzfuss, 2014. Probabilistic forecasting, Annu. Rev. Stat. App. 1, 125-151.

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

[68] Goh, K.I., Barabasi, A.L., 2008. Burstiness and memory in complex systems, Europhys. Lett. 81(4), 48002.

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

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

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

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

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

[74] Gomez-Capera, A.A., D?Amico, M., Lanzano, G., Locati, M., Santulin, M., 2020. Relationships between ground motion parameters and macroseismic intensity for Italy. Bull. Earthq. Eng. 18, 5143?5164. https://doi.org/10.1007/s10518-020-00905-0.

[75] Gutenberg, B., Richter, C.F., 1954. Seismicity of the Earth and associated phenomena. Princeton University Press, Princeton, New Jersey. Second edition.

[76] Hanks, T.C. and A. Johnston, 1992. Common features of the excitation and propagation of strong ground motion for North American earthquakes, Bull. Seism. Soc. Am. 82, 1-23.

[77] Harris, R.A. and R.W. Simpson, 1998. Suppression of large earthquakes by stress shadows: A comparison of Coulomb and rate-and-state failure, J. Geophys. Res., 103(B10), 24439-24451.

[78] Hartigan, J.A., 2006. Classification - 1, Encyclopedia of Statistical Sciences, 1-9, John Wiley and Sons, Inc.

[79] Hastie, T., Tibshirani, R., and Friedman, J., 2009. The elements of statistical learning: Data mining, inference, and prediction, 2nd ed, Springer, New York.

[80] Hauksson, E., K. Felzer, D. Given, M. Giveon, S. Hough, K. Hutton, H. Kanamori, V. Sevilgen, S. Wei, and A. Yong, 2008. Preliminary report on the 29 July 2008 Chino Hills, eastern Los Angeles Basin, California, earthquake sequence, Seism. Res. Lett. 79(6), 855-866.

[81] Hearn E. H., Pollitz, F. F., Thatcher, W. R., Onishi, C. T., 2013. How do “ghost transients” from past earthquakes affect GPS slip rate estimates on southern California faults?, Geochem. Geophys. Geosyst. 14, 828?838.

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

[83] Hough, S.E. 2000. On the scientific value of “unscientific” data, Seism. Res. Lett., 71, 483-485.

[84] Hough, S.E., 2012. Initial assessment of the intensity distribution of the 2011 Mw5.8 Mineral, Virginia, earthquake, Seismol. Res. Lett. 83:4, 649-657.

[85] Hough, S. E., 2013. Spatial variability of ‘Did You Feel It?’ intensity data: Insights into sampling biases in historical earthquake intensity distributions. Bull. Seismol. Soc. Am. 103(5), 2767-2781.

[86] Hough, S.E., 2014. Earthquake intensity distributions: a new view. Bull. Earthq. Eng. 12, 135-155.

[87] Hough, S.E., Elliot, A., 2004. Revisiting the 23 February 1892 Laguna Salada earthquake. Bull. Seismol. Soc. Am. 94(4), 1571-1578.

[88] Hough, S.E., Hutton, K., 2008. Revisiting the 1872 Owens Valley, California, earthquake. Bull. Seismol. Soc. Am. 98(2), 931-949.

[89] Hough, S. E., Page, M., 2011. Toward a consistent model for strain accrual and release for the New Madrid Seismic Zone, central United States. J. Geophys. Res. 116, B03311.

[90] Hough, S. E., Tsai, V. C., Walker, R, 2017. Was the Mw 7.5 1952 Kern County, California, earthquake induced (or triggered)? J. Seismol. 21,1613-1621.

[91] Hough, S.E., Martin, S.S., 2018. A proposed rupture scenario for the 1925 Mw 6.6 Santa Barbara, California, earthquake. Tectonophysics 747, 211-224.

[92] Hough, S.E. and Graves, R.W., 2018, December. The 1933 Long Beach, California, earthquake: Ground motions and rupture scenario. In AGU Fall Meeting Abstracts 2018, S54B-07.

[93] Huo, J-R., Hwang, H. H. M., 1995. Seismic fragility analysis of equipment and structures in a Memphis electric substation. National Center for Earthquake Engineering Research, Technical Report NCEER-95-0014.

[94] Hutton, K., Woessner, J., Hauksson, E., 2010. Earthquake monitoring in southern California for seventy-seven years (1932-2008), Bull. Seism. Soc. Am. 100(2), 423-446.

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

[96] Kagan, Y. Y., Jackson, D. D., 1991. Long-term earthquake clustering, Geophys. J. International. 104, 117-133.

[97] Kanamori, H., Jennings, P.C., 1978. Determination of local magnitude, ML, from strong-motion accelerograms, Bull. Seism. Soc. Am. 68(2), 471-485.

[98] Karsai, M., Kimmo, K., Barabasi, A., Kertesz, J. 2012. Universal features of correlated bursty behavior, Sci. Rep. 2, 397.

[99] Kaufman, L. and Rousseeuw, P.J., 1990. Finding groups in data: An introduction to cluster analysis, Wiley-Interscience, New Jersey.

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

[101] LaMontagne, M. and K.B.S. Burke, 2019. Status and future of macroseismic information in Canada, Seism. Res. Lett. 90(2B), 1014.

[102] Li, Q., Liu, M., Stein, S., 2009. Spatial-temporal complexity of continental intraplate seismicity: Insights from geodynamic modeling and implications for seismic hazard estimation, Bull. Seismol. Soc. Am. 99.

[103] Liu, M., Stein, S., 2016. Mid-continental earthquakes: Spatiotemporal occurrences, causes, and hazards, Earth-Sci. Rev. 162, 364-386.

[104] 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.

[105] Locati, M. and A. Cassera, 2010. MIDOP: Macroseismic Intensity Data Online Publisher. Rapporti Tecnici INGV, 123, 92.

[106] Locati M., Camassi R., Rovida A., Ercolani E., Bernardini F., Castelli V., Caracciolo C.H., Tertulliani A., Rossi A., Azzaro R., D?Amico S., Antonucci A., 2019. Database Macrosismico Italiano (DBMI15), versione 2.0. Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/DBMI/DBMI15.3.

[107] Loreto, M. F., Zitellini, N., Ranero, C. R., Palmiotto, C., and Prada, M., 2021. Extensional tectonics during the Tyrrhenian back?arc basin formation and a new morpho?tectonic map. Basin Research, 33(1), 138-158.

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

[109] Mak, S., R.A. Clements, D. Schorlemmer, 2014. The statistical power of testing probabilistic seismic-hazard assessments, Seism. Res. Lett. 85, 781-783.

[110] Mak, S. and D. Schorlemmer, 2016. A comparison between the forecast by the U.S. National Seismic Hazard Maps with recent ground?motion records, Bull. Seism. Soc. Am. 106(4), 1817-1831.

[111] Martin S.M., Hough, S.E., Hung, C., 2015. Ground motions from the 2015 Mw 7.8 Gorkha, Nepal, earthquake constrained by a detailed assessment of macroseismic data. Seism. Res. Lett., 86(6), 1524-1532.

[112] Martin, S.M. and Szeliga, W. (2010). A catalog of felt intensity data for 570 earthquakes in India from 1636 to 2009. Bull. Seismol. Soc. Am.,100(2), pp. 562?569. doi: 10.1785/0120080328.

[113] Martindale, D., Evans J.P., 2002, Historigraphical analysis of the 1857 Ft. Tejon earthquake, San Andreas Fault, California: Preliminary results, AGU Fall Meeting Abstracts, S12C-05.

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

[115] Marzocchi, W. and T.H. Jordan, 2014. Testing for ontological errors in probabilistic forecasting models of natural systems, Proc. Natl. Acad. Sci. U.S.A. 111(33), 11973-11978.

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

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

[118] Medvedev, S.V., Sponheuer, W. and V. Karnik, 1967. Seismic intensity scale version 1964, Punl. Inst. Geody. Jena, 48.

[119] Meltzner, A.J., Wald, D.J., 1999. Foreshocks and aftershocks of the great 1857 California earthquake. Bull. Seismol. Soc. Am. 89 (4), 1109-1120.

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

[121] Miyazawa, M. and J. Mori, 2009. Test of seismic hazard map from 500 years of recorded intensity data in Japan, Bull. Seism. Soc. Am. 99, 3140-3149.

[122] Mucciarelli, M., 2007. A test for checking earthquake aperiodicity estimates from small samples, Nat. Hazard Earth Sys. Sci. 7(3), 399-404.

[123] Murphy, L.M., Cloud, W.K., 1954. United States Earthquakes 1952, Department of Commerce, United States Government Printing Office, 117, Washington DC.

[124] Murtagh, F. and Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?, Journal of Classification 31, 274-295.

[125] Musson, R. M., Grunthal, G., Stucchi, M., 2010. The comparison of macroseismic intensity scales. J. of Seismol., 14(2), 413-428.

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

[127] Nath, S,K, and K.K.S. Thingbaijam, 2012. Probabilistic seismic hazard assessment of India. Seis. Res. Let. 83(1). doi: 10.1785/gssrl.83.1.135.

[128] Naylor, M., I. G. Main, and S. Touati, 2009. Quantifying uncertainty in mean earthquake interevent times for a finite sample, J. Geophys. Res. 114, doi: 10.1029/2008JB005870.

[129] Nekrasova, A., V. Kossobokov, A. Peresan and A. Magrin, 2014. Comparison of NDSHA, PSHA seismic hazard maps and real seismicity for the Italian territory, Nat. Hazards 70(1), 629-641.

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

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

[132] Oakshotte, G.B., ed., 1955. Earthquakes in Kern County California during 1952. California Div. of Mines Bull. 171.

[133] Page, M. and K. Felzer, 2015. Southern San Andreas fault seismicity is consistent with the Gutenberg-Richter magnitude-frequency Distribution. Bull. Seismol. Soc. Am. 105(4). doi: 10.1785/0120140340.

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

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

[136] Peers, G.A., 1955. Damage to electrical equipment caused by Arvin-Tehachapi earthquake. Earthquakes in Kern County California during 1952. California Div. of Mines Bull. 171.

[137] Petersen, M.D., M.P. Moschetti, P.M. Powers, C.S. Mueller, K.M. Haller, A.D. Frankel, Y. Zeng, S. Rezaeian, S.C. Harmsen, O.S. Boyd, E.H. Field, R. Chen, K.S. Rukstales, N. Luco, R.L. Wheeler, R.A. Williams, and A.H. Olsen, 2014. Documentation for the 2014 update of the United States National Seismic Hazard Maps. USGS OFR 2014-1091.

[138] Petersen, M.D., Mueller, C.S., Moschetti, M.P., Hoover, S.M., Llenos, A.L., Ellsworth, W.L., Michael, A.J., Rubinstein, J.L., McGarr, A.F., and Rukstales, K.S., 2016, 2016 One-year seismic hazard forecast for the Central and Eastern United States from induced and natural earthquakes: U.S. Geological Survey Open-File Report 2016?1035. doi: 10.3133/ofr20161035.

[139] Petersen, M.D, A.M. Shumway, P.M. Powers, C.S. Mueller, M.P. Moschetti, A.D. Frankel, S. Rezaeian, D.E. McNamara, N. Luco, O.S. Boyd, K.S. Rukstales, K.S. Jaiswal, E.M. Thompson, S.M. Hoover, B.S. Clayton, E.H. Field, Y. Zeng, 2020. The 2018 update of the US National Seismic Hazard Model: Overview of model and implications. Earthq. Spectra 36(1), 5-41.

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

[141] Ram, T.D. and W. Guoxin., 2013. Probabilistic seismic hazard analysis in Nepal. Earthq. Eng. and Eng. Vib. 12(4), 577-586. doi: 10.1007/s11803-013-0191-z.

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

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

[144] Richter, C.F., 1958. Elementary Seismology. W.F. Freeman and Company, San Francisco, and Bailey Bros. and Swinfen Ltd., London.

[145] Rockwell, T.K., Lindvall, S., Herzberg, M., Murbach, D., Dawson, T., Berger, G., 2000. Paleoseismology of the Johnson Valley, Kickapoo, and Homestead Valley faults: Clustering of earthquakes in the eastern California shear zone, Bull. Seismol. Soc. Am. 90(5), 1200-1236.

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

[147] Rogers, G., Dragert, H., 2003. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip, Science 300(5627), 1942-1943.

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

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

Rukstales, K.S. and M.D. Petersen, 2019. Data release for 2018 update of the U.S. National Seismic Hazard Model: U.S. Geological Survey data release,`https://doi.org/10.5066/P9WT5OVB`. (`https://www.sciencebase.gov/catalog/item/5cbf47c4e4b0c3b00664fdef`, accessed 12/30/2019).

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

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

[152] Salditch, L., S.E. Hough, S. Stein, B.D. Spencer, E.M. Brooks, J.S. Neely and M.C. Lucas, 2018. The 1952 Kern County, California earthquake: A case study of issues in the analysis of historical intensity data for estimation of source parameters, Phys. Earth Planet. In. 283, 140-151, `https://doi.org/10.1016/j.pepi.2018.08.007`.

[153] Salditch, L., Gallahue, M.M., Lucas, M.C., Neely, J.S., Hough, S.E., Stein, S., 2020. California Historical Intensity Mapping Project (CHIMP): a consistently re-interpreted dataset of maximum shaking for the past 162 years and implications for PSHA. Seismol. Res. Lett., 91(5), p. 2631-2650. doi: 10.1785/0220200065.

[154] Salditch, L., Stein, S., Neely, J., Spencer, B.D., Agnon, A., Liu, M., 2020. Earthquake supercycles and Long-Term Fault Memory. Tectonophysics. 774, 228-289. doi: 10.1016/j.tecto.2019.228289.

[155] Satake K., 2015. Geological and historical evidence of irregular recurrent earthquakes in Japan, Phil. Trans. R. Soc. A 373, 20140375.

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

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

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

[159] Shao, J., Tu, D., 1995. The Jackknife and Bootstrap. New York: Springer.

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

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

[162] Sibson, R. , 1981. A brief description of natural neighbor interpolation, chapter 2 in Interpolating Multivariate Data. New York: John Wiley and Sons. 21-36.

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

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

[165] Sira, C., S. Auclair, I. Bouroullec, C. Cornou, I. Douste-Bacque, O. Guilhot, H. Jomard, E. Koufoudi, K. Manchuel, A. Montabert, A. Pothon, L. Provost, G. Senechal and A. Schlupp, 2019. Exploratory workshop to study the effect of time on macroseismic analyses. 7th International Colloquium on Historical Earthquakes and Paleoseismicity Studies, Barcelona, Spain.

[166] Sirovich L., F. Pettenati, F. Cavallini and M. Bobbio, 2002. Natural neighbor isoseismals, Bull. Seism. Soc. Am. 92, 1933? 1940.

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

[168] Sokolov, V.Y. and Y.K. Cherov, 1998. On the correlation of seismic intensity with Fourier amplitude spectra, Earthquake Spectra 14, 679-694. doi.org/10.1193/1.1586022.

[169] Stein, S., Brooks, E.M., Spencer, B.D., Liu, M., 2018. Should all of Nepal be treated as having the same earthquake hazard? From: J.H. Kruhl et al. (eds.), Living Under the Threat of Earthquakes, Springer Natural Hazards. doi.org/10.1007/978-3-319-68044-62.

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

[171] Stein, R.S., Thatcher, W., 1981. Seismic and aseismic deformation associated with the 1952 Kern County, California, earthquake and relationship to the quaternary history of the White Wolf fault. J. Geophys. Res 86(B6), 4913-4928.

[172] Stein, S., Okal, E.A., 2007. Ultralong period seismic study of the December 2004 Indian Ocean earthquake and implications for regional tectonics and the subduction process, Bull. Seismol. Soc. Am. 97(1A), S279-S295.

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

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

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

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

[177] Stein, S., Spencer, B. D. and Brooks, E. M., 2015. Metrics for assessing earthquake hazard map performance, Bull. Seismol. Soc. Am. 105(4), 2160?2173.

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

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

[180] Stein, S., Spencer, B. D., Brooks, E. M., 2015. Metrics for assessing earthquake hazard map performance. Bull. Seismol. Soc. Am. 105(4), 2160-2173.

[181] Stein, S. and G. Sella, 2005. Pleistocene change from convergence to extension in the Apennines as a consequence of Adria microplate motion, in The Adria Microplate: GPS Geodesy, Tectonics and Hazards, Nat. Sci, Series, 21-34, edited by Pinter, N., G. Grenerczy, J. Weber, S. Stein, and D. Medak, Springer.

[182] Steinbrugge, K.V., Moran, D.F., 1954. An engineering study of the Southern California earthquake of July 21, 1952 and its aftershocks. Bull. Seism. Soc. Am. 44(2B), 201-462.

[183] Stevens, V.L., Shrestha, S.N., Maharjan, D.K., 2018. Probabilistic seismic hazard assessment of Nepal. Bull. Seis. Soc. Am. doi: 10.1785/0120180022.

[184] Stirling, M. and M. Gerstenberger, 2010. Ground motion-based testing of seismic hazard models in New Zealand, Bull. Seism. Soc. Am. 100(4), 1407-1414.

[185] Stirling, M.W. and M. Petersen, 2006 . Comparison of the historical record of earthquake hazard with seismic-hazard models for New Zealand and the continental United States, Bull. Seism. Soc. Am. 96, 1978-1994.

[186] Stover, C.W. and J.L Coffman, 1993. Seismicity of the United States, 1568-1989. U.S. Geological Survey Professional Paper 1527. United States Government Printing Office, Washington, D.C.

[187] Stucchi, M., P. Albini, C. Mirto and A. Rebez, 2004. Assessing the completeness of historical earthquake data, Ann. Geophys. 47(2/3), 659?673.

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

[189] Tasan, H., C. Beauval, A. Helmstetter, A. Sandikkaya and P. Guguen, 2014. Testing probabilistic seismic hazard estimates against accelerometric data in two countries: France and Turkey, Geophys. J. Int. 198, 1554-1571.

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

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

[192] Toppozada, T.R., C.R. Real, S.P. Bezore and D.L. Parke, 1981. Preparation of isoseismal maps and summaries of reported effects for pre-1900 California earthquakes. US Geological Survey.

[193] Toppozada, T. and D. Branum, 2004. California earthquake history, Annals of Geophysics 47(2/3).

[194] Turcotte, D.L., 1997. Fractals and chaos in geology and geophysics, Cambridge University Press.

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

[196] Trifunac, M.D. and A.G. Brady, 1975. On the correlation of seismic intensity scales with peaks of recorded ground motion, Bull. Seism. Soc. Am., 65(1), 139-162.

[197] Trifunac, M.D., 2003. Nonlinear soil response as a natural passive isolation mechanism: The 1933 Long Beach, California earthquake, Soil. Dyn. And Earthq. Eng., 23, 549-562.

[198] U.S. Department of Commerce, Environmental Services Administration, Coast and Geodetic Survey, 1966. Abstracts of earthquake reports for the Pacific Coast and the Western Mountain Region, MSA-74 April, June, July 1952.

[199] Vanneste, K., S. Stein, T. Camelbeeck and B. Vleminckx, 2018. Insights into earthquake hazard map performance from shaking history simulations, Sci. Re. 8, 1855.

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

[201] Vick, S.G., 2002. Degrees of belief: Subjective probability and engineering judgment. ASCE Publications.

[202] Wald, D.J., V. Quitoriano, L. Dengler and J.W. Dewey, 1999. Utilization of the internet for rapid community intensity maps, Seism. Res. Lett. 70(6), 680-693.

[203] Wald, D.J., Dengler, L., Dewey, J.W., 1999. Utilization of the internet for rapid community intensity maps. Seismol. Res. Lett. 70(6), 680-693.

[204] Wald, D.J., V. Quitoriano, T.H. Heaton, H. Kanamori, C.W. Scrivner, Worden, C.B., 1999b. TriNet 'ShakeMaps': Rapid generation of peak ground-motion and intensity maps for earthquakes in Southern California. Earthq. Spectra 15(3), 537-556.

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

[206] Walls, C., Rockwell, T., Mueller, K., Bock, Y., Williams, S., Pfanner, J., Dolan, J., Fang, P., 1998. Escape tectonics in the Los Angeles metropolitan region and implications for seismic risk. Nature, 394(6691), 356.

[207] Wang, Z., 2015. Predicting or forecasting earthquakes and the resulting ground motion hazards: a dilemma for earth scientists, Seism. Res. Lett. 86, 1-5.

[208] Ward, J.H., 1963. Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association 58(301), 236-234.

[209] Ward, S.N., 1995. Area-based tests of long-term seismic hazard predictions. Bull. Seism. Soc. Am. 85, 1285-1298.

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

[211] Weldon, R., Scharer, K., Fumal, T., Biasi, G., 2004. Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work, GSA Today 14(9), 4-10.

[212] 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.

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

[214] Wesnousky, S.G., 2020. Great pending Himalaya earthquakes. Seismol. Res. Lett. 91 (6): 3334 ?3342. doi: 10.1785/0220200200.

[215] Wessel, P., Smith, W.H.F., 1991. Free software helps map and display data. Eos Trans Am. Geophys. Union 72, 441.

[216] Woessner, J., Danciu L., D. Giardini and the SHARE consortium, 2015. The 2013 European Seismic Hazard Model: key components and results, Bull. Earthq. Eng., doi:10.1007/s10518-015-9795-1.

[217] Wood, H.O., Neumann, F., 1931. Modified Mercalli intensity scale of 1931, Bull. Seismol. Soc. Am. 21(4), 277-283.

[218] Worden, C.B., Gerstenberger, M.C., Rhoades, D.A., Wald, D.J., 2012. Probabilistic relationships between ground-motion parameters and modified Mercalli intensity in California. Bull. Seism. Soc. Am. 102, 204-221.

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

[220] Wyss, M., Gupta, S., Rosset, P., 2018. Casualty estimates in repeat Himalayan earthquakes in India. Bull. Seis. Soc. Am., 108(5A), 2877-2893. doi: 10.1785/0120170323.