Decoding Seismic Hazard Enigmas via Next-Generation Paleoseismic Analyses: With Application to the Cascadia Subduction Zone and Beyond

Ryan A. Rasanen

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Brett Maurer, Chair

Pedro Arduino

Joseph Wartman

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Abstract

Decoding Seismic Hazard Enigmas via Next-Generation Paleoseismic Analyses: With Application to the Cascadia Subduction Zone and Beyond

Ryan A. Rasanen

Chair of the Supervisory Committee:
Brett Maurer
Department of Civil and Environmental Engineering

In regions of infrequent moderate-to-large earthquakes, historic earthquake catalogs are often insufficient to provide inputs to seismic-hazard analyses (i.e., fault locations and magnitude-frequency relations) or to inform ground-motion predictions for certain seismic sources. As a result, analyses of macroseismic evidence (e.g., landslides, liquefaction, and felt intensities) are commonly used to help elucidate the seismic record, thereby reducing seismic-hazard uncertainty. However, while paleoseismic studies have been performed widely, existing inverse-analysis techniques (i.e., to constrain the causative earthquake magnitude and/or ground motions) have several shortcomings. Foremost, they are generally deterministic and require that the location of fault rupture be known, when in reality there are numerous uncertainties, and the source location is typically unknown. Accordingly, this dissertation (1) proposes a novel, flexible inversion methodology for analyzing paleoseismic evidence, resulting in probabilistic constraint of the seismic-source parameters (e.g., rupture location and magnitude), to the degree that the observational evidence permits. Following from (1), the methodology is tested using simulated paleoseismic studies of modern (2) liquefaction and (3) landslides. Following these simulated studies, from which lessons for forward use are gleaned, the methodology is (4) applied to evidence produced by the 1886 Charleston, SC earthquake, thereby constraining its seismic source in a more complete and transparent manner than previously available. Next, paleoliquefaction evidence in the U.S. Pacific Northwest (PNW) is (5) compiled into a community database and liquefaction in an M_w9 Cascadia Subduction Zone (CSZ) event is forecasted regionally using 30 different physics-based ground motion realizations and a first-order geospatial liquefaction model. These products are (6) made available as an online database so that others may study, and benefit from, the compiled data. Lastly, liquefaction resulting from M_w9 CSZ earthquakes is predicted at select locations using subsurface test data and more advanced geotechnical methods. As a precursor, (7) liquefaction models based on the cone penetration test (CPT) are tested in the PNW via compilation and analysis of case histories from the 2001 Nisqually, WA, earthquake. Following model testing, (8) M_w9 liquefaction is predicted for each of the 30 ground-motion simulations at 400 locations where CPT data is available. These predictions suggest that liquefaction could be both pervasive in the region and affect numerous significant hubs of population, transportation, and commerce. These predictions are compared to the evidence from (5) to assess which of the 30 simulations represent more and less plausible realizations of the motions experienced in 1700. This comparison suggests that all simulations predict far more liquefaction than has been found and that shaking intensities in 1700 may have been less in inland hubs (e.g., Seattle, Portland, and Vancouver), than prevailing forecasts indicate. The uncertainties associated with this possibility are discussed in detail.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First and foremost, my deepest gratitude goes to my research advisor, Dr. Brett Maurer, whose unwavering guidance, mentorship, and friendship have shaped my PhD journey. Dr. Maurer has gone above and beyond his role as an advisor, helping me grow not only as a researcher but also as a teacher and individual. I wholeheartedly endorse Brett as an advisor and mentor for aspiring researchers in geotechnical earthquake engineering.

I would also like to thank my committee members, Dr. Pedro Arduino, Dr. Marine Denolle, and Dr. Joseph Wartman, for their invaluable insights that have enriched this dissertation. I am particularly grateful to Dr. Arduino for his instrumental role as a second mentor during my time at UW, helping me develop my leadership skills as the President of the Geo-Institute Graduate Student Society (GIGSS) for four years. Moreover, I appreciate the late-night discussions with Pedro about research, GIGSS, and life. Special gratitude goes to my Masters advisor, Dr. Steve Kramer, who first taught me the intricacies of liquefaction. When I first arrived, I had limited knowledge of this fluid subject, but Steve skillfully liquefied those barriers, helping me comprehend its complexities. I also want to acknowledge Dr. Mertcan Geyin for his invaluable assistance at the beginning of my PhD and for being a close friend throughout this journey.

The support I received from the U.S. Geological Survey (USGS) throughout my PhD has been irreplaceable. From Dr. Brian Atwater's guidance, taking me to explore paleoliquefaction features and enlightening me about the 1700 CE Cascadia earthquake, to the research assistance and mentorship provided by Dr. Alex Grant and Dr. Erin Wirth during the final year of my PhD. I am especially grateful to Alex for his tremendous effort in facilitating the use of the USGS CPT truck, significantly advancing my research. In addition, I am eternally grateful to the National Science

Foundation Graduate Research Fellowship for providing three years of funding, without which this PhD would not have been possible.

Lastly, I would like to thank my family and friends for their unwavering support throughout my grad school journey and my life. My mom, Laurie, has been an incredible source of inspiration, embodying true work ethic and always being there when I needed someone to talk to. My dad, Rich, has shared his humor, love for sports, and provided a welcome respite with enjoyable conversations during breaks from research. To my partner, Tracy Arietta, who has been a constant pillar of support, helping me through the toughest times of my PhD, even when I would go months on end working until 10:00pm. I am also grateful to my sister, Julie, for her ongoing support throughout life. Additionally, I extend my thanks to the entire Rasanen family for their unconditional love and support, making every holiday memorable and filled with spirited Catan games and 4th of July parties. I am also grateful to all the UW Geotech students who have made my time at UW truly special. There are many others to thank as well, for all of those not mentioned here, thank you.

DEDICATION

To my family, the Rasanens, Martens, and Moses.

Chapter 1. INTRODUCTION

This dissertation follows the "manuscript format," consisting of 9 publications that include datasets, conference papers, and journal articles. The introductory chapter first discusses the background and motivations behind the collective body of work. Next, the publications comprising the dissertation are listed and the connections between them are illustrated schematically. Lastly, the motivations and objectives of each publication are succinctly stated. Later, in Chapters 2 and 3, each publication's principal products and findings are summarized in an expanded, yet still brief format. Finally, all publications are provided in-full in separate appendices.

1.1 Background and Motivations

In many seismic zones, the return period of moderate-to-large earthquakes is longer than the duration of recorded seismicity (e.g., the New Madrid Seismic Zone, South Carolina Coastal Plain, and Cascadia Subduction Zone, among others in the United States). As a result, historic earthquake catalogs may be insufficient to inform the inputs requisite for seismic-hazard analyses (i.e., the locations, magnitudes, and frequencies of fault ruptures). The uncertainties of these parameters are more than an academic curiosity, given that they influence the computed seismic hazard (e.g., the US National Seismic Hazard Map) (Vidale et al., 2011), which is adopted by building codes and used by engineers and policy makers. If the uncertainty is biased toward greater-than-actual hazards, capital is spent superfluously on seismic design, construction, and rehabilitation of infrastructure, hindering the economy (Murphy, 2011; Stein, 2010). Conversely, if seismic hazards are underestimated, lives, infrastructure, and property are at risk. For example, a 2005 study predicted that a repeat of the 1886 Charleston, South Carolina, earthquake would cause 900 deaths, 44,000 injuries, and economic losses of $20 billion (Wong et al., 2005).

In addition to being affected by this uncertainty, many seismic zones have another commonality: pre-historic and/or pre-instrumental earthquakes induced macroseismic evidence, such as soil liquefaction (or "paleoliquefaction") that was later used to infer information about the causative ruptures (Figure 1). Paleoliquefaction has been studied in several regions of the United States, examples of which are listed in Table 1. Unsurprisingly, relic liquefaction also influences the US National Seismic Hazard Maps (Petersen et al., 2008, 2014). In addition, other forms of paleoseismic evidence, such as landslides and felt intensities, can similarly be used to elucidate the seismic record. However, paleoseismic research to date has focused more on field investigations, with less effort focused on advancing the techniques for inverse analyzing field evidence. Both components are important, but the focus of this study is on the latter.

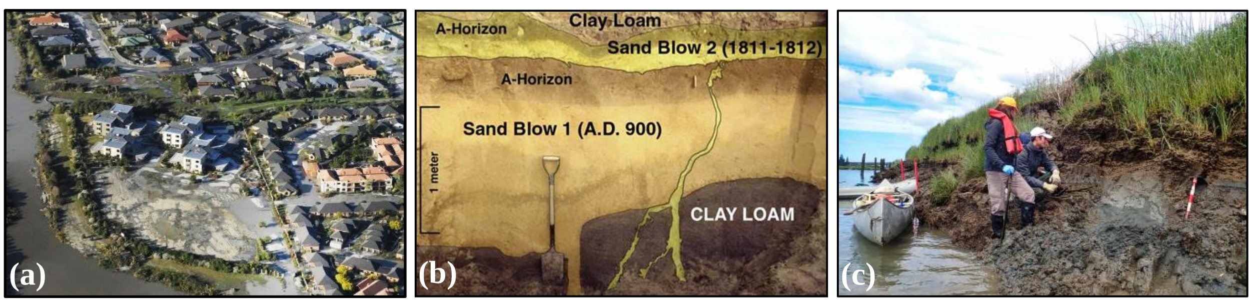


Figure 1. Soil liquefaction, as in (a) Christchurch, NZ, tends to be preserved in the geologic record; for example: (b) Missouri, USA; (c) Washington, USA. Photos (a) and (b) are courtesy of NSF Geotechnical Extreme Events and Reconnaissance (GEER) and the United States Geological Survey (USGS), respectively.

Table 1. Select regions of paleoliquefaction study in the United States.

Seismic Zone	Location	Representative References
New Madrid	Missouri, Arkansas, Tennessee	Tuttle et al. (2005)
Wabash Valley	Illinois, Indiana	Green et al. (2005)
South Carolina Coastal Plain	South Carolina	Talwani and Schaeffer (2001)
Cape Ann	Massachusetts	Ellis and de Alba (1999)
Clarendon	New York	Tuttle et al. (2002)
Cascadia Subduction Zone	Washington, Oregon	Obermeier and Dickenson (2000)
Mississippi Embayment	Arkansas	Cox et al. (2004)
San Diego	California	Kuhn (2005)
Willamette Valley	Oregon	Peterson et al. (2014)
Puget Sound	Washington	Bourgeois and Johnson (2001)

Notably, existing paleoseismic analysis methods suffer from up to three shortcomings. *First*, existing methods are underpinned by the need to locate the earthquake that induced the observed field of evidence (e.g., earthquake magnitude cannot otherwise be constrained). This requires assumptions in many settings (e.g., where faults are blind or historically quiescent) and, in turn, may lead to erroneous characterizations of the seismic hazard. Moreover, existing techniques do not explicitly consider the influence of source-location uncertainty, which can be significant, on the uncertainty of magnitude estimates derived from paleoseismic analyses. *Second*, existing methods have traditionally analyzed locations *with* observed paleoseismic evidence (e.g., positive observations of liquefaction or landslides) but not locations *without* observed evidence (e.g., negative observations of liquefaction or landslide). As a result, derivative estimates of shaking intensity and/or earthquake magnitude (e.g., using "magnitude-bound curves") are inherently lower-bound estimates, even if this is not explicitly stated. *Third*, existing methods have been deterministic in nature, whereas paleoseismic analyses involve multiple uncertainties that should be accounted for probabilistically to arrive at a complete and transparent characterization of the seismic hazard. An approach that addresses these shortcomings would thus greatly benefit paleoseismic analyses and could provide new insights in regions of enigmatic seismic hazard.

The South Carolina Coastal Plain is one such region. The 1886 Charleston, SC earthquake caused widespread liquefaction (Amick et al. 1990), damaged structures in multiple U.S. states, including most structures in Charleston, SC (Dutton 1889; Wong et al., 2005), produced perceptible shaking over 1500 km away in Canada (Bakun et al., 2002), and was larger in magnitude than any earthquake to since occur in Eastern North America (ENA). However, because the 1886 rupture predated seismic instruments and did not manifest at the surface (Nuttli et al., 1986), its exact location and magnitude remain uncertain, as do the regional intensities of resultant ground motions. Published confidence intervals of the 1886 magnitude range from M_w6.4 to M_w7.8, as interpreted from felt-intensity reports (e.g., Bakun and Hopper, 2004; Cramer and Boyd, 2014) and liquefaction evidence (e.g., Martin and Clough, 1994; Hayati and Andrus, 2008). Liquefaction features also suggest a history of recurrent earthquakes in the region extending back 6000 years (Talwani and Schaeffer, 2001) with wide ranging magnitude estimates of M_w5.1 to M_w7.8 (e.g., Hu et al., 2002; Gheibi et al., 2020). Considering the existing literature, Petersen et al. (2014, 2020) assigned a magnitude probability distribution to the Charleston source that ranged from M_w6.7 to M_w7.5 in the most recent U.S. National Seismic Hazard Model Project (NSHMP) maps. While much has been learned about the 1886 earthquake from macroseismic evidence, prior analyses suffer from each of the limitations above. That is, they tend to be deterministic and assume the event was caused by a particular fault (i.e., the "Woodstock Fault") even though the fault(s) responsible for the event are debated and the characteristics of the Woodstock Fault are uncertain. Moreover, the analysis of paleoseismic evidence requires a series of models for predicting the relevant phenomena. For example, regionally distributed ground motions must be predicted, conditioned on a hypothetical source. The probability of field observations (e.g., liquefaction response) must then be computed, conditioned on the expected ground-motions. In this regard, major modelling advances have recently been made. The NGA-East project (Goulet et al., 2018; Harmon et al., 2019), for example, resulted in the most advanced understanding of ENA ground motions and site response to-date, but these and other advances have not been incorporated in any existing analysis of the Charleston earthquake. A comprehensive revisit of the 1886 event that includes a new inversion methodology and the most up-to-date models would confirm or refine the magnitude distribution assumed by the NSHMP.

In a region with similar uncertainty, the Cascadia Subduction Zone (CSZ) is thought to be capable of producing M_w9 earthquakes (Goldfinger et al., 2012), yet all prior events predate seismic instruments, with the most recent interface rupture occurring in 1700 CE (e.g., Atwater et al. 1991, Atwater et al., 2005). While past CSZ earthquakes have been dated, their magnitudes and the intensities of resultant ground motions remain unknown. Given the premise of a full-margin M_w9 CSZ earthquake, various ground-motion predictions have been made (e.g., USGS, 2017). These include – most recently – a suite of broadband synthetic seismograms (i.e., ground-motion time histories) (Frankel et al., 2018; Wirth et al., 2018) that predict motions on a 1-km grid across the Pacific Northwest (PNW). This suite includes 30 different realizations to reflect the uncertainty of key parameters (e.g., the down-dip limit of fault rupture; the slip distribution and location of asperities; and hypocenter location). Given that the last CSZ event occurred in 1700, these parameters are unknown for all such events. Shown in Figure 2 are two such realizations. While both simulate a full-margin M_w9 CSZ rupture, predicted ground motions vary significantly in some locales (e.g., peak ground velocities may vary by 400%). By corollary, the expected impacts on the built and living environments would also be very different (Marafi et al., 2019, 2020).

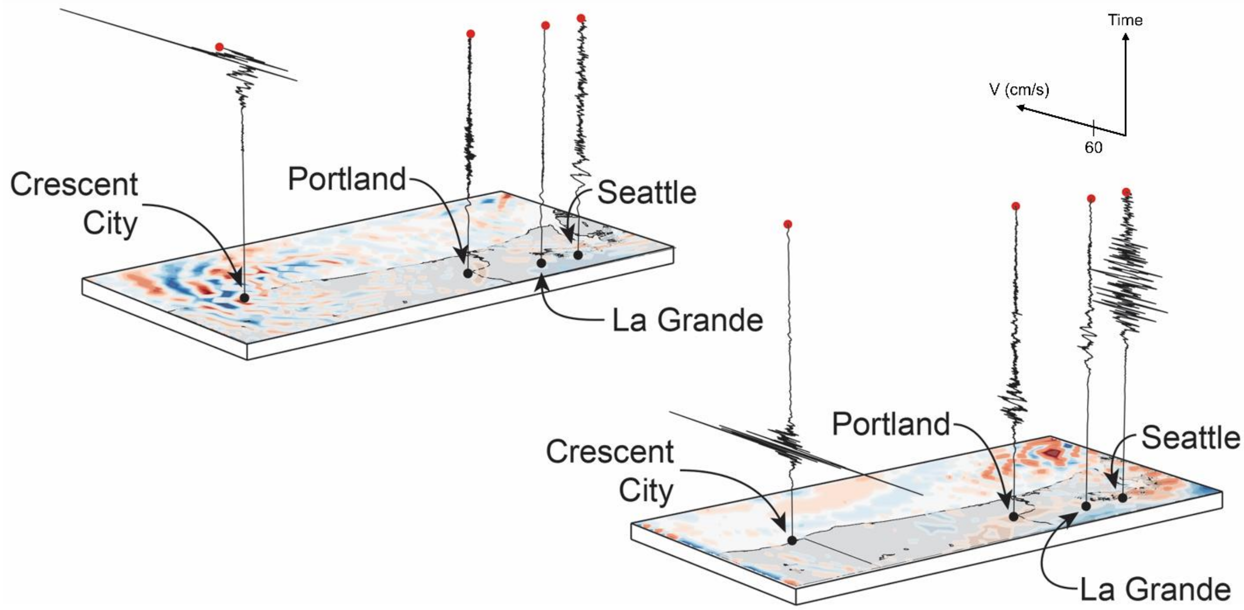


Figure 2. Two ground-motion predictions of a M_w9 CSZ earthquake, reflecting the influence of salient modeling uncertainties (traces are ground velocity for the duration of the event; colors show a snapshot of seismic waves at a moment in time). The sensitivity of predictions to modelling uncertainties is readily apparent via the large discrepancies at some locales (e.g., in Seattle). Simulations by Frankel et al. (2018) and Wirth et al. (2018).

Notably, the actual ground motions experienced in 1700 (and in other paleo events) can be determined only through inverse analysis of the coseismic evidence. But, of the many types of coseismic evidence that have been documented in Cascadia, only soil liquefaction and landslides are presently capable of "recording" the intensity of ground motions. This distinction arises because mechanistic models exist for predicting these phenomena as a function of ground-motion intensities, whereas other forms of evidence may suggest that an earthquake occurred but can do little to quantify the ground motions experienced. While paleoliquefaction has been identified at over 180 locations in the PNW, including many from the 1700 Cascadia event, paleolandslides have yet to be linked to any CSZ earthquake. Thus, liquefaction presently provides the best evidence from which the intensity of shaking in PNW paleoearthquakes may be quantified. However, the documented evidence exists across numerous publications and has not been centrally compiled or digitized, leading to an absence of regional scale paleoliquefaction studies in the PNW.

In addition to compiling and studying existing paleoliquefaction evidence in the PNW, it is also critical that additional evidence be discovered and analyzed, since better constraint of regional ground motions in past earthquakes will result. However, a field search of the entire region would be extremely cost-prohibitive and would not necessarily be useful. There are an infinite number of locations, for example, where the confirmed presence or absence of 1700 liquefaction would do little to inform or constrain ground-motion predictions (e.g., because various predictions lead to similar expectations of liquefaction). For field pursuits to be conducted more efficiently and strategically, it would thus be helpful to identify locations where uncertainties in ground-motion simulations (e.g., Frankel et al., 2018; Wirth et al., 2018) give rise to significant differences in liquefaction predictions.

Accordingly, the research presented herein is motivated by the value of paleoseismic evidence (i.e., liquefaction, landslides, felt intensities), which has the potential to provide new insights in regions of seismic-hazard uncertainty, but which has arguably not been exploited fully. Specifically, this dissertation will (1) propose a new methodology to inverse analyze paleoseismic evidence, resulting in probabilistic constraint of the seismic-source parameters (e.g., location and magnitude), to the degree that the observational evidence permits. Following from (1), the methodology will be tested via simulated paleoseismic studies of modern (2) liquefaction and (3) landslides. Following these simulated studies, from which lessons for forward use will be gleaned, the methodology will be (4) applied to evidence produced by the 1886 Charleston, SC earthquake. Next, paleoliquefaction evidence in the PNW will (5) be compiled into a GIS database and liquefaction in an M_w9 CSZ event will be forecasted regionally using 30 different physics-based ground motion realizations and a first-order geospatial liquefaction model. These products will (6) be made available as an online database so that others may study, and benefit from, the compiled data. Lastly, liquefaction resulting from M_w9 CSZ earthquakes will be predicted at select locations using subsurface test data and more advanced geotechnical methods. As a precursor, (7) liquefaction models based on the cone penetration (CPT) test will be tested in the PNW via compilation and analysis of case histories from the 2001 Nisqually, WA, earthquake. Following model testing, (8) M_w9 liquefaction will be predicted for each of the 30 ground-motion simulations at 400 locations where CPT data is available. These predictions will be compared to the evidence from (5) to assess which of the 30 simulations represent more and less plausible realizations of the motions experienced in 1700.

1.2 Dissertation Structure and Contents

Table 2 lists publications resulting from this dissertation, parsed into two themes/chapters. Chapter 2 focuses on improving paleoseismic analytics while Chapter 3 focuses on field application in two seismic zones. The principal products and findings of publications are succinctly summarized in the ensuing chapters; these publications are also provided in their complete form in separate appendices. The workflow giving rise to the publications in Table 2, and the connections between them, is shown in Figure 3.

Table 2. Publications resulting from this dissertation.

No.	**Type**	**Summary in Ch.**	**Publication Name**	**Current Status**
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.
[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

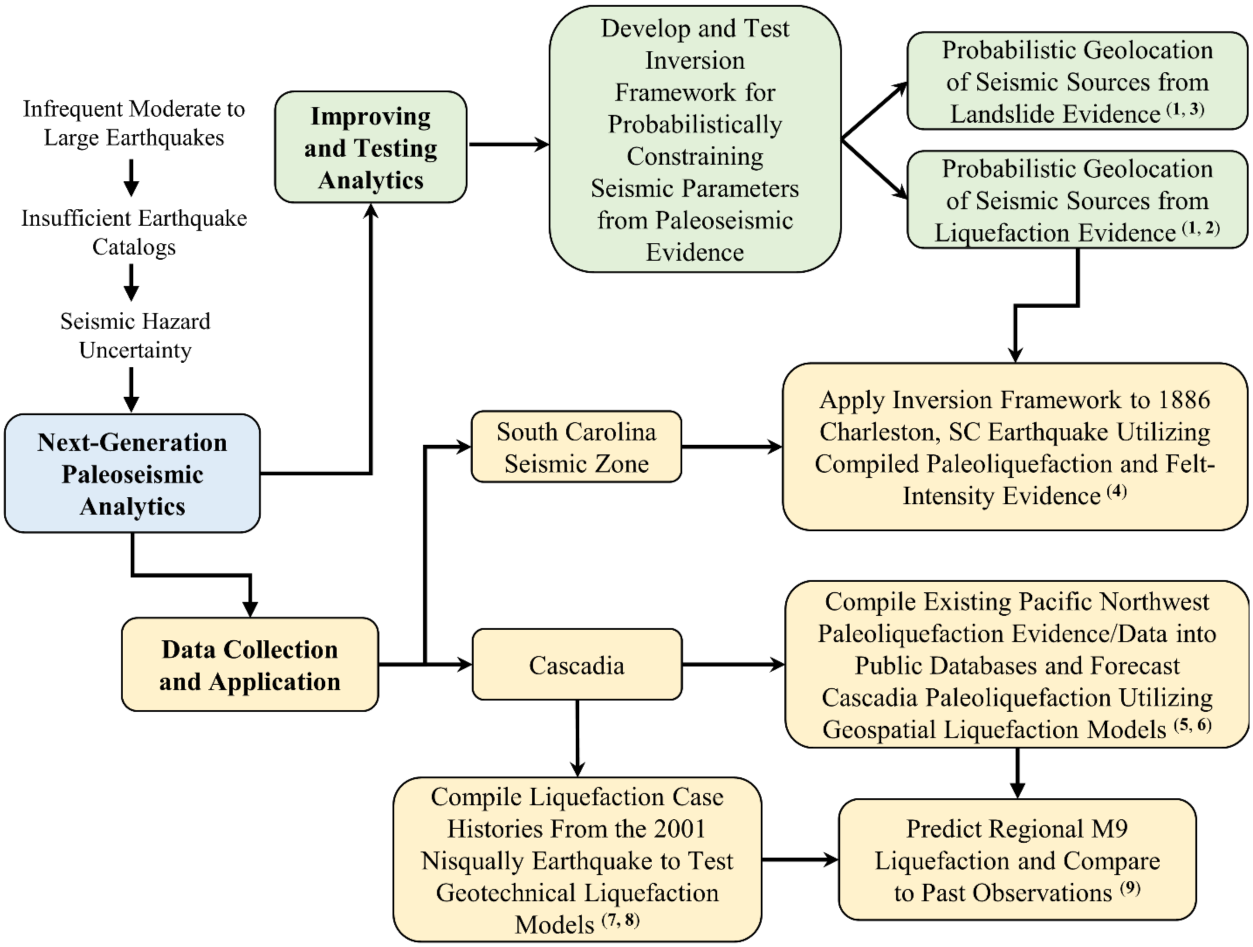


Figure 3. Schematic showing the flow and interconnection of research subjects comprising this dissertation. The three colors correspond to the three chapters of the dissertation. Numbers in each box correspond to the publications appearing in the first column of Table 2.

1.3 A Succinct Summary of the Ensuing Chapters

Chapter two summarizes the development and testing of a flexible inversion methodology that probabilistically constrains seismic source parameters from macroseismic evidence. With minimal modification, the methodology is applicable to soil liquefaction, landslides, and felt-intensity data. This methodology significantly improves upon past inversion techniques and can produce probability distributions of rupture location, geometry, and magnitude. To test the methodology and glean lessons for forward use, it is tested on modern earthquakes with known seismic source models. *First*, the inversion methodology is applied to liquefaction induced by three earthquakes in Canterbury, New Zealand, utilizing the extensive CPT-based liquefaction dataset resulting from these events. This is followed by further testing on 5 global earthquakes with fewer liquefaction case histories. *Second*, the methodology is tested on landslides from four earthquakes in California. These simulated analyses illustrate the proposed methodology's provocative potential, while also highlighting caveats and lessons for forward use. Chapter three summarizes the compilation and analysis of field data in two regions where seismic hazards are enigmatic. *First*, publicly available paleoliquefaction and felt-intensity data from the 1886 Charleston, SC earthquake are compiled and analyzed, thereby probabilistically constraining the seismic-source location, geometry, and magnitude, to the extent permitted by the field evidence. *Second*, all existing paleoliquefaction evidence in the PNW is compiled into a GIS database, resulting in data from 185 study sites (e.g., feature locations, types, sizes, and ages). Moreover, maps are developed – specific to the CSZ – that forecast liquefaction for 30 different simulations of an M_w9 CSZ event. These maps can be used to guide field explorations for new evidence, such that they are conducted efficiently and strategically, considering the apparent utility of evidence toward constraint of CSZ ground-motion models. Of additional utility, the resulting maps of expected shaking intensity and liquefaction may be useful in downstream hazard modelling, regional loss estimation, policy development, and science communication. *Third*, CPT-based liquefaction case histories from the 2001 Nisqually, WA earthquake are compiled to test geotechnical liquefaction models for subsequent use in the PNW. While this event induced liquefaction in several urban centers, and although the value of liquefaction case-history data to model training and testing is widely recognized, no such data has previously been published from this event. Accordingly, using existing CPTs and reconnaissance reports, a digital dataset of 24 liquefaction case histories is compiled and analyzed. *Fourth*, liquefaction is predicted in an M_w9 CSZ rupture at select sites using subsurface test data and more advanced geotechnical methods for predicting profile response. Predictions are made for each of the 30 ground-motion simulations at 400 locations in the PNW where CPT data is available. These predictions are compared to the paleoliquefaction evidence compiled in (5) to assess which of the 30 simulations represent more and less plausible realizations of the motions experienced in 1700. This comparison suggests that all simulations predict far more liquefaction than has been found and that shaking intensities in 1700 may have been less in inland hubs (e.g., Seattle, Portland, and Vancouver), than prevailing forecasts indicate. The uncertainties associated with this possibility are discussed in detail.

Chapter 2. DEVELOPMENT AND TESTING OF AN INVERSION METHODOLOGY TO PROBABILISTICALLY CONSTRAIN SEISMIC SOURCE PARAMETERS FROM PALEOSEISMIC EVIDENCE

A novel, flexible inversion methodology that probabilistically constrains seismic-source parameters from paleoseismic evidence, and which greatly improves upon prior approaches, is proposed and tested. These efforts are summarized in the following three publications.

2.1 Publication 1: Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress* 2022.

- This conference paper introduces the proposed methodology and demonstrates how it can be used to probabilistically constrain rupture locations and magnitudes from paleoseismic evidence (in this case, from liquefaction and landslides). The methodology is tested on two earthquakes, yielding preliminary results.
- Past inverse-analyses of paleoseismic evidence have several shortcomings, namely, they: (i) require the location of an earthquake's source to be known, even though this is typically unknown for a paleoearthquake; (ii) tend to provide only a lower-bound estimate of seismic parameters, and (iii) are generally deterministic in nature, even though the analysis of paleoseismic data involves numerous uncertainties.
- Accordingly, the proposed methodology: (i) inherently constrains the earthquake source, (ii) provides an interpretable estimate of seismic parameters (e.g., provides a probability density

function (PDF) of magnitude), and (iii) is probabilistic in nature, allowing for any and all uncertainties to be accounted for in a complete and transparent manner.

- The methodology utilizes likelihood analysis to constrain an earthquake's source parameters from a near-infinite number of possibilities. Specifically, seismic parameters are identified by analyzing the likelihood that a rupture with given location, geometry, and magnitude would produce a series of field observations (sites with and without evidence), wherein various measurement and modelling uncertainties may be included. In general, the likelihood of a parameter having a value, given a set of observations, is the product of the probabilities of those observations, conditioned on the parameter value. Thus, the likelihood of a rupture having some location (L), geometry (G), and magnitude (M_w), given a set (x) of field observations at N different sites, is:

$$Likelihood\ (L, G, M_w|x) = P(X = x|L, G, M_w) = \prod_{i=1}^{N} P(X_i = x_i|L, G, M_w) \qquad (1)$$

where $P(X_i = x_i|L, G, M_w)$ is the probability of the observation at site i (the observed presence or absence of field evidence) given an earthquake with parameters L, G, and M_w. By repeating for a near-infinite number of possibilities, the most likely, actual rupture parameters are probabilistically constrained via the likelihood function (product of the probabilities of N observations), such that different combinations of L, G, and M_w will be found more and less likely to produce the observed field evidence. In Equation 1, the probability of a field observation, conditioned on all input variables, is computed at sites with and without field evidence using Equations 2 and 3, respectively:

$$P(\text{Field Evid}|\text{EQK: } L, G, M_w) = \int_{IM} P(\text{Field Evid}|IM, M_w)\, f(IM|L, G, M_w) \cdot \mathrm{d}IM \qquad (2)$$

$$P(\text{No Field Evid}|\text{EQK: } L, G, M_w) = 1 - \int_{IM} P(\text{Field Evid}|IM, M_w)\, f(IM|L, G, M_w) \cdot \mathrm{d}IM \; (3)$$

where $f(IM|L, G, M_w)$ is a probability density function (PDF) of an intensity measure (IM) computed by a ground motion model (GMM) considering site response, and conditioned on fault parameters L, G, and M_w; and $P(\text{Field Evid}|IM, M_w)$ is the probability of observing field evidence at a site, given ground-motion parameters IM and M_w, as computed by a hazard model (e.g., a liquefaction, landslide, or rockfall prediction model).

- To test the inversion methodology for paleoliquefaction and paleolandslide evidence, simulated tests on two modern earthquakes were performed. These simulated paleoseismic analyses considered only the paleoseismic evidence left behind and assumed no knowledge of rupture location, geometry, or magnitude. In addition, because the two earthquakes are modern, the number of study sites available from reconnaissance (i.e., sites with and without observed evidence) far exceeds the number of field sites conceivable of a regional paleoseismic study. Accordingly, analyses are performed on a random sample of 100 study sites from each event (i.e., 50 sites with evidence and 50 sites without evidence).
- CPT-based liquefaction case histories from the 2011 M_w6.3 Christchurch, New Zealand, earthquake were analyzed using the inversion methodology.
- Ground motion intensities were predicted by a regionally appropriate ground motion model (GMM), while liquefaction triggering and surface manifestation were predicted by the Boulanger and Idriss (2014) and Geyin and Maurer (2020) models, respectively.
- Select results are in Figures 4 and 5a, which respectively show contours of epicenter likelihood and the PDFs of inverted M_w. Figure 4 shows the most likely source location (as given by

contours of highest likelihood), which coincides with the actual surface projection of the fault responsible for the earthquake (i.e., the geolocation error, the distance between the location of highest likelihood and the surface projection of the actual fault, is zero). In addition, the analysis indicates that other known mapped faults in the region have near-zero likelihood. Figure 5a provides two PDFs of M_w including: (i) the M_w distribution conditioned only on the most likely source (i.e., the PDF of M_w at the location of highest likelihood); and (ii) the M_w distribution obtained if source-location uncertainty is included (i.e., the overall PDF of M_w obtained from aggregating PDFs across all potential source locations). As shown in Figure 5a, when conditioned on the most likely source, the 95% confidence interval (CI) is M_w6.1-M_w6.65 and the median M_w is M_w6.4, very close to the actual M_w6.3. When considering all potential sources, the 95% CI is M_w6.1-M_w7.9 with a median of M_w6.8. Unless there is on-fault or other geologic evidence to constrain the causative fault, the latter PDF would be the more appropriate product of a paleoseismic analysis.

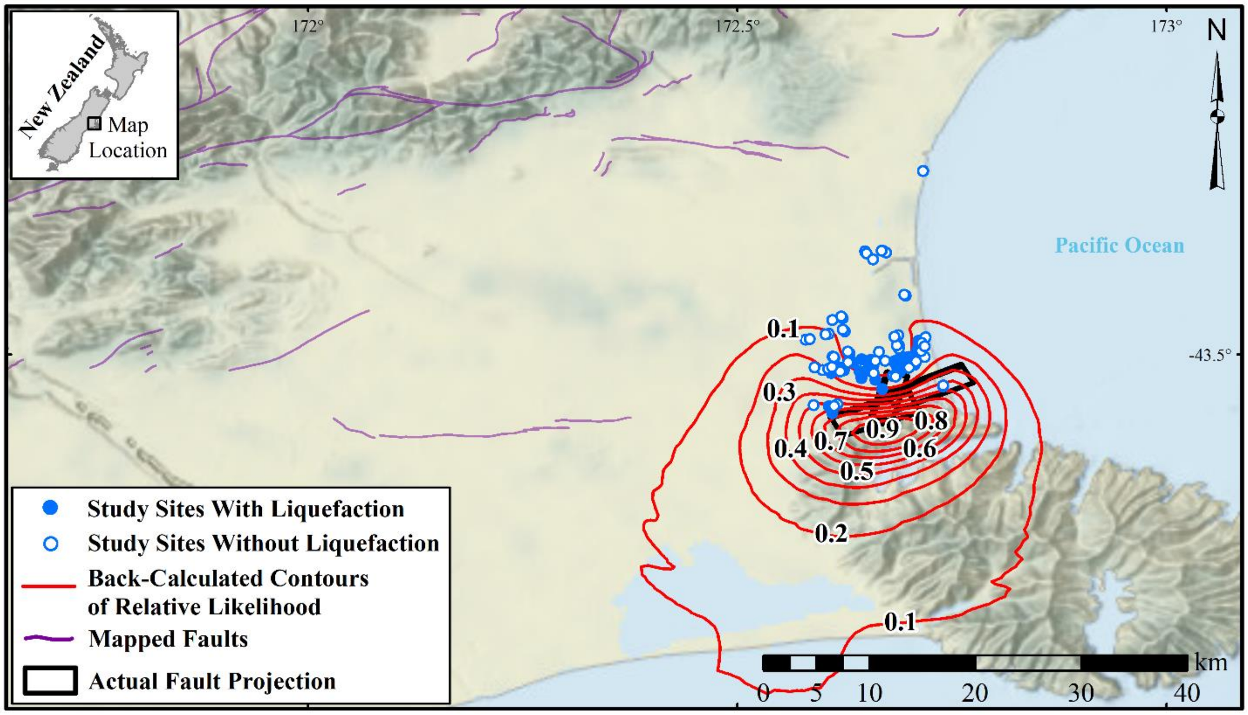


Figure 4. Representative analysis using 100 random study sites from the 2011 $M_w 6.3$ Christchurch earthquake showing the relative likelihood contours of earthquake source locations. Black rectangles delineate the actual fault projection (Beavan et al. 2012); faint purple lines are other mapped faults in the region (Langridge et al. 2016).

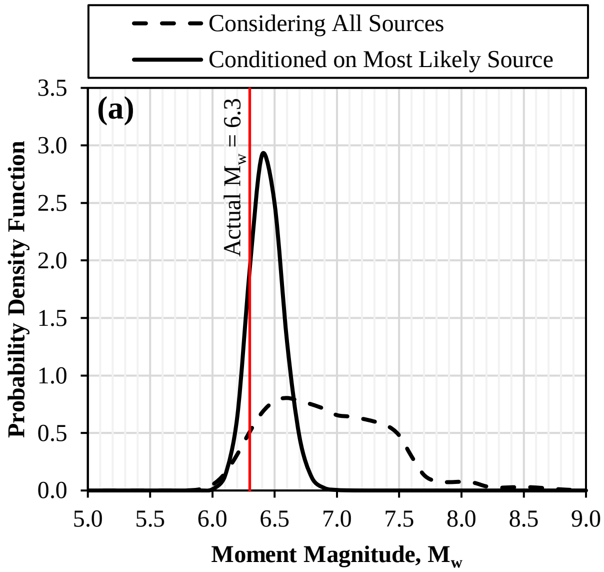


Figure 5. PDFs of earthquake M_w inverted from: (a) liquefaction evidence in the 2011 M_w6.3 Christchurch earthquake; and (b) landslide evidence in the 1971 M_w6.6 San Fernando earthquake. Shown are PDFs conditioned on the most likely source location (solid line) as well as the aggregate PDFs considering source-location uncertainty (dashed lines).

- Landslide case histories from the 1971 San Fernando, California, earthquake were analyzed. Ground motion intensities were predicted by a regionally appropriate GMM, while landsliding was predicted by the Nowicki et al. (2018) model.
- Select results are in Figures 5b and 6. As shown in Figure 6, the analysis identified a relatively well-constrained area as the likely source, aligning with the actual surface projection of the rupture (i.e., geolocation error = zero). Figure 5b shows the M_w PDFs produced by the analysis. When conditioned on the most likely source, the 95% CI is M_w6.25-M_w6.80 and the median M_w is M_w6.6, matching the actual M_w6.6 of the event. When considering all potential source locations (i.e., considering source location uncertainty), the 95% CI is M_w6.25-M_w6.9 with a median of M_w6.6.

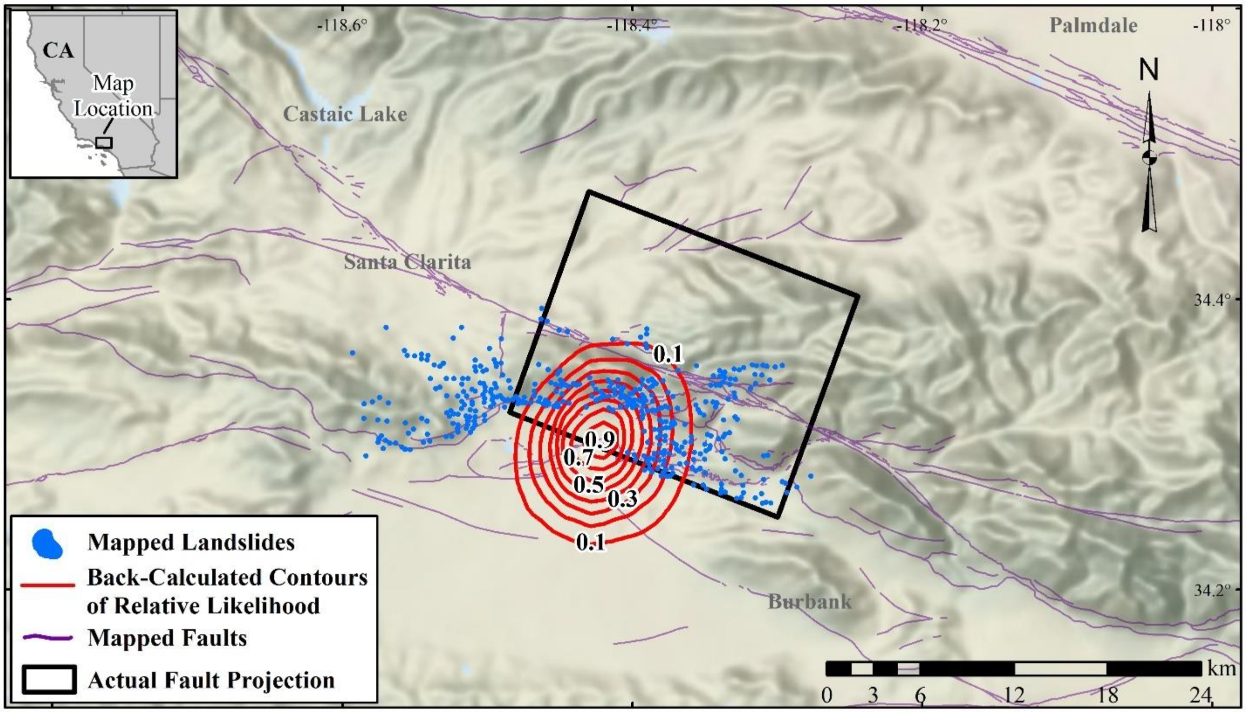


Figure 6. Representative analysis using 100 random study sites from the 1971 M_w6.6 San Fernando, California earthquake showing the relative likelihood contours of earthquake source locations. Black rectangles delineate the actual fault projection (Carena and Suppe, 2002); faint purple lines are other mapped faults in the region (California Geological Survey, 2014).

- This study introduced a methodology to probabilistically constrain seismic source parameters from regional liquefaction and landslide evidence. This methodology results in: (i) a geospatial likelihood surface that constrains the rupture location; and (ii) a probability distribution of the rupture magnitude.
- The simulated paleoseismic studies demonstrate the methodology’s potential, but further tests are needed. These tests, and the lessons learned, are discussed in Publications 2 and 3.

2.2 PUBLICATION 2: RASANEN, R.A. AND MAURER, B.W. (2022). "PROBABILISTIC SEISMIC SOURCE LOCATION AND MAGNITUDE VIA INVERSE ANALYSIS OF PALEOLIQUEFACTION EVIDENCE." *EARTHQUAKE SPECTRA*, DOI: 10.1177/87552930211056355, EERI.

- Further simulated paleoliquefaction analyses were performed on seven additional earthquakes. Results from two of these events are shown to highlight extremes in the methodology's performance.
- The most important models used within the methodology are the GMM, liquefaction triggering model, and liquefaction manifestation model. In this study, regionally appropriate GMMs (used to compute the distribution of peak ground acceleration (PGA) at study sites) are adopted. These include the Bradley (2013) model for New Zealand events and the Chiou and Youngs (2014) model for all others. Liquefaction triggering is predicted using the CPT-based Boulanger and Idriss (2014) model and liquefaction manifestation is predicted using the Iwasaki et al. (1978) liquefaction potential index (LPI) in conjunction with the fragility functions of Geyin and Maurer (2020), which relate the probability of surficial liquefaction manifestation to LPI. However, because the methodology is modular, other models (including those developed in the future) can be substituted as appropriate.
- 46 CPT-based liquefaction case histories (50% with liquefaction manifestation and 50% without) from the 1989 M_w6.9 Loma Prieta, California earthquake were analyzed. Contours of source-location likelihood are given in Figure 7a. In this case, the identified, most likely source was ~30 km north of the actual rupture; moreover, the median magnitude was M_w7.9, an overestimate of M_w1.0. These mispredictions are attributable to unusual site-response effects in the San Francisco Bay Area, which are variable and difficult to predict without site-specific

ground motion analyses. In particular, the northernmost study sites (e.g., in the environs of Alameda and Richmond) experienced PGAs far exceeding predictions by the Chiou and Youngs (2014) GMM. That is, liquefaction was not expected at these sites given GMM predictions of the actual rupture, causing the inverted source to be larger in magnitude and more northward than actual. To confirm, a revised analysis using site-specific ground motion amplification factors was performed. This reanalysis, shown in Figure 7b, resulted in a geolocation error of 0 km. The median M_w at the most likely source location was M_w6.9 with 95% CI of M_w6.6–M_w7.3, as shown in Figure 8a. The overall CDF considering source-location uncertainty resulted in a similarly accurate but slightly wider distribution, having a median of M_w7.05 and 95% CI of M_w6.5–M_w7.7. This demonstrates the potential benefit of performing site-specific ground motion analyses in paleoliquefaction studies (i.e., shear-wave velocity (Vs) profiling and site-response analysis), rather than relying solely on empirical GMMs.

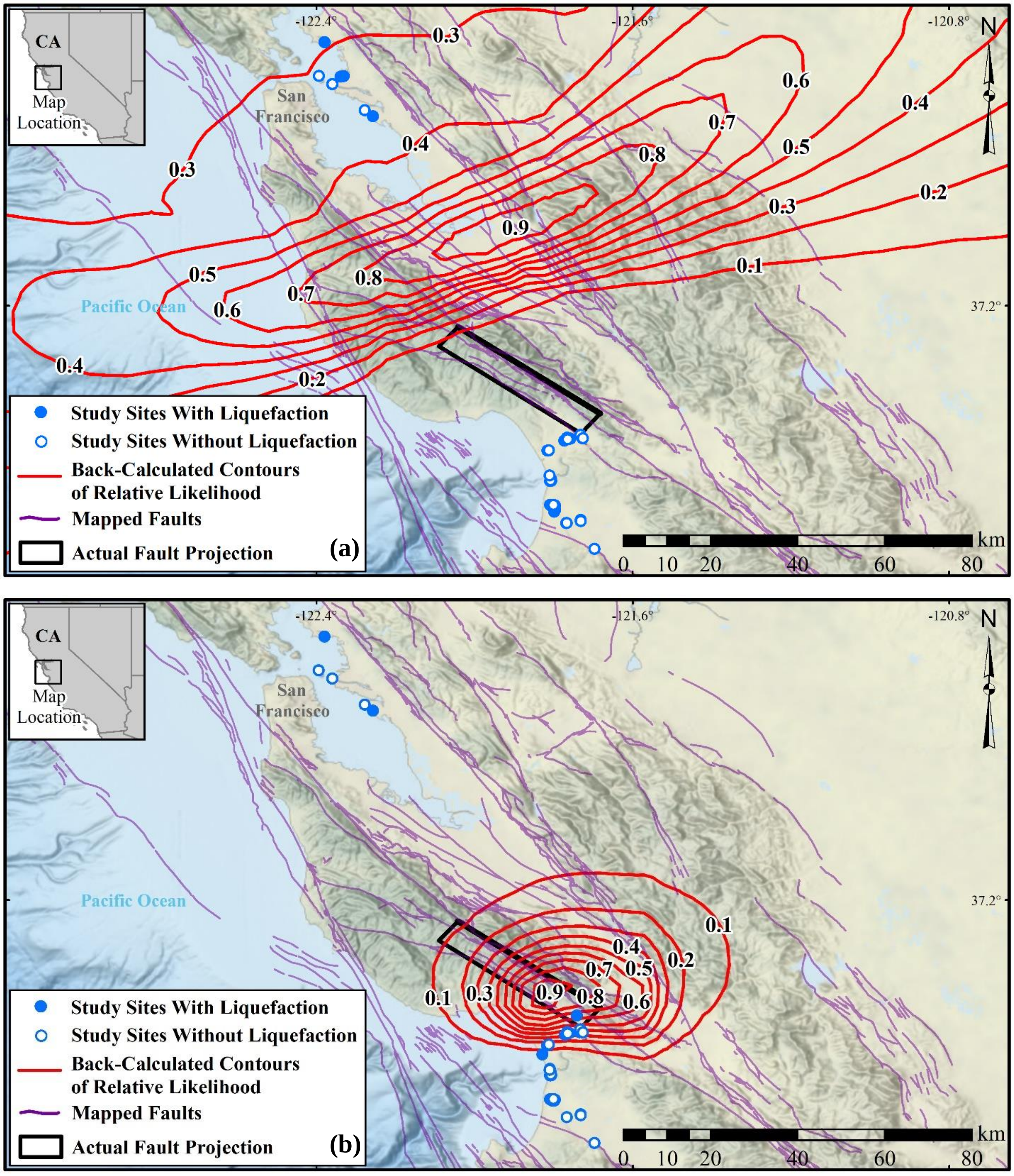


Figure 7. Analysis of 46 study sites from the 1989 M_w6.9 Loma Prieta, California earthquake: (a) contours showing the relative likelihood of source locations; (b) contours showing the relative likelihood of source locations, after inclusion of site-specific ground motion amplification factors. Black rectangles delineate the actual fault projection (Wald et al., 1991); faint purple lines are other mapped faults in the region (California Geological Survey, 2014).

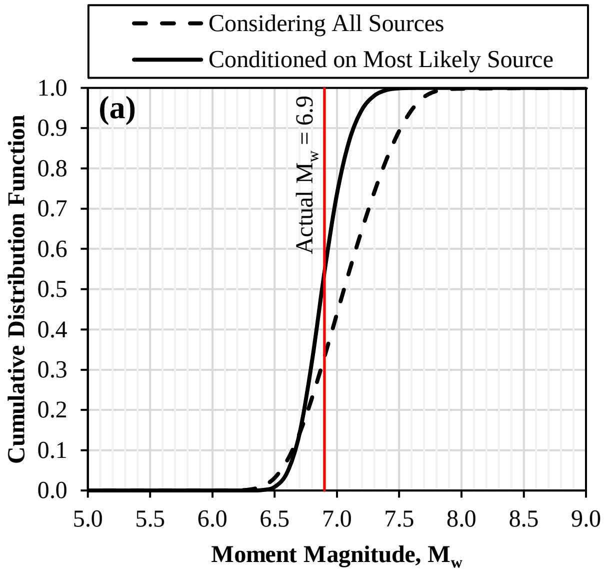


Figure 8. CDFs of earthquake M_w inverted from liquefaction evidence in the: (a) 1989 M_w6.9 Loma Prieta, California earthquake, after inclusion of site-specific ground motion amplification factors; and (b) 1979 M_w6.5 Imperial Valley, California earthquake. Shown are CDFs conditioned on the most likely source location (solid line) as well as the aggregate CDFs considering source-location uncertainty (dashed lines).

- 6 CPT-based liquefaction case histories (50% with liquefaction manifestation and 50% without) from the 1979 M_w6.5 Imperial Valley, California earthquake were analyzed. Contours of source-location likelihood are given in Figure 9. The inversion identified a relatively small but weakly constrained source area, with the most likely source falling near the fault projection (geolocation error = 2 km). The corresponding median magnitude, shown in Figure 8b, was M_w6.1 with a 95% CI of M_w5.2–M_w7.3, versus the actual M_w6.5. The Imperial Valley event thus demonstrates that seismic source parameters may be accurately predicted even when few study sites are available. However, as evident from Figure 9, earthquakes located elsewhere throughout the study area also have appreciable potential to produce the field observations. As a result, the overall CDF considering source-location uncertainty has a median of M_w8.0 and 95% CI of M_w6.1–M_w8.8. This large uncertainty should not be interpreted as a failing of the methodology, but rather, as an objective quantification of the actual uncertainty, given the limited quantity of field data. The potential benefit of coupling the results in Figure 9 with additional geologic evidence and interpretation is apparent, given the large uncertainty that ensues when the source location is weakly constrained.

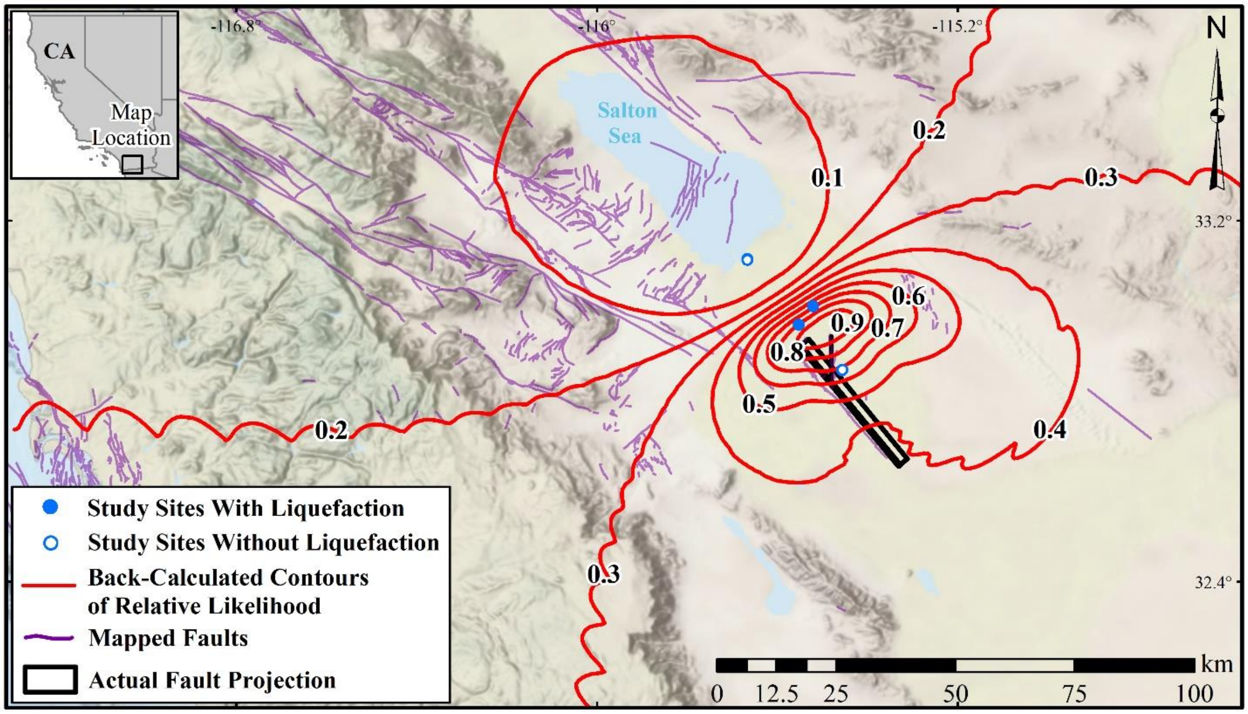


Figure 9. Analysis of 6 study sites from the 1979 M_w6.5 Imperial Valley, California earthquake showing contours of relative likelihood of source locations. Black rectangles delineate the actual fault projection (Archuleta, 1984); faint purple lines are other mapped faults in the region (California Geological Survey, 2014).

- In total, eight simulated paleoliquefaction analyses were performed on modern earthquakes. These analyses produced several lessons for forward use, including:
    - (i) Accurate results can be obtained using only a few study sites, albeit with larger uncertainty.

- (ii) Study sites should be regionally distributed. Similar to the geolocation of epicenters using seismometers, which relies on seismic waves having different arrival times, the proposed inversion methodology relies on waves having different arrival amplitudes. As amplitudes become more similar at two sites, it becomes more difficult to determine the direction from which motions are propagating, and thus more difficult to locate the seismic source. Similar arrival amplitudes can arise at sites with similar site-to-source distances or at sites relatively far from the fault rupture, where the rate of ground-motion attenuation is less, complicating detection of source direction.
- (iii) As a modular methodology, performance depends on component models to predict ground motions (e.g., GMM and site response model) and liquefaction response (e.g., liquefaction triggering and manifestation models). The inaccuracy of any one of these models can lead to poor performance. However, because the methodology is modular, performance should improve over time as newer models are developed. In addition, the modularity allows the methodology to be transferrable to other types of evidence, such as landslides and felt intensity (see publications 3 and 4, respectively).

- Given a set of field observations, the inversion methodology provides an objective and transparent quantification of uncertainty, leading to an interpretable estimate of M_w. In this regard, the methodology accounts for the uncertainty of the source location and the uncertainty of ground motions, conditioned on that source location. However, additional uncertainties not accounted for (e.g., geologic field interpretation) could also be incorporated via the total probability theorem.

- While potential earthquake sources were treated as epicenters (and then converted to statistical fault realizations), the proposed methodology could directly assess the plausibility of three-dimensional faults (e.g., with differing length, orientation, and depth – See Publication 4).
- The probabilistic characterizations of source location and M_w are direct inputs to probabilistic seismic hazard analyses. Therefore, the proposed methodology has the potential to provide new insights in enigmatic seismic zones worldwide (see Publication 4).

2.3 Publication 3: Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419, DOI: 10.1007/s10346-021-01780-9, Springer.

- Continuing from Publication 1, further simulated paleolandslide analyses were performed on three additional modern earthquakes. Results from two of these events are shown to highlight extremes in the methodology's performance. Some of the lessons learned from analyzing landslides are similar to those gleaned from the study of liquefaction in Publication 2.
- The main component models used in this study include a regionally appropriate GMM (Chiou and Youngs, 2014) and a probabilistic landslide model (Nowicki et al., 2018), which is an empirical model that uses geospatial variables. While empirical hazard models like Nowicki et al. (2018) have outperformed mechanistic models in some settings (Geyin et al. 2020), it is generally assumed that the adopted model is not the most accurate available. However, this model allows for the proposed methodology to be readily implemented on multiple events, each with numerous study sites. In a paleolandslide study, wherein site or region-specific profiling and material shear-strengths may be available, an improved empirical or mechanistic landslide model could be employed (e.g., limit equilibrium, sliding block, or nonlinear dynamic analysis). In cases where the mode of landsliding may be discerned, mode-dependent models, which potentially offer improved prediction efficiency, could be used.
- 100 randomly selected landslide case histories (50% with landslides and 50% without) from the 1994 M_w6.7 Northridge, California earthquake were analyzed to produce the likelihood contours mapped in Figure 10. The analysis identified a relatively well-constrained area as the likely source ~5 km from the actual fault projection, with locations of up to 40% relative

likelihood within the projection (i.e., they are not unlikely). The median M_w at the most-likely source was M_w6.8 with a 95% CI of M_w6.5–M_w7.0, versus the actual M_w6.7, as given in Figure 11a. The overall CDF considering source-location uncertainty has a similar M_w distribution as the most likely source.

- For comparison, the magnitude-bound curve of Keefer (1984), which is an existing, popular method for paleolandslide analysis, was also used to constrain the M_w. Magnitude-bound curves use the site-to-source distance of the most distal feature (in this case, a landslide) to produce a lower-bound estimate of M_w. As commonly done in past studies, the assumption of the earthquake source being located at the centroid of the mapped landslides was made. Inputting the maximum site-to-source distance into the Keefer (1984) curves generates an inverted lower bound M_w5.65, well below the events actual M_w6.7.

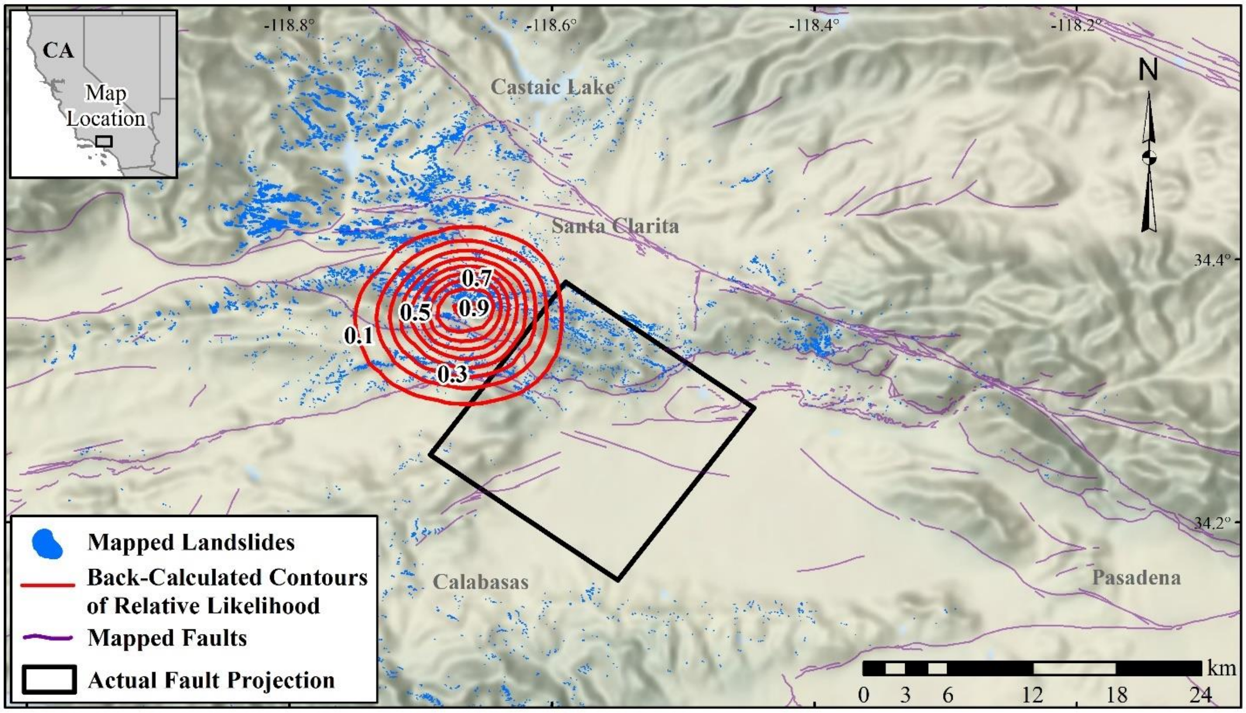


Figure 10. Representative analysis using 100 random study sites from the 1994 $M_w 6.7$ Northridge, California earthquake showing the relative likelihood contours of earthquake source locations. Black rectangles delineate the actual fault projection (Updike et al., 1996); faint purple lines are other mapped faults in the region (California Geological Survey, 2014); blue polygons are the complete inventory of mapped landslides.

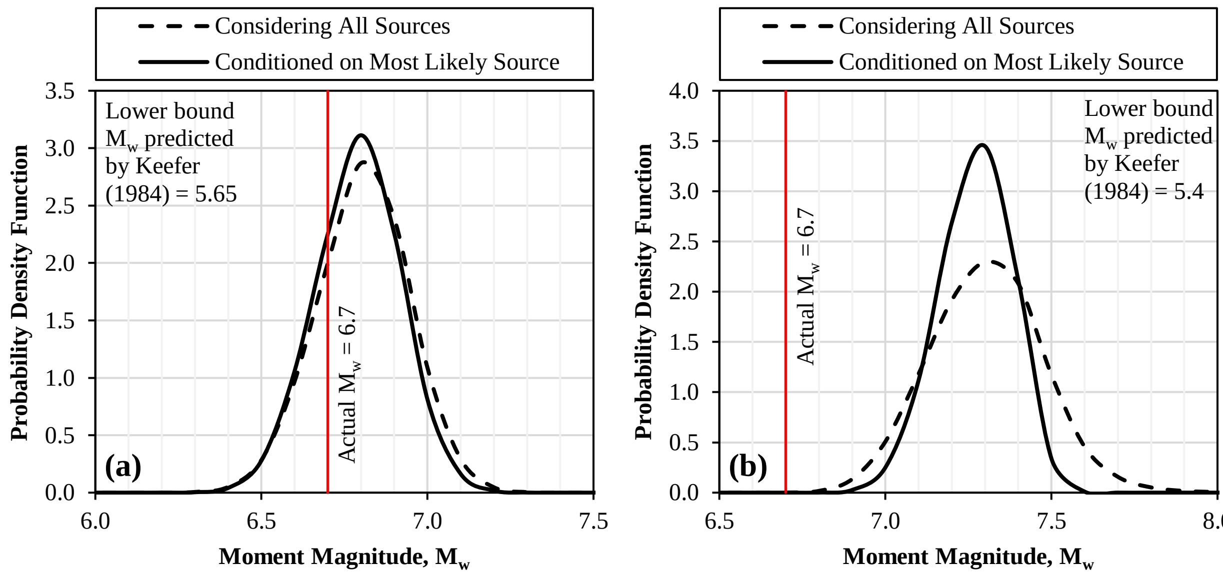


Figure 11. PDFs of earthquake M_w inverted from regional landslide evidence in the: (a) 1994 M_w6.7 Northridge, California earthquake; and (b) 1983 M_w6.7 Coalinga, California earthquake. Shown are PDFs conditioned on the most likely source location (solid line) as well as the aggregate PDFs considering source-location uncertainty (dashed lines).

- 100 randomly selected landslide case histories (50% with landslides and 50% without) from the 1983 M_w6.7 Coalinga, California earthquake were analyzed to produce the likelihood contours in Figure 12. The inversion accurately located the source, with the identified, most-likely source falling within the actual fault projection. However, the median inverted M_w here was M_w7.3 with 95% CI of M_w7.0–M_w7.4, as shown in Figure 11b. In contrast, the actual magnitude, M_w6.7, was deemed to have near-zero probability of producing the evidence field. Aggregating conditional PDFs resulted in a similar overall PDF with median M_w7.3 and 95% CI of M_w6.9–M_w7.6, versus the actual M_w6.7. Thus, the inversion methodology significantly overpredicted the rupture magnitude, despite accurately predicting its location.
- The discrepancy in M_w could have one or more causes: (i) the epicentral grid points converted to fault realizations are not an accurate representation of the true source; (ii) the landslide

evidence is constrained by topography, which can make it more difficult to detect the direction and rate of ground-motion attenuation; and (iii) the inversion methodology relies on accurate predictions of ground motions and landslides. The methodology can thus be expected to perform poorly when one or more component models perform poorly.

- Nonetheless, had the analysis been based on the Keefer (1984) magnitude-bound curve, a lower-bound estimate of M_w5.4 would result, providing an equally poor constraint.

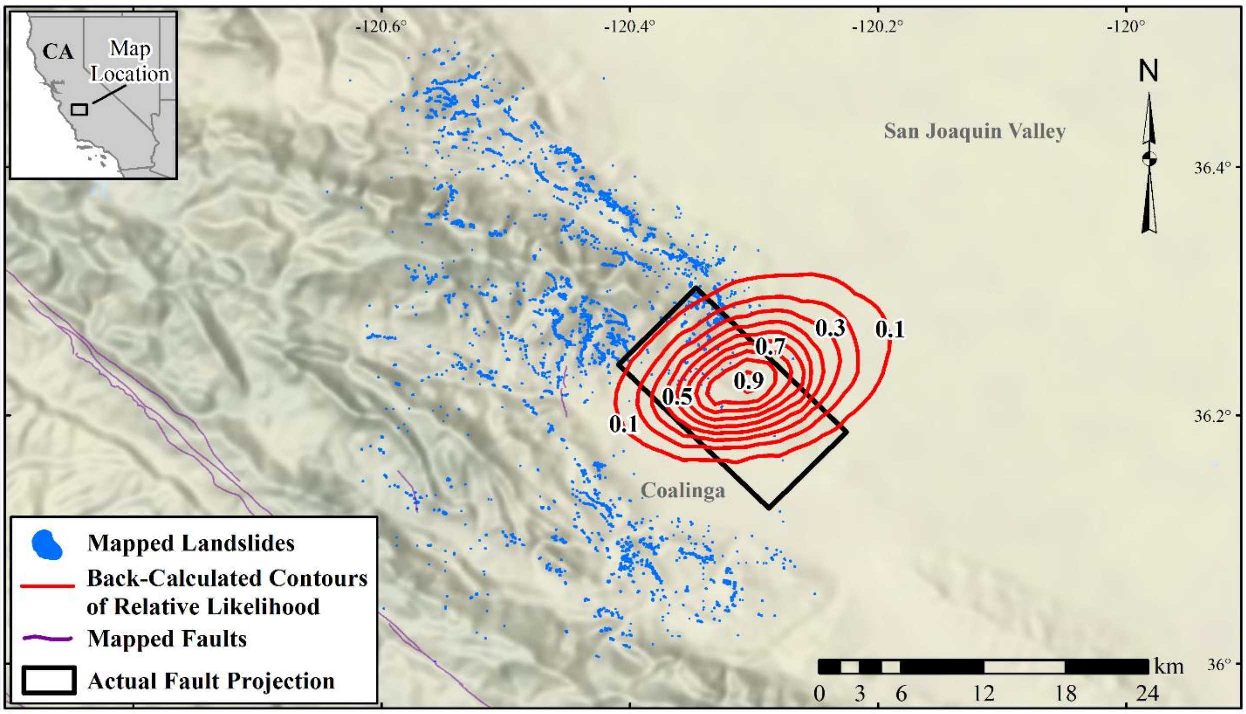


Figure 12. Representative analysis using 100 random study sites from the 1983 M_w6.7 Coalinga, California earthquake showing the relative likelihood contours of earthquake source locations. Black rectangles delineate the actual fault projection (Wentworth and Zoback, 1990); faint purple lines are other mapped faults in the region (California Geological Survey, 2014); blue polygons are the complete inventory of mapped landslides.

- The four earthquakes in this study were all in California, USA and had similar magnitudes. Thus, these earthquakes do not span the range of geologic, seismologic, and climatic settings found on Earth. However, because the methodology is modular in nature, it can accommodate alternative models (e.g., GMMs and landslide prediction models) as needed in different regions or given the availability of data (e.g., site-specific information). Most importantly, the methodology remains conceptually valid for all earthquakes in all locations, so long as ground motions and landslides are predictable.

- It was shown that the seismic-source location and magnitude can be probabilistically constrained from landslide evidence. However, in events where the evidence field is constrained by topography and/or far from the source, it may be more difficult to detect the direction and rate of ground–motion attenuation and therefore more difficult to constrain the source parameters. In this regard, study sites should be spatially distributed to maximize the potential for well constrained rupture parameters.
- Focusing specifically on differences between landslides and liquefaction, potential issues that could hinder the methodology's application to landslides are: (i) topographic amplification effects, which can greatly increase IM values at landslide locations, but which are currently not considered by GMMs; (ii) the current paucity of discovered and dated paleolandslides; and (iii) the challenge of determining whether a paleolandslide was seismically induced. Nonetheless, tests of the inversion methodology demonstrate its ability to provide new insights in enigmatic seismic zones (see Publication 4).
- As with the study of liquefaction evidence, potential earthquake sources were treated as point sources and then statistically converted to fault realizations, such that the fault geometry is not explicitly considered. In the future (see publication 4), faults could be modeled more explicitly to determine whether a particular rupture geometry is more likely to produce an observed field of evidence, all else being equal.

Chapter 3. DATA COLLECTION AND FIELD APPLICATION

The analysis of a paleoearthquake inherently relies on paleoseismic evidence. However, this evidence is often discovered by an array of researchers and exists across numerous publications. Accordingly, Chapter 3 focuses on data compilation and field application toward the analysis of the: (i) 1886 Charleston, SC; and (ii) 1700 Cascadia Subduction Zone earthquakes.

3.1 PUBLICATION 4: RASANEN, R.A. AND MAURER, B.W. (2023). PROBABILISTIC SEISMIC SOURCE INVERSION OF THE 1886 CHARLESTON, SOUTH CAROLINA, EARTHQUAKE FROM MACROSEISMIC EVIDENCE: A MAJOR UPDATING." *ENGINEERING GEOLOGY, 312: 106958, 10.1016/J.ENGGEO.2022.106958, ELSEVIER.*

- This study applies the methodology developed and tested in Chapter 2 (see Publications 1, 2, and 3) on paleoliquefaction and felt-intensity evidence compiled from the 1886 Charleston, SC earthquake.
- Prior studies of Modified Mercalli Intensity (MMI) and liquefaction data from this event have several limitations. Namely, these studies tend to: (i) either be deterministic or account for uncertainties in an informal manner (e.g., it is often unclear what published uncertainty bounds represent and which uncertainties are, and are not, included); (ii) assume that the 1886 event was caused by a particular fault (i.e., the Woodstock Fault) without investigating the uncertainty of this assumption or the ability of the field data to constrain source traits beyond magnitude (i.e., fault location, geometry); and (iii) rely on models for predicting various phenomena (e.g., ground motions, site response, liquefaction response, MMI) that have since

been superseded or augmented (e.g., by the NGA East project). Collectively, existing studies of the MMI data have reported estimates of M_w6.4 to M_w7.8.

- Accordingly, this study presented probabilistic seismic-source inversions of the 1886 earthquake using the latest predictive models and the methodology discussed previously.
- This study also introduced a new "fault search" method, where seismic sources are explicitly modelled as faults with differing lengths and orientations. In previous studies, faults were modeled as point sources and then converted to statistical fault realizations using site-to-source distance correlations. While the former approach produces an "average" fault, given some magnitude (i.e., an average depth, length, and orientation), the latter allows for fault characteristics to be constrained more explicitly. It can be determined, for example, whether some fault strikes are more likely to produce a field of evidence than other fault strikes. In all, three separate analysis methods for representation of seismic sources were considered. These were: (i) the "epicenter search" (used in Publications 1, 2, and 3), (ii) the "fault search", and (iii) the Woodstock Fault analysis, wherein the source was assumed to be the Woodstock Fault proposed by Durá-Gómez and Talwani (2009a,b). Given that the source of the 1886, Charleston SC earthquake is unknown, the analyses summarized here are those of the "fault search"; analyses conditioned on the Woodstock Fault may be found in the complete paper.
- To study MMI evidence, the equation for probability of observation is modified as follows:

$$P(\text{obs}|\text{EQK}: L, G, M_w) = \int_{IM}\int_{MMI} P(\text{obs}|MMI)\, f(MMI|IM) f(IM|L, G, M_w) \cdot \text{d}MMI \cdot \text{d}IM \quad (4)$$

where $f(IM|L, G, M_w)$ is the probability of an IM value conditioned on fault parameters L, G and M_w, as computed by a GMM that considers site response; $f(MMI|IM)$ is the probability of an MMI value conditioned on the IM value, and possibly on other parameters, as computed by an IM-MMI model; and $P(\text{obs}|MMI)$ is the binomial probability that the predicted MMI is equal to the observed MMI. In this work, we consider predictions and observations to agree if within ± 0.5 MMI (e.g., predicted MMIs of 5.51 and 6.49 agree with an observed MMI of 6). Thus, we assign an implicit, uniform measurement uncertainty of ± 0.5 MMI, but do not otherwise model the uncertainty of reported MMIs. However, the uncertainties of source location, ground-motion intensity, and predicted MMI are each accounted for.

- This study used the Bakun et al. (2002) dataset of 1,034 intensity reports and four different IM-MMI models, such that the epistemic uncertainty of model selection is accounted for.
- The results of the "fault search" method using the Worden et al. (2012) (WEA12) IM-MMI model are shown in Figure 13. Faults more likely to produce the MMI observations have thicker line weight and are darker in color. In addition, results using the three other IM-MMI models are shown as single lines denoting the singular faults deemed most likely by each model. The WEA12 model produces a median M_w7.6 and tends to agree with the location and orientation of the proposed Woodstock Fault (Durá-Gómez and Talwani, 2009a,b), but the other three models produce median magnitudes ranging from M_w7.20 to M_w7.90 and are oriented perpendicular to the Woodstock Fault and located partly offshore. The CDFs of each model (when considering source-location uncertainty), an ensemble CDF of the models, and the current CDF assigned to the Charleston source in the National Seismic Hazard Model Project (NSHMP) (Peterson et al., 2014, 2020), are given in Figure 14. This ensemble CDF has a median M_w7.25 and 95% CI of M_w7.05-M_w7.85. Collectively, the results do not provide strong support for the hypothesized Woodstock Fault, which is found to be a relatively unlikely source of the field evidence. This is not to say the fault does not exist, but rather, that the position of the 1886 source model cannot be constrained with the available data and models.

In the absence of other evidence, other sources would warrant consideration. Nonetheless, the results generally uphold the magnitude-distribution assumed by the NSHMP, while also indicating that $M_w > 7.5$ events should merit consideration.

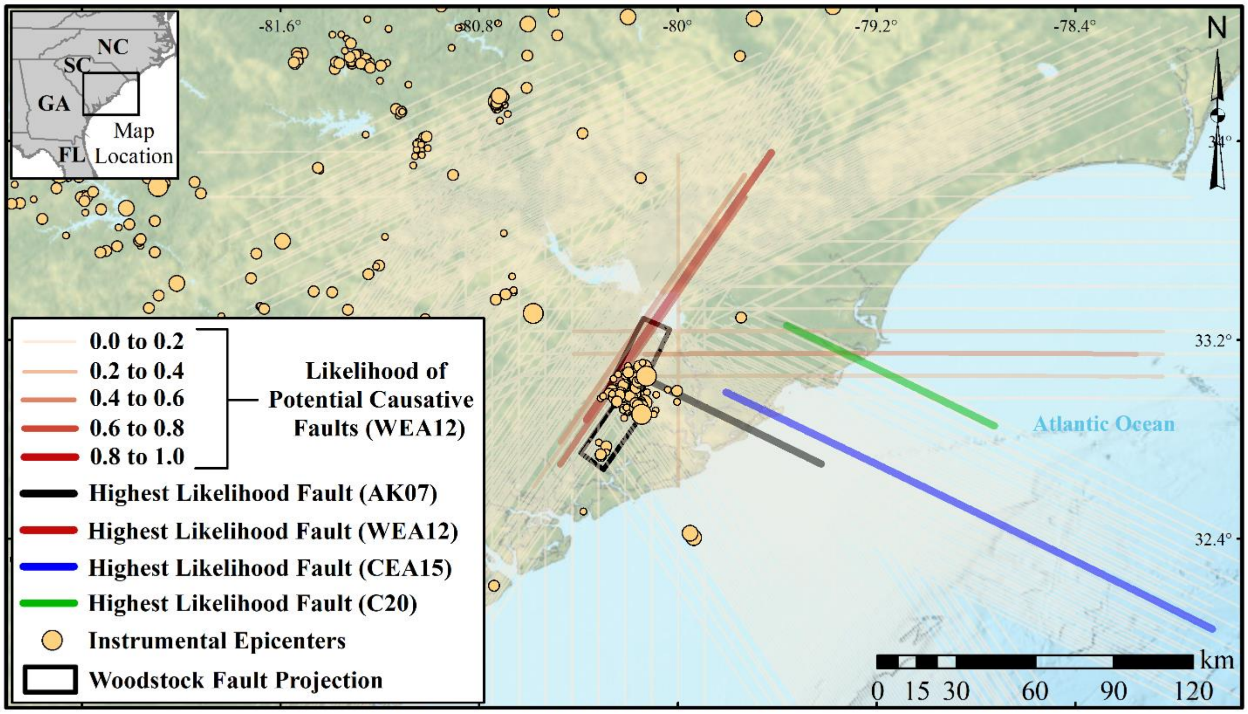


Figure 13. Likelihood of hypothetical faults with differing orientations and lengths to produce the MMI evidence, as obtained using the WEA12 model. Also shown are the singular faults deemed most likely using the AK07, CEA15, and C20 models; AK07 = Atkinson and Kaka (2007), WEA12 = Worden et al. (2012), CEA15 = Caprio et al. (2015), C20 = Cramer (2020); black rectangles = the Woodstock projection (Durá-Gómez and Talwani 2009a,b); tan circles = instrumental epicenters (M_w>1.0; 1900-present) (USGS, 2022).

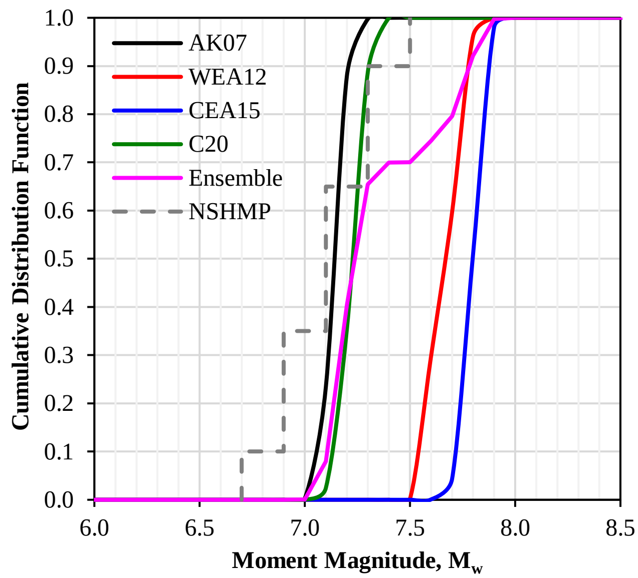


Figure 14. CDFs of earthquake M_w inverted from MMI observations using four IM-MMI models, wherein seismic sources are modelled as faults. AK07 = Atkinson and Kaka (2007), WEA12 = Worden et al. (2012), CEA15 = Caprio et al. (2015), C20 = Cramer (2020), and NSHMP = National Seismic Hazard Model Project (Peterson et al., 2014, 2020).

- This study also analyzed 24 sites where the 1886 liquefaction response was documented (50% with liquefaction manifestation and 50% without). Three separate models to account for soil aging effects, which can alter the expected response of Pleistocene soils, were adopted and ultimately ensembled, thereby accounting for the epistemic uncertainty of model selection.
- Results of the "fault search" method are shown in Figure 15. Regardless of whether soil aging is accounted for, the most likely fault location is SE of the proposed Woodstock Fault (Durá-Gómez and Talwani (2009a,b) and has a strike nearly perpendicular to that proposed. Shown in Figure 16 are the CDFs of M_w inverted from the liquefaction data. The large uncertainty of a M_w inferred from the liquefaction evidence is apparent, as is the influence of soil aging and its correction. To account for the uncertainty of selecting aging-correction models, each CDF was weighted to produce an ensemble CDF, as shown in Figure 16. This ensemble has a median

of M_w6.95 and 95% CI of M_w6.15 to M_w7.85, suggesting a similar median magnitude as the NSHMP (Petersen et al., 2014, 2020) but much greater uncertainty. Thus, similar to the analysis of MMI observations, the liquefaction evidence does not constrain the source of the 1886 earthquake to be the Woodstock Fault, although it does lend credence to the distribution of M_w currently considered by the NSHMP.

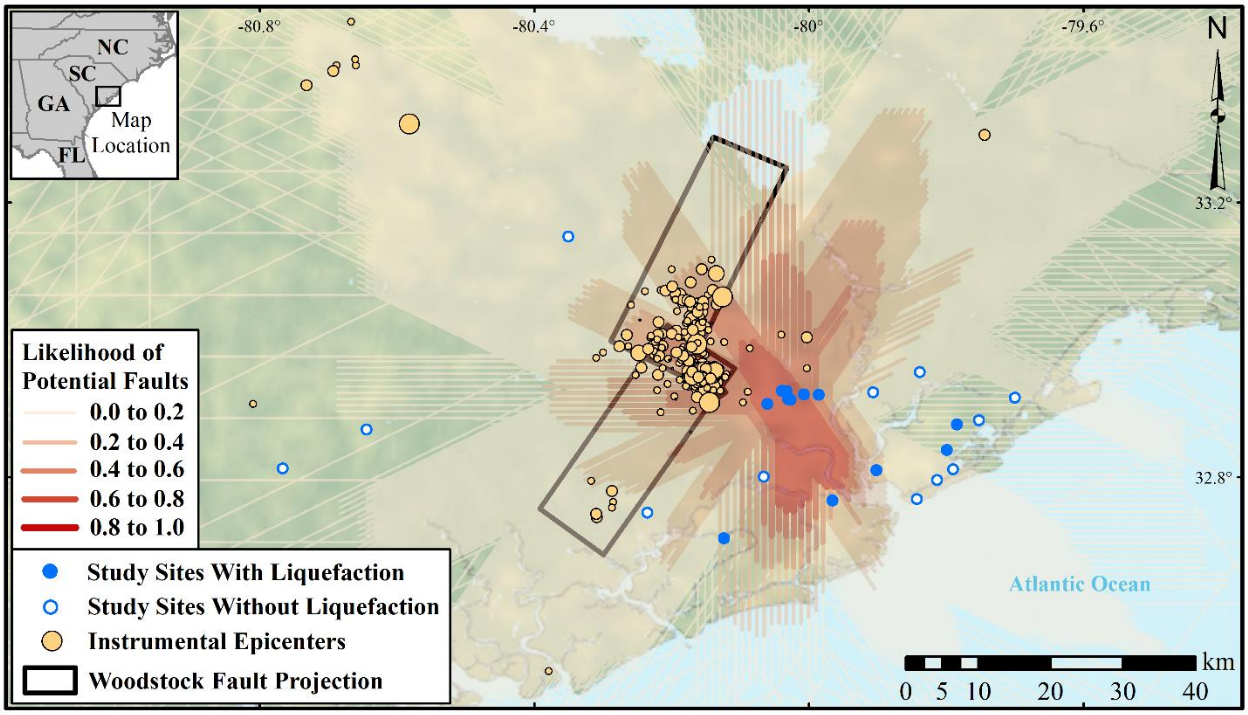


Figure 15. Likelihood of hypothetical faults with differing orientations and lengths to produce liquefaction evidence (no aging correction). Black rectangles = the Woodstock projection (Durá-Gómez and Talwani 2009a,b); tan circles = instrumental epicenters (M_w>1.0; 1900-present) (USGS, 2022).

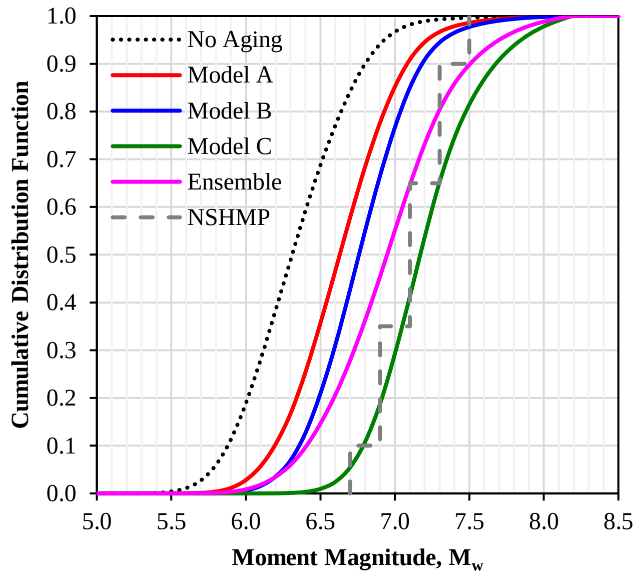


Figure 16. CDFs of earthquake M_w inverted from liquefaction evidence using three age correction models, wherein seismic sources are modelled as faults. NSHMP = National Seismic Hazard Model Project (Peterson et al., 2014, 2020).

- The results in this study do not consider every source of uncertainty. Omitted, for example, are the uncertainties of: (i) fault depth and length, (ii) fault dip, (iii) site V_{S30}, and (iv) aging-correction model uncertainty (for each of these inputs a median prediction was used). The uncertainties that are considered include: (i) source location uncertainty, (ii) ground motion uncertainty, (iii) epistemic uncertainty of IM-MMI and aging correction models, (iv) and the uncertainty of predicted MMI.
- This study finds the location and orientation of the 1886 fault rupture could not be confidently constrained by the available MMI and liquefaction evidence. Furthermore, the analyses do not show strong support for the hypothesized Woodstock Fault. However, this result may change if additional evidence were available for analysis, or if the component models were less uncertain.

- Collectively, the results largely support the M_w distribution adopted by the NSHMP, which ranges from M_w6.7-M_w7.5 with a median of M_w7.1. While analyses indicate that M_w < 6.7 ruptures have potential to produce the observed liquefaction response (particularly when the uncertainty of selecting an aging-correction model is considered), the MMI evidence suggests a near-zero likelihood of such magnitudes. Conversely, both types of evidence suggest that M_w > 7.5 events have potential to produce the field evidence. While this conclusion hinges on which component models are adopted to predict soil-aging effects and MMI, we fail to find conclusive evidence for outright rejecting some models in favor of others. In the absence of such evidence, M_w > 7.5 ruptures would merit consideration.

3.2 Publication 5: Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, DOI: 10.1016/j.enggeo.2021.106253, Elsevier.

- Publications 1, 2, and 4 established the importance of macroseismic evidence for the constraint of seismic parameters in regions of infrequent moderate-to-large earthquakes. The U.S. PNW is one such region. The Cascadia Subduction Zone (CSZ), for example, is inferred to have produced numerous M_w9 ruptures, yet the most recent of these (1700 CE) predates the historic record. While paleoliquefaction evidence is relatively abundant (in contrast to the paucity of landslide and felt intensity evidence), its documentation is spread across numerous publications by various researchers.
- Accordingly, this study: (i) reviews current paleoliquefaction inverse-analysis methods and their limited, prior applications in the PNW; (ii) compiles all PNW paleoliquefaction evidence from the literature into a community database, resulting in data from 185 study sites (e.g., feature locations, types, sizes, and ages); and (iii) develops maps – specific to the CSZ – that forecast paleoliquefaction for 30 different simulations of a CSZ interface earthquake.
- A comprehensive literature review reveals that inverse-analysis of paleoliquefaction in the U.S. PNW is limited. A few studies (e.g., Bourgeois and Johnson, 2001; Obermeier and Dickenson, 2000) utilized the magnitude-bound and site-specific geotechnical analysis methods, respectively, but these studies: (i) used models that have been significantly modified and improved over the past 20 years; (ii) analyzed relatively few paleoliquefaction sites; and (iii) were deterministic in nature.

- A more modernized, advanced analysis of paleoliquefaction in the PNW could thus help to elucidate the regional distribution of shaking intensity in past CSZ earthquakes (discussed more later in Publication 9)
- All publicly available paleoliquefaction evidence in the PNW was compiled from 24 publications. The information compiled from each reference includes (when available): study site locations and feature morphologies (e.g., sand-blow thickness and dike width); dating information; important comments from the author; and citations to all original source documents, data, and figures. In total, 185 study sites were compiled. This is the first database of its kind compiled for the U.S. PNW.
- Paleoliquefaction evidence was forecasted for 30 different physics-based ground motion simulations of an $M_w 9$ CSZ earthquake (Frankel et al., 2018; Wirth et al., 2018) where regional-scale wave-propagation based site-response analyses (Marafi et al., 2021) were used to account for expected near-surface site effects. The geospatial liquefaction model of Rashidian and Baise (2020), an update of the Zhu et al. (2017) model, was used to forecast paleoliquefaction on a regional scale. The Rashidian and Baise (2020) model allows for computation of a probability of liquefaction and liquefaction spatial extent (LSE, the percent area of ground covered by liquefaction manifestation) using free and readily available predictor variables. Both the probability of liquefaction and LSE are given in an online database (see Publication 6) for each of the 30 different simulations of a CSZ earthquake. Figure 17 maps the locations of compiled paleoliquefaction evidence. These data are superimposed with the computed LSE for one of the 30 $M_w 9$ CSZ simulations. This realization leads to some of the largest LSE values in the Fraser River delta of British Columbia. Notably, confirmed 1700 liquefaction has yet to be discovered anywhere in the vicinity, despite there being other, older

features at several study sites in the area, indicating that “csz013” may be an unlikely realization (although site-specific geotechnical analysis would need to be used to confirm this).

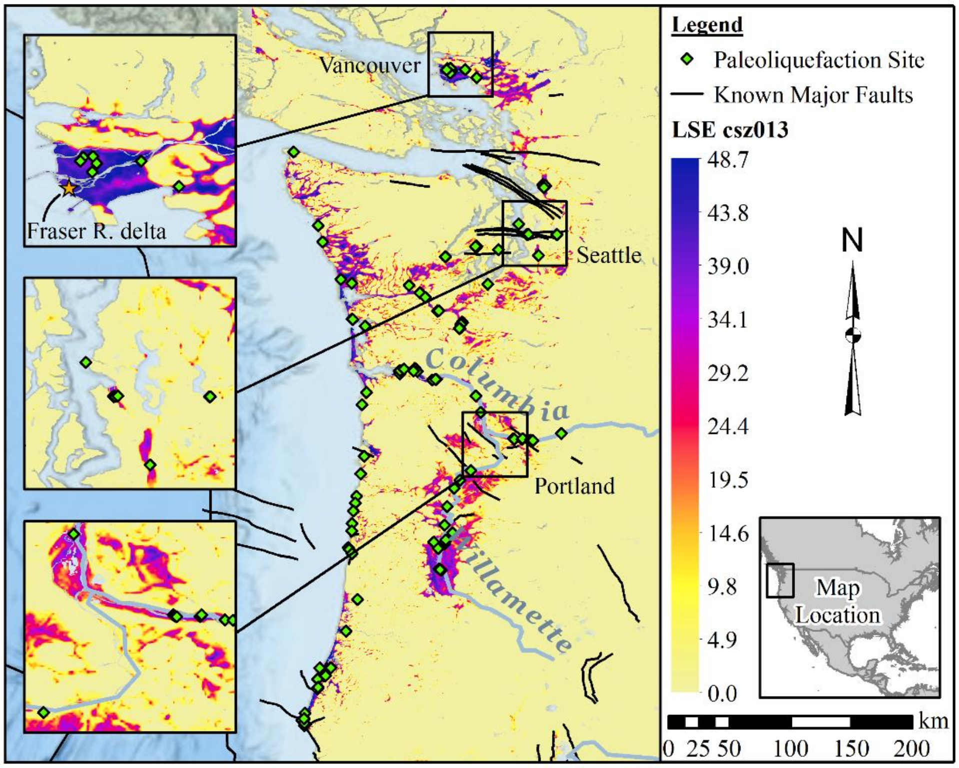


Figure 17. Forecasted liquefaction spatial extent, LSE (%) computed for CSZ M_w9 ground-motion simulation “csz013” developed by Frankel et al. (2018) and Wirth et al. (2018).

- Repeating the prior predictions for all 30 ground motion simulations, Figure 18 shows the 5-to-95 percentile variation in LSE across the simulations. Figure 18 can be used to identify where the presence or absence of liquefaction would have greater potential to constrain the intensity of past shaking (i.e., where differences in LSE are greatest). In these locales, the presence or absence of features would be strongly at odds with some subset of the ground motion realizations, thereby diminishing their plausibility. Elsewhere, where differences in LSE are small, the presence or absence of features would potentially do little or nothing to constrain ground-motion models. Specific to the three urban centers highlighted in Figure 18, relatively large LSE differences are predicted along the Columbia and Willamette Rivers in Portland, along the Duwamish River in Seattle, and in the Fraser River delta near Vancouver. In these urban areas, liquefaction manifestations are likely for some realizations of a CSZ rupture, but unlikely for others. Paleoliquefaction features have been documented in each of these urban areas, yet few, if any, of these were conceivably caused by the 1700 CSZ event. If these sites are judged not to have liquefied in 1700, the maximum shaking intensity could be provisionally constrained via geotechnical testing (i.e., if the intensity were any larger, liquefaction would be expected). Such analyses would also have the potential to inform regional ground-motion models, given that some subset of realizations would be less likely to represent that which occurred in 1700. Given the importance of geotechnical testing (e.g., CPTs) towards constraint of shaking intensity, an additional strategy is to search for the presence or absence of paleoliquefaction where testing has already been performed. That is, at locations where the subsurface is well documented, and where a significant cost of inverse analysis is already paid for. Towards this end, provided in the online database and mapped in Figure 18 are the locations of many hundreds of CPTs presently available from the Washington Dept. of Natural Resources (DNR). Many of these CPTs, the data for which can be obtained from Jeschke et al. (2019), will be analyzed as part of publication 9, wherein more advanced models will be used to predict liquefaction in CSZ events.

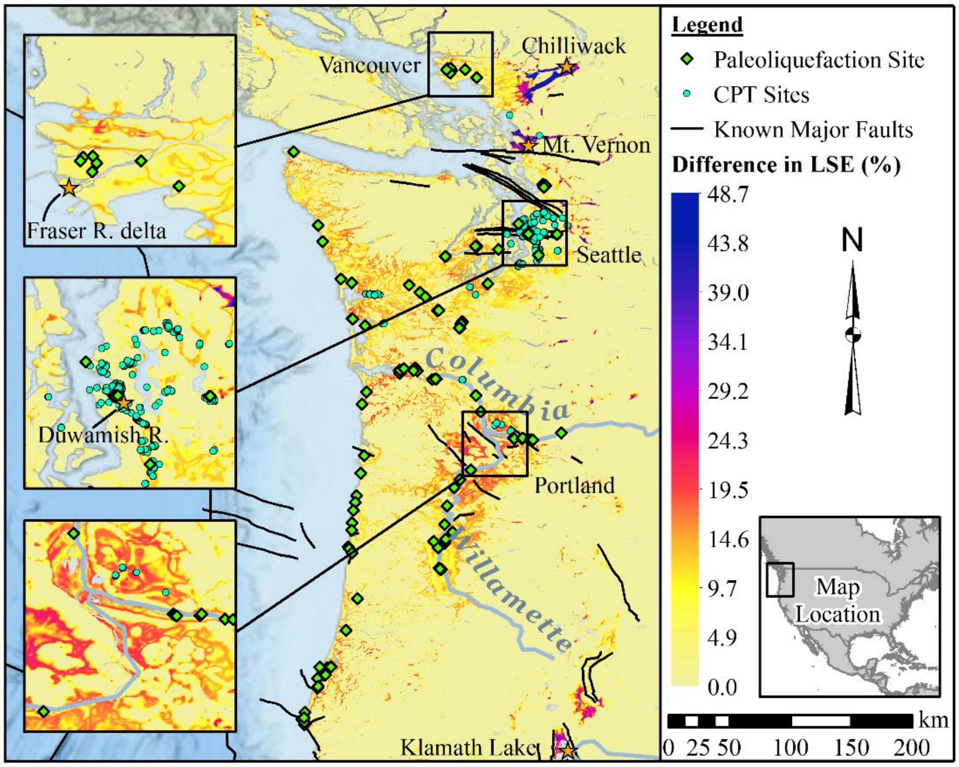


Figure 18. The 5-to-95 percentile variation in expected LSE, as computed from 30 different ground-motion simulations of a M_w9 CSZ earthquake; paleoliquefaction and CPT sites are also mapped.

- Collectively, these maps can be used to guide field explorations by engineering geologists for new evidence, considering both: (i) where evidence is likely to be found; and (ii) whether evidence is likely to provide meaningful constraint of CSZ ground motion models. Of additional utility, this process resulted in the first ever suite of physics-based M_w9 CSZ ground-motion predictions that account for the response of near-surface sediment. In making these predictions, equivalent linear site-response analyses were used to modify the Frankel et al. (2018) simulations for local, near surface conditions. Together, the maps of expected shaking intensity and liquefaction may be useful in regional loss estimation, scenario planning, and science communication. These maps, in addition to the paleoliquefaction database and Washington State CPTs, are given in the online database discussed in Publication 6.

- Ultimately, as more paleoliquefaction evidence is identified and analyzed, better constraint of regional ground-motions in past earthquakes will result (discussed more in Publication 9).

3.3 Publication 6: Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, DOI: 10.17603/ds2-fqkr-h615.

- This dataset includes all products developed in Publication 5. These products include:
    - All publicly available paleoliquefaction data in the U.S. PNW and southeastern British Columbia sourced from 185 total study sites documented in 24 publications.
    - The first ever suite of site-adjusted M_w9 CSZ ground-motion predictions and forecasted liquefaction impacts for the Cascadia Region. These predictions include PGA, PGV, the probability of liquefaction, and LSE.
    - Locations of CPTs from the Washington State Department of Natural Resources.
- An example figure from this online dataset is provided in Figure 19 and shows the locations of compiled paleoliquefaction evidence.

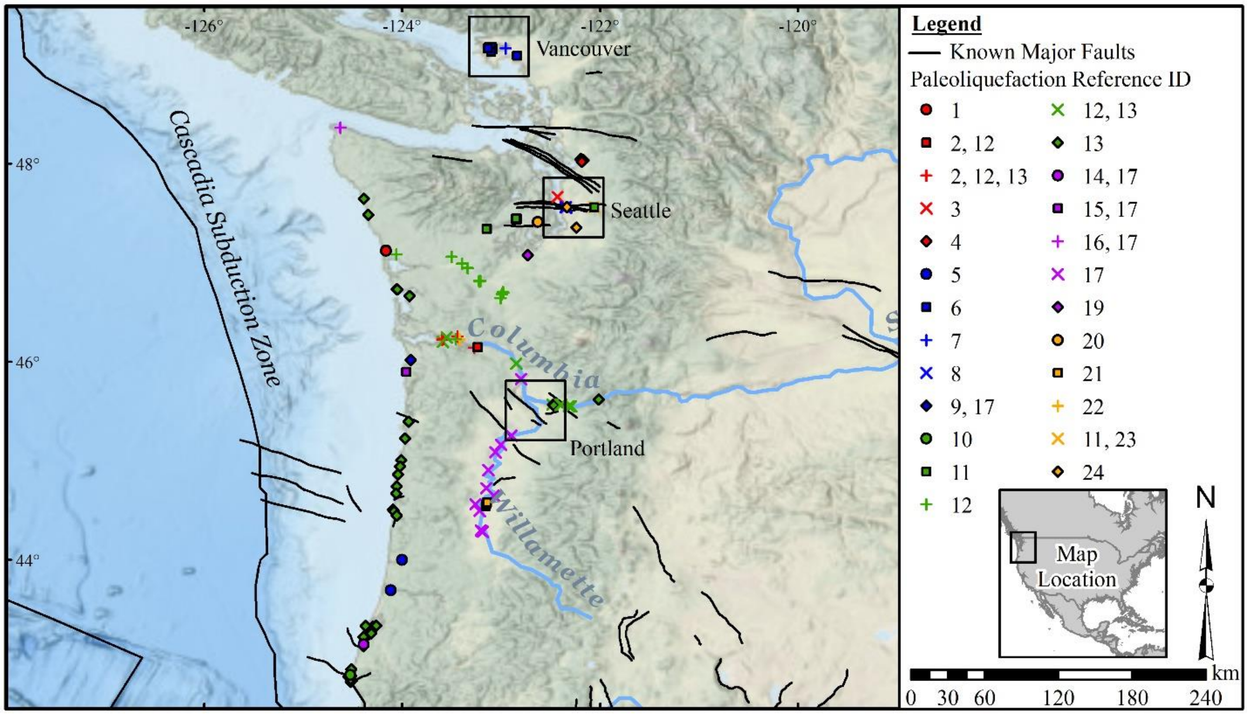


Figure 19. Compilation of paleoliquefaction data. Reference IDs are given in Appendix F.

3.4 Publication 7: Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington earthquake." *Earthquake Spectra*, In Press, 10.1177/87552930231174244, EERI.

- The ultimate goal of the PNW research is to predict liquefaction in M_w9 events using the best available models, and to constrain the ground motions experienced in 1700 CE via comparison to compiled paleoliquefaction evidence. For this purpose, nonlinear site-response models and CPT-based liquefaction triggering and manifestation models are used to study specific sites of interest. However, prior to using CPT-based models, these models were tested for use in the PNW through analysis of liquefaction case histories.
- In this study, 24 locations where the liquefaction response was well documented during the 2001 Nisqually, WA, earthquake, and where CPT testing has been performed, were studied. A total of 18 CPT-based models from the literature were tested and their predictions were compared to observations. This effort is also notable in that liquefaction case-histories from the PNW have not been published, even though widespread liquefaction has occurred in several modern events. Furthermore, these cases are among very few associated with subduction zone earthquakes, and potentially the only cases from an intraslab rupture. This compilation is also the second largest from any U.S. earthquake and the fourth largest from any earthquake outside of Canterbury, New Zealand. Only the 1989 Loma Prieta (U.S), 1999 Chi-Chi (Taiwan), and 2012 Emilia (Italy) earthquakes have produced more CPT case histories to date. In this regard, only ~275 CPT-based liquefaction case histories have been compiled from all earthquakes combined outside the Canterbury region of New Zealand. The compilation of 24 case histories thus represents a significant addition and can be used by the research community to train and test future liquefaction models.

- Specifically, a liquefaction case history consists of: (i) documentation of whether liquefaction manifestations were, or were not, observed; (ii) subsurface CPT testing at the site of observation; and (iii) measurement or estimation of the ground water depth and PGA experienced at the site.
- Case histories were compiled with emphasis on free-field level-ground sites. Nine of the twenty-four case history sites were classified as "positive," meaning that surficial manifestations of liquefaction were observed. The remaining fifteen sites were classified as "negative," meaning that no surficial evidence of liquefaction was observed. Sources of reconnaissance information included Bray et al. (2001), Seattle Office of Emergency Management (2019), Pacific Northwest Seismic Network (2001), Hausler and Koelling (2004), Walsh (2001), Geoengineers, Inc. (2003), and personal records and communication with individuals from the City of Seattle, regional utility companies, the University of Washington, and those who previously published reconnaissance reports in the public domain.
- Subsurface CPT tests and ground water table depth measurements were compiled from existing geotechnical tests available in the Jeschke et al. (2019) database, which contained more than 100,000 subsurface tests and observations from Washington State, including CPTs, standard penetration tests, borings, shear-wave velocity measurements, laboratory tests, and well logs. This required digitization of the 24 CPT profiles and thorough analysis of geotechnical reports.
- The methodology developed by Bradley and Hughes (2012) was used to compute conditional ground motions at case history sites (i.e., to predict ground motions, conditioned on nearby recordings). The Bradley and Hughes (2012) methodology computes a conditional probability distribution of PGA (i.e., conditional median PGA and standard deviation) by statistically coalescing ground motion records with empirical ground motion predictions, wherein the inter-

and intra-event uncertainties are used to develop the conditional distribution. In essence, when near a ground-motion recording station (where ground motions are known), the conditional median PGA approaches the measured PGA and the uncertainty approaches zero. However, as the distance from a ground-motion recording station increases, the uncertainty of the conditional median PGA increases, and its value becomes closer to that of the GMM.

- The Bradley and Hughes (2012) methodology was first applied in Canterbury, New Zealand. To apply the Bradley and Hughes (2012) methodology to the Nisqually earthquake, several modifications were required. These included: (i) ground motion recordings from throughout the PNW; (ii) the rupture's proposed seismic-source model; (iii) a subduction zone GMM that includes site response; and (iv) a spatial correlation model appropriate for subduction zones. For example, the new Next Generation Attenuation Subduction (NGA-Sub) GMM was implemented with its regional adjustment factors for the Cascadia region. The NGA-Sub GMM itself had several important inputs necessary for computations (e.g., V_{S30} which was obtained from the Washington DNR National Earthquake Hazards Reduction Program (NEHRP) seismic site class map).
- Application of the Bradley and Hughes (2012) methodology to the Nisqually earthquake is shown in Figures 20 and 21 for the Seattle area. Specifically, the conditional median and conditional uncertainty of PGA values are computed on a 0.1 km grid. Highlighted via markers in Figure 20 are the conditional median PGAs at liquefaction case-history sites and the observed PGAs at ground-motion recording stations. As expected, computed conditional median PGAs have similar values to those of the recording stations when near them, and differing values when moving further away. Figure 21 shows contours of conditional sigma values where values approach zero near stations and increase when moving away from stations.

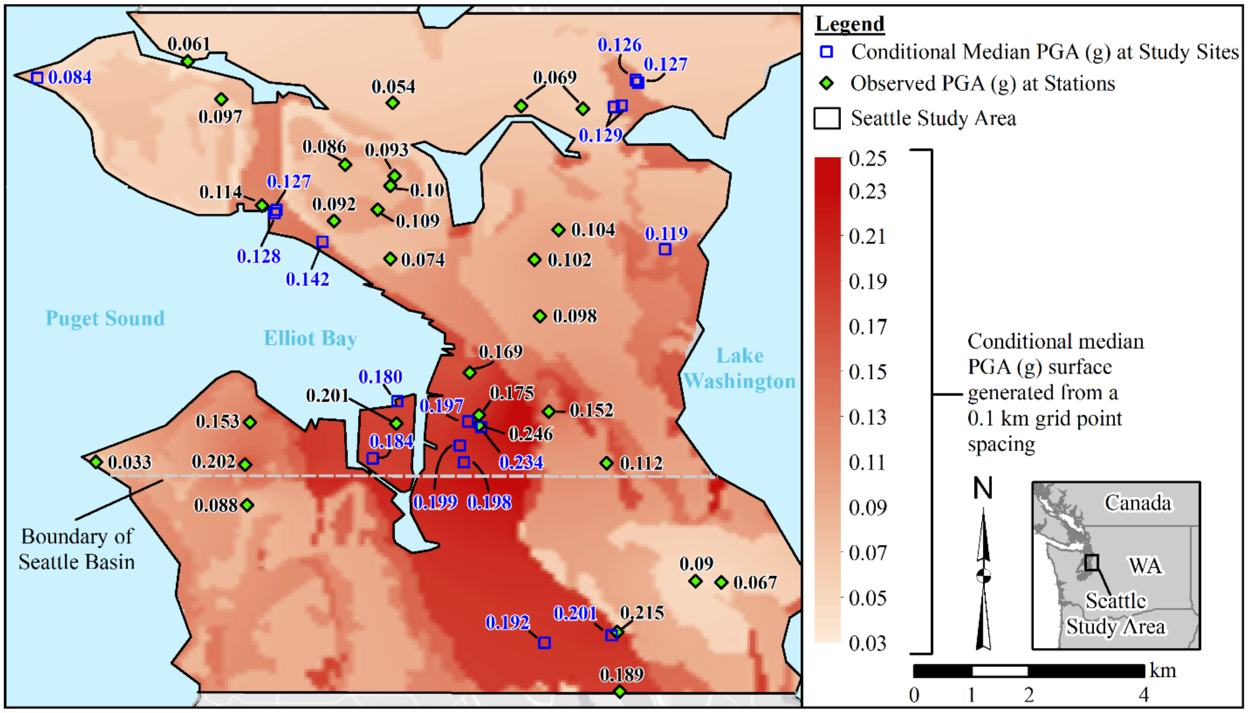


Figure 20. Conditional median PGA (g) surface overlain by conditional median PGA (g) at case history study sites and observed PGA (g) at ground motion recording stations in the Seattle study area. PGA = peak ground acceleration; km = kilometers.

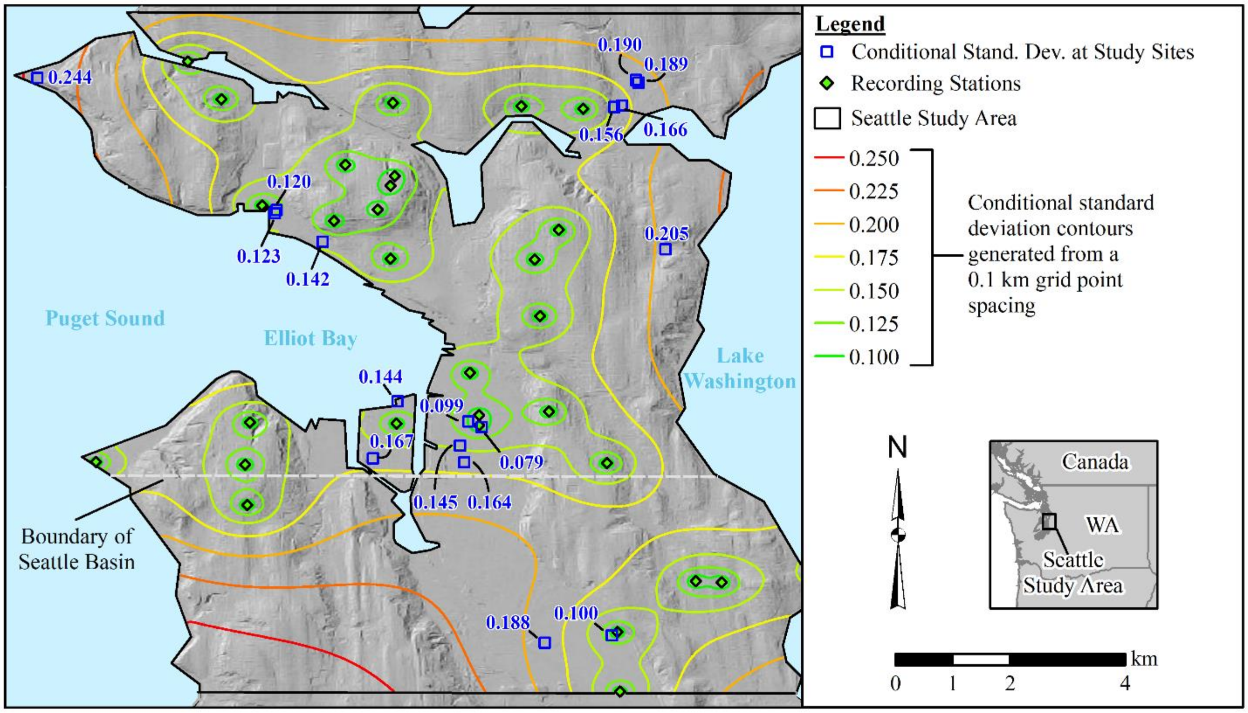


Figure 21. Contours of conditional lognormal standard deviation of PGA (g) and values at case history study sites in the Seattle study area. PGA = peak ground acceleration; km = kilometers.

- An ensemble of 18 CPT-based liquefaction models were used to predict liquefaction response at the 24 case history sites. Specifically, the liquefaction triggering models of Robertson and Wride (1998) [RW98], Architectural Institute of Japan (2001) [AIJ01], Moss et al. (2006) [Mea06], Idriss and Boulanger (2008) [IB08], Boulanger and Idriss (2016) [BI16], and Green et al. (2019) [Gea19] were each separately used in series with the liquefaction manifestation models of Iwasaki et al. (1978) [LPI], van Ballegooy et al. (2014) [LSN], and Maurer et al. (2015) [LPI_{ISH}]. Using global case-history data, Geyin and Maurer (2020a) conditioned fragility functions on each of these 18 models, thereby permitting probabilistic predictions of liquefaction surface manifestation to be made.

- Prior to using any liquefaction models, liquefaction susceptibility was inferred from the CPT soil behavior type index (I_c) (Robertson and Wride, 1998). However, because the relationship between I_c and susceptibility is uncertain, the I_c – susceptibility model of Maurer et al. (2019) was used to probabilistically predict susceptibility, as defined by Boulanger and Idriss (2006), from the measured I_c. Utilizing the conditional ground motion distributions, and considering the uncertainty in the liquefaction susceptibility threshold, the probability of observing liquefaction manifestation at the ground surface was computed at each of the study sites as:

$$P(Manifestation) = \int_{PGA} \int_{I_c} P(\text{Manifestation}|PGA, I_c)\, f(I_c) f(PGA) \cdot \mathrm{d}I_c \cdot \mathrm{d}PGA \quad (5)$$

  where $f(I_c)$ is the PDF of the I_c threshold for discriminating susceptibility (Maurer et al., 2019); $f(PGA)$ is the PDF of conditional PGA, and $P(\text{Manifestation}|PGA, I_c)$ is the probability of surficial manifestation, which is a conditioned on PGA and the threshold I_c for determining susceptibility, among numerous other inputs. In total, 18 different models for predicting the manifestation of liquefaction were implemented at each site.
- Averaging the liquefaction predictions across all 18 models for each case history, sites with observed manifestations have average probabilities of manifestation ranging from 0.12 to 0.79 whereas sites without observed manifestations have average probabilities ranging from 0.00 to 0.40. It may be observed that the predictions made by the 18 models often vary considerably for a given case history. The difference between the highest and lowest predicted probability, for example, exceeds 0.50 for 6 cases and exceeds 0.30 for 12 cases. These large differences indicate that the cases may be useful for resolving discrepancies amongst the models (i.e., that the soil is near the modeled threshold for triggering over several meters of the subsurface).

- To quantify model efficiency, receiver operating characteristic (ROC) analyses were performed and are summarized in terms of area under the ROC curve (AUC). The top-performing models include BI16-LPI (i.e., the BI16 triggering model used with the LPI manifestation model) and RW98-LPI, which each have an AUC of 0.926. For context, an AUC of 0.5 corresponds to random guessing (i.e., there is a 50% probability that sites with manifestations have higher model predictions than sites without manifestations). In the case of a perfectly efficient model, this probability is 100% (i.e., AUC is 1). Nearly as efficient as these models are BI16-LPIISH and Gea-LPI (AUC = 0.919). The least efficient model for this dataset is Mea06-LSN (AUC = 0.793), albeit its performance is still nearer to a perfect model than to random guessing. Notably, the computed AUC values have relatively large finite-sample uncertainties, given the limited number of compiled case histories.
- Accordingly, to evaluate whether the differences in AUC are statistically significant, p-values were computed per DeLong et al. (1988) to compare each model's AUC to all others. p-values are the probability that an observed difference in model performance is due to finite sampling (as opposed to real differences in performance). The most salient findings are: (i) no model is significantly better than all others; (ii) most of the measured differences in model performance are not statistically significant (i.e., of the 153 model comparisons, just 13 are significant) and the overall conclusion is that little can be confidently said about model superiority; and (iii) the models with significantly different performance are largely models based on LSN, which when compared to the best-performing models (e.g., BI16-LPI), are significantly less efficient.
- Finally, the optimum operating point (OOP), or computed probability of manifestation that most optimally separates cases with and without observed manifestations, was extracted from the ROC analyses. An OOP of 0.19 was computed for the probability of manifestation

averaged across all 18 models, suggesting that liquefaction manifestations may have occurred more readily than expected by the Geyin and Maurer (2020) fragility functions. That is, the liquefaction hazard may have been underpredicted based on calibrations of the 18 models against global case histories. However, like AUC values the finite-sample uncertainty of OOPs should be considered, given the limited number of case histories. To account for this uncertainty, bootstrap sampling of the 24 cases was used to compute a 95% confidence interval for the OOP averaged across all models. This interval ranged from 0.12 to 0.56, and as such, it cannot be confidently concluded that the OOP is less than 0.5 (i.e., that the Nisqually case histories deviate from median global behavior). Thus, while liquefaction manifestations may have occurred more readily than expected for the Nisqually dataset, additional case histories would be needed to confirm this, or to draw any definitive conclusion about performance and bias.

- This study compiled 24 liquefaction case histories, which is the 2nd largest compilation from a U.S. earthquake to date and are among the very few in literature associated with subduction zone earthquake. The Nisqually case histories provide a useful addition for model training and testing but must be combined with existing case histories (e.g., Brandenberg et al., 2020; Montalva et al., 2021; Geyin et al., 2021) and others not yet collected.
- The 18 different liquefaction prediction models were tested on the 24 Nisqually case histories. These prediction models revealed differences in model efficiency, albeit these differences were rarely statistically significant when considering finite-sample uncertainty. This conclusion alludes to the general inappropriateness of championing models based on individual earthquakes or otherwise small datasets, as is frequently done in the literature without statistical backing.

- To disseminate the case history data for model training and testing, the 24 case histories and their metadata were made publicly available on DesignSafe (see Publication 8).
- Because no CPT-based liquefaction model was shown to be superior, all 18 models are used in the subsequent analyses to predict the liquefaction from M_w9 CSZ ruptures at 400 sites where detailed site-specific data was compiled (see Publication 9).

3.5 Publication 8: Rasanen, R.A., Geyin, M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, DOI: 10.17603/ds2-nsf8-7944.

- This online dataset includes the CPT data for all 24 liquefaction case histories and a summary of the case-history metadata developed in Publication 7. These products include:
    - A comprehensive summary of the case-history metadata (Summary Table S1.xlsx). This comprehensive summary includes: the CPT depth; CPT cone tip resistance; CPT sleeve friction; CPT excess pore pressure (if recorded); the site name; latitude; longitude; citations for the geotechnical testing and post-earthquake reconnaissance; the CPT test date; manifestation classification; the ground water table depth; the CPT pre-drill depth; GMM inputs for each case-history site (e.g., V_{S30}, $Z_{1.0}$, $Z_{2.5}$, basin flags); conditional median PGA and standard deviation, and notes and commentary specific to each case history.
    - The CPT data for each case history is provided in 24 separate files (.xlsx format), where the most pertinent metadata specific to each case is again given (e.g., ground water table and pre-drill depths, median and sigma of conditional PGA, etc.).

3.6 PUBLICATION 9: RASANEN, R.A. ET AL. (2023). "IMPLICATIONS OF PHYSICS-BASED M9 GROUND MOTIONS ON LIQUEFACTION-INDUCED DAMAGE IN THE CASCADIA SUBDUCTION ZONE: LOOKING FORWARDS AND BACKWARDS." *IN PREPARATION.*

- Building on prior publications, Publication 9: (i) uses the tested geotechnical liquefaction models to predict the range of expected responses for 30 different physics-based ground motion simulations of a M_w9 CSZ earthquake (Frankel et al., 2018; Wirth et al., 2018); and (ii) compares these expected responses to current CSZ paleoliquefaction evidence.
- Sites to be studied include those of 400 CPTs in the PNW, several of which are located near well-studied locations where the presence or absence of 1700 CE liquefaction has been documented. To date, no regional scale prediction of liquefaction occurrence and consequence in an M_w9 CSZ event has been made using subsurface test data and the most advanced models for predicting ground motions and liquefaction response.
- Of the 400 study sites, 393 subsurface CPT tests and ground water table depth measurements were compiled from existing geotechnical tests collectively available from the Jeschke et al. (2019) database and numerous geotechnical consulting reports. The CPT profiles from these sources required digitization and careful analysis of the respective geotechnical reports. Data for 7 additional CPTs was collected via the USGS CPT truck at well-studied 1700 CE sites, which required substantial permitting and communication with various government and private entities.
- Given the premise of a full-margin M_w9 CSZ earthquake, physics-based seismograms (i.e., ground-motion time histories) were recently developed at the USGS (Frankel et al. 2018; Wirth et al. 2018) to predict motions across the PNW. The suite of simulations includes 30 different

realizations that reflect uncertainties in key modeling parameters, such as the down-dip limit of fault rupture, the slip distribution and location of asperities, the rupture direction, and the hypocenter location. Since the last CSZ interface earthquake occurred in 1700 CE, the actual motions are unknown, and thus, are quite uncertain. For example, Figures 22a and 22b show the geometric mean PGAs for simulations csz-005 and csz-025, respectively, at the 400 CPT locations (discussed subsequently). Figure 22a shows that simulation csz-005 predicts weaker PGAs inland in the north (e.g., in Seattle, WA and Vancouver, BC) and stronger PGAs inland in the south (e.g., Portland). Meanwhile, a different rupture scenario is shown in Figure 22b for simulation csz-025, which predicts larger PGAs in Seattle, WA, which are up to 260% greater than predicted by csz-005 at the same location. Both simulations predict strong shaking on the coast, but these predictions often differ greatly in the far North and far South. In fact, the differences in PGA in the near field are often larger than in the far field. For example, csz-005 at Cape Blanco predicts 2.15 g while csz-025 predicts 0.23 g.

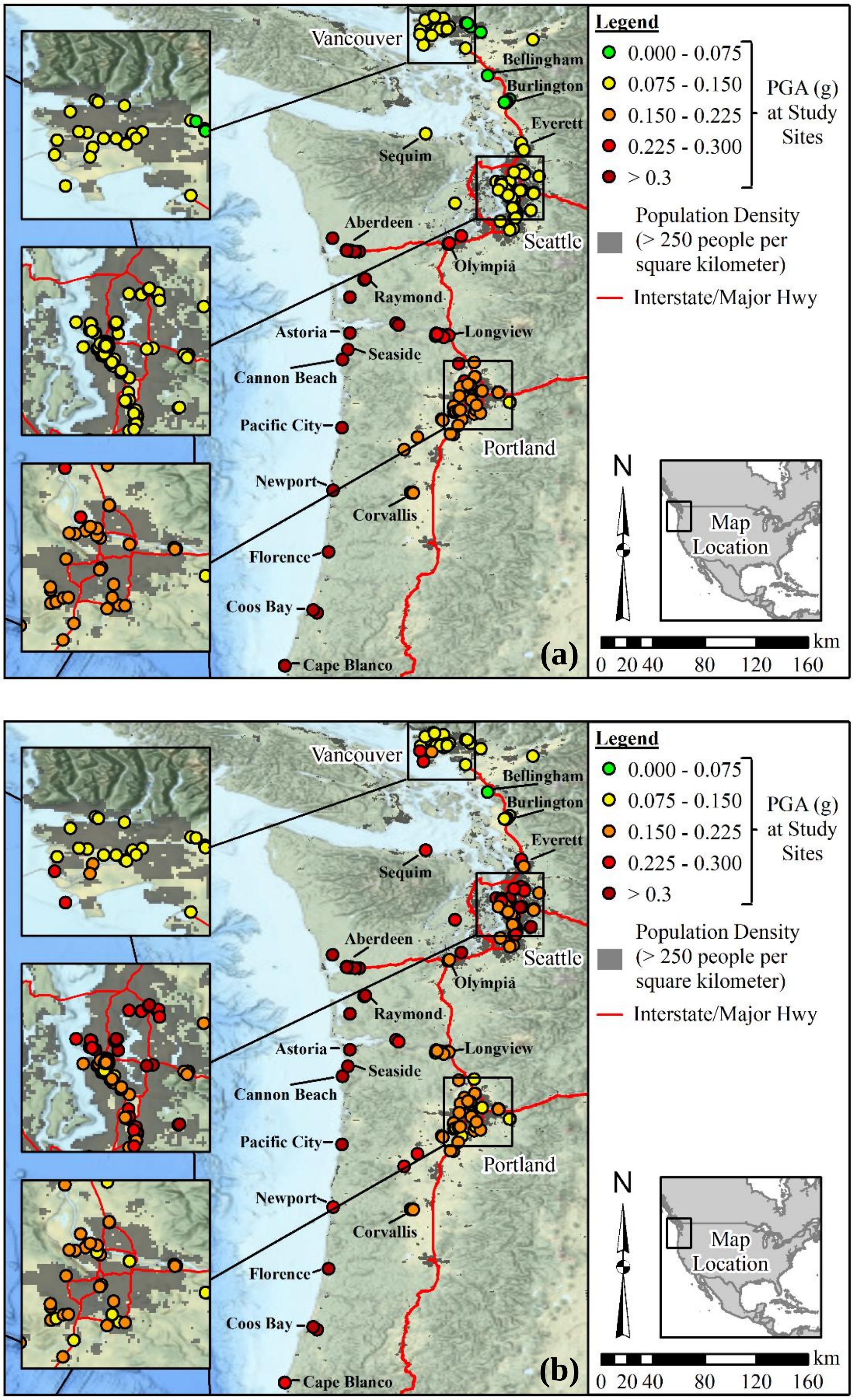


Figure 22. PGAs at 400 CPT study sites for ground-motion simulation: (a) csz-005 and (b) csz-025. PGA = peak ground acceleration; csz = Cascadia Subduction Zone.

- Importantly, the PGAs shown in Figures 22a and 22b assume a surficial V_s of 600 m/s and have not been adjusted for local site conditions that could significantly alter the motions. That is, the predictions are for rock conditions and do not consider the potential for local soil conditions to alter the amplitude, duration, and frequency content of incoming motions. Therefore, to predict liquefaction, the CSZ ground-motion simulations need to be modified for surficial conditions through site response analysis. In Publication 5, site response was handled using the Marafi et al. (2021) soil velocity model and a simple 1D equivalent-linear site response analysis was performed. However, for $M_w 9$ motions, which can cause large shear strains in liquefiable soils (i.e., > 0.4%), a site response model that more completely captures the true nonlinear behavior of soil should be used to obtain more accurate ground motions at study sites. Failure to do so (i.e., by relying on equivalent-linear analyses) can potentially result in misprediction of high-frequency ground motions (e.g., PGA) at sites of strong shaking. In total, 30 CSZ simulations were propagated through the soil at 400 CPT study sites.
- However, to run nonlinear site response analysis a profile of shear wave velocity (V_s) is needed. For 394 CPTs, the V_s profiles were generated by fitting and optimizing a function of the form of the Wirth et al. (2021) Soil Velocity Model (SVM) on estimates of V_s versus depth. These estimates included: (i) V_s values averaged from 6 CPT-V_s models; and (ii) V_s values predicted by the Wirth et al. (2021) Pacific Northwest SVM. The CPT-V_s model directly computes V_s from CPT cone tip resistance, among other variables. A downside to the CPT-V_s model is CPTs oftentimes do not extend to a deep enough depth to reach a V_s = 600 m/s. Thus, the region-specific Wirth et al. (2021) SVM, which only requires V_{s30} as an input, was used to generate V_s profiles to 600 m/s. Therefore, to obtain a reasonable V_s profile down to 600 m/s, the two estimation methods were meshed and refit with the Wirth et al. (2021) functional form.

For the remaining 6 CPTs, which are located near 1700 CE paleoliquefaction sites, active and passive multi-channel analysis of surface waves (MASW) testing was performed to produce a site-specific V_s profile down to a depth of V_s = 600 m/s.

- Using the 400 V_s profiles, nonlinear time-domain site response analyses were performed at each study site using Seismosoil (Asimaki and Shi 2017), which employs the finite-difference method. This approach includes: (i) stress-strain and damping behaviors described by the hybrid hyperbolic model (Shi and Asimaki 2017), which has shown improved performance over the popular modified hyperbolic (MKZ) model; and (ii) hysteresis behavior governed by the model of Li and Assimaki (2010), rather than by Masing rules. The resulting time histories at the ground surface account for the local soil conditions, which are different for each site, instead of assuming each CPT was performed on soft rock.
- To illustrate the differences between nonlinear and equivalent linear site response analysis, Seismosoil was also used to generate time histories via equivalent linear site response analysis with frequency dependent modulus and damping. Equivalent linear assumes constant soil modulus and damping in each layer throughout the duration of shaking, whereas nonlinear site response analysis does not, which leads to the differences between the two. In total, 24,000 time histories were adjusted for soil conditions for both equivalent linear and nonlinear site response analysis. The geometric mean horizontal PGA was then computed from the horizontal components at each CPT-simulation combination giving 12,000 predictions of PGA.
- To illustrate the differences between nonlinear and equivalent linear site response analysis for each CPT-simulation combination, their geometric mean horizontal PGAs are compared in Figure 23. The two methods generally agree at peak shear strains less than 0.4%, but in general as the peak shear strain increases, typically when PGAs are large, equivalent linear tends to

predict higher PGAs. In addition, there is more uncertainty in both the equivalent linear and non-linear results as the peak shear strain increases. This discrepancy becomes especially apparent when peak shear strains are greater than 1%. Therefore, using equivalent linear site response analysis for sites along the coast where shaking is strong (and thus where higher peak shear strains are more likely to occur) could lead to erroneous predictions. However, in inland locations such as Seattle, Portland, and Vancouver, use of equivalent linear analysis may be adequate for determining site adjusted PGA if the peak shear strain is less than 0.4%.

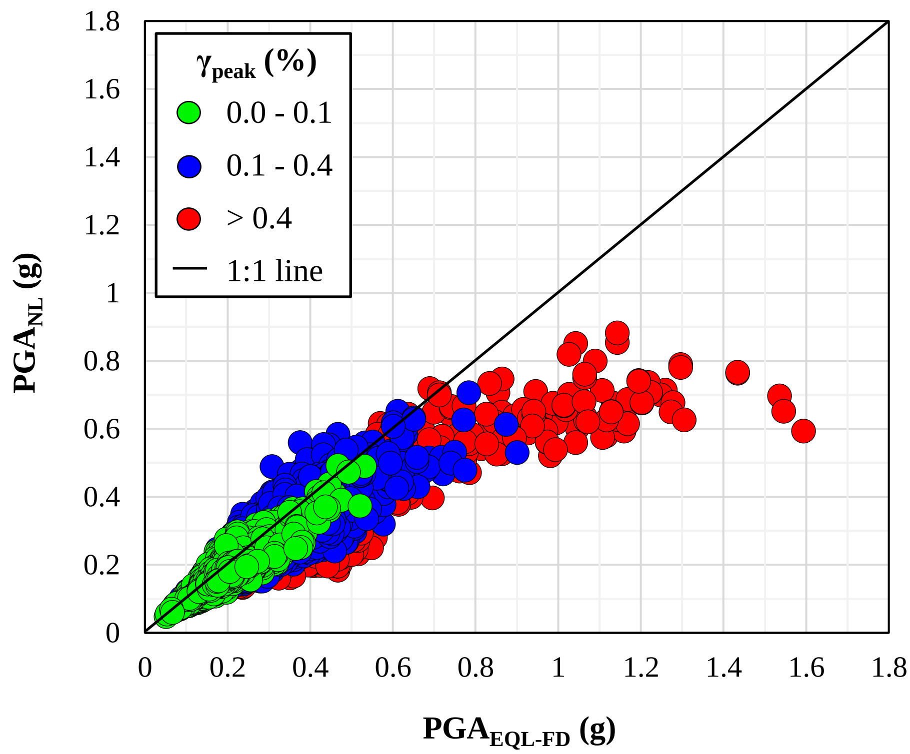


Figure 23. Comparison of Peak Ground Acceleration (PGA) predicted by non-linear (NL) and equivalent linear (EQL-FD) site response analyses.

- To predict the probability of liquefaction manifesting at the ground surface, 18 liquefaction models were used and equally weighted. No one model was weighted more based on the lack of model superiority in the PNW, as discussed in Publication 7.

- In total, liquefaction predictions were made at 400 CPT locations for 30 different simulations, resulting in 12,000 predictions of the probability of ground failure (i.e., liquefaction manifesting at the ground surface). Figure 24 shows the median probability of ground failure at the 400 CPTs, as predicted from the 30 simulations. This median is above 50% at many locations in Seattle, Portland, and Vancouver, amongst many other towns and cities. To further illustrate the expected liquefaction response, Figures 25a and 25b show the 2.5th percentile and median probability of ground failure in Seattle, WA. Even for the weakest shaking, Figure 25a shows high probabilities of ground failure in areas of Seattle (e.g., the Greater Duwamish, Interbay, and University districts). These tend to be locations where artificial fill was placed in the past. This should not be interpreted to mean that only locations of fill will liquefy, as liquefaction has been locally documented in alluvium and lahar deposits (e.g., Davis 2018). The 97.5th percentile predictions in Seattle, as well as various other predictions of liquefaction response, are given in the complete paper. These analyses show that similar to the differences in PGA predicted across the 30 simulations, the differences in liquefaction response also often vary considerably. Therefore, it would be useful to determine which, if any, simulation or set of simulations best represent the ground motion intensities experienced in the 1700 CE rupture.

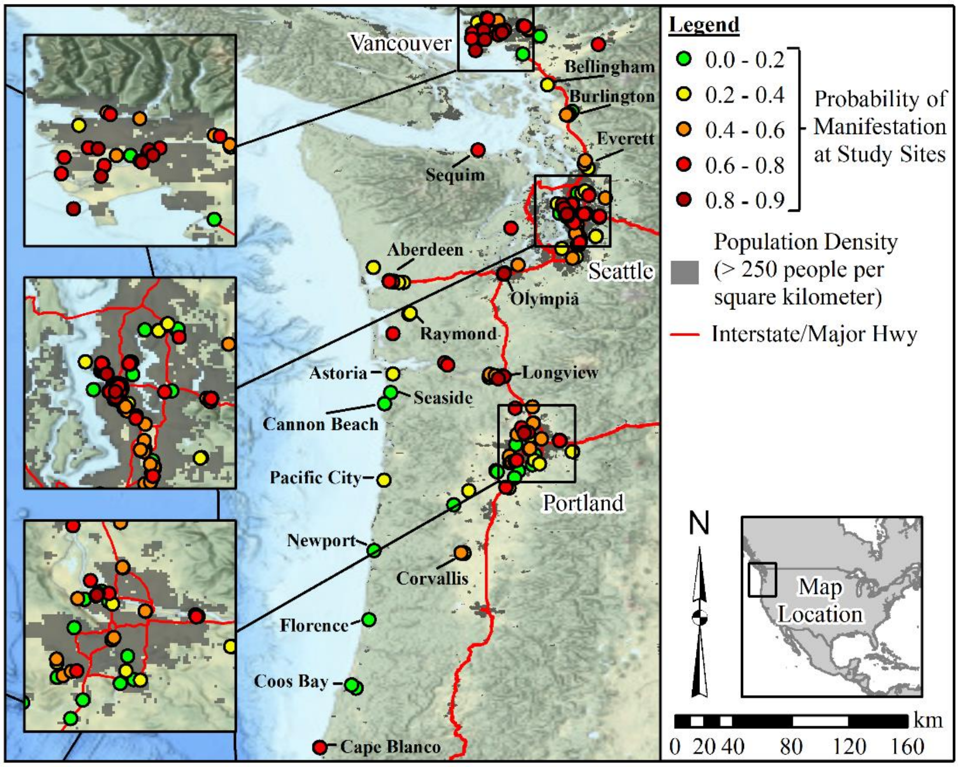


Figure 24. Median probability of liquefaction surface manifestation at 400 CPT sites considering 30 ground motion simulations.

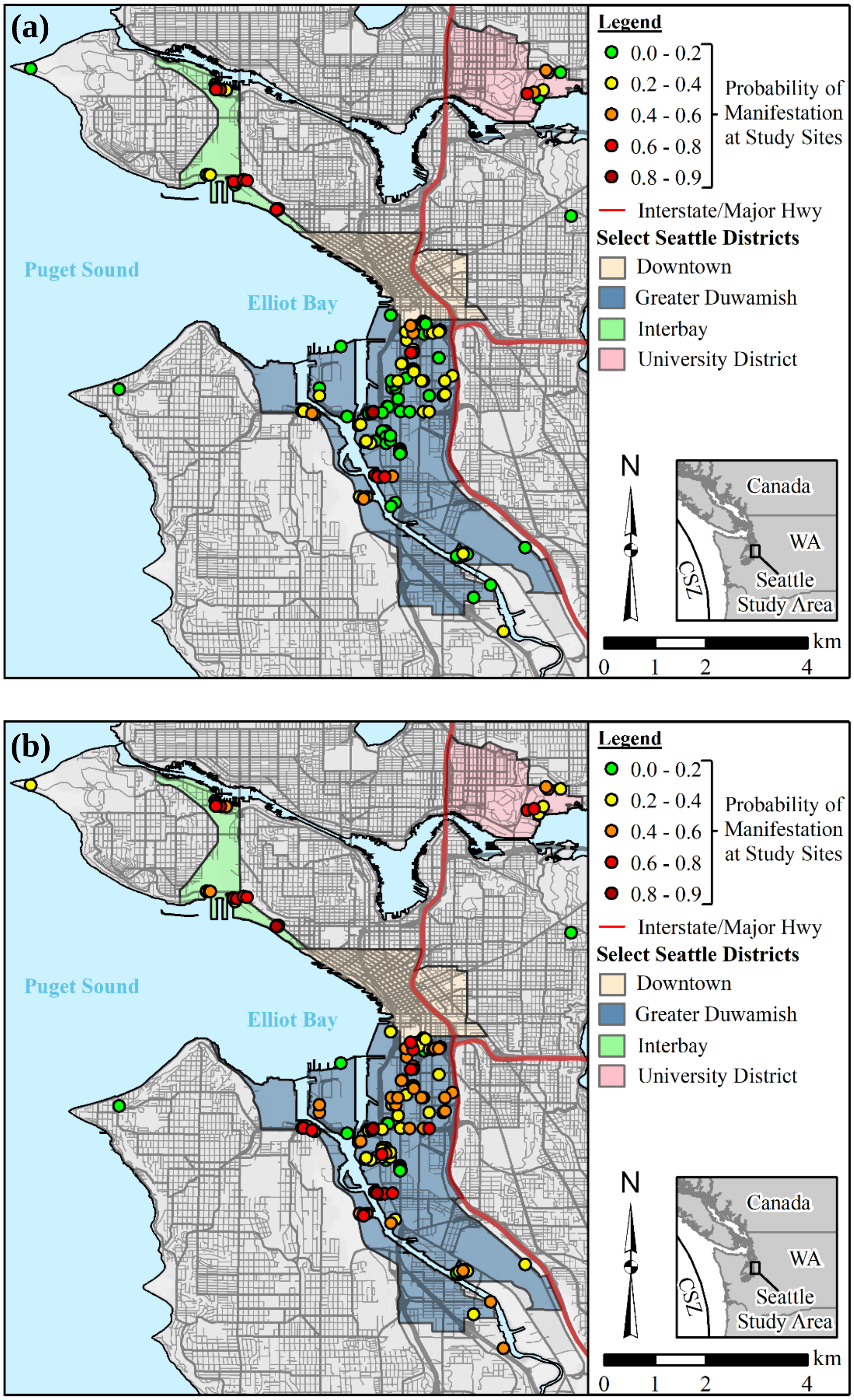


Figure 25. Probability of liquefaction surface manifestation at the (a) 2.5th and (b) 50th percentiles at study sites in Seattle, WA, considering 30 ground motion simulations.

- To explore which simulations are more and less likely to produce the 1700 CE paleoliquefaction evidence, techniques introduced in Chapter 2 were applied to sites where 1700 liquefaction has and has not been observed in rigorous field studies. Specifically, the likelihood that a certain simulation produced the observed 1700 paleoliquefaction evidence is computed.
- In Publication 5, all PNW paleoliquefaction evidence discovered to date was compiled. This included several well-studied 1700 CE paleoliquefaction sites, including: Price and Hunting Islands along the Columbia River, WA (Atwater 1994); Cape Blanco State Park on the coast of Oregon (Kelsey et al. 2002); Kellogg Island in Seattle, WA (Davis 2018); the Snohomish River Delta near Everett, WA (Bourgeois and Johnson 2001); Lynch Cove near Belfair, WA (Martin and Bourgeois 2012); Annacis Island in Vancouver, BC (Clague et al. 1997); and the Sandy River, east of Portland, OR (Peterson and Madin 1997), amongst several others. To account for the uncertainties of whether positive sites (i.e., liquefaction observed) and negative sites (i.e., no liquefaction observed) are truly positive or negative, classification uncertainties were assigned to each site and incorporated into the likelihood analysis. These uncertainties were assigned considering the rigor of prior field studies, the age constraints obtained by these studies, and conversations with the original authors. Nonetheless, these classification uncertainties are ultimately subjective. Therefore, several different analyses were performed to explore this uncertainty, as will be further discussed.
- To account for presence of artificial fill at several study sites (i.e., to correct for its absence in 1700), the in-situ CPT profile and ground water table at the time of testing were used to stress-normalize CPT measurements, whereas the trimmed profile with typically ~ 1-2 m of fill removed was used to infer saturation and compute the cyclic stress ratio at the time of shaking.

In this regard, the 1700 ground water table was assumed to be at the same elevation as present day (i.e., it was ~ 1-2 m closer to the ground surface in 1700). While this introduces uncertainty, the results are ultimately insensitive to this assumption and would only change if the ground water table was considerably deeper in 1700 CE than at present, which is highly unlikely.

- In the first analysis of 1700 evidence, classification uncertainties were not included. That is, field observations (i.e., liquefaction did or did not manifest in 1700) were considered certain. This analysis showed csz-005 to be the most likely simulation to have produced the 1700 paleoliquefaction evidence. Simulation csz-005, shown previously in Figure 22a, predicts weaker shaking inland in the north (e.g., in Seattle, WA and Vancouver, BC) and stronger shaking inland in the south (e.g., Portland, OR). It is important to note that Figure 22a shows csz-005 prior to site-response analysis (i.e., all sites are considered to be on soft rock). The absence of site effects allows for a clearer picture of the regional prediction trends.
- In a second analysis, classification uncertainties were assigned values at each site using the holistic approach previously described. This analysis also indicated that csz-005 is the most likely simulation. However, because the classification uncertainties are ultimately subjective, several additional possibilities were considered. These included: a variety of different plausible classification uncertainties; raising and lowering the ground water table depth at each site by two meters; removing and adding certain study sites from the analyses; and foregoing trimming of CPT profiles where artificial fill was placed following the 1700 earthquake. Ultimately, the most likely simulation was always csz-005, whereas the least likely simulation was almost always csz-025. Simulation csz-005 produces weaker shaking on average at the study sites relative to the other simulations.

- While simulation csz-005 fits the paleoliquefaction observations the best among the 30 simulations, it still fits the observations poorly. That is, relatively high probabilities of liquefaction manifestation are computed at sites where none was observed from the 1700 event. A simulation that would match the paleoliquefaction evidence better would be one that does not predict liquefaction in the Puget Sound and northward. Such a simulation might be similar to csz-005, but the shaking in Washington would have to be less.
- Importantly, these findings are based on 1700 CE paleoliquefaction evidence, and thus are for the 1700 CE earthquake. However, the CSZ has ruptured many times in the past and ruptures prior to 1700 may have had different rupture properties and therefore different ground motions and different liquefaction fields. To elucidate the shaking in these pre-1700 ruptures, paleoliquefaction evidence tied to each specific rupture would need to be analyzed. In this regard, it is crucial for researchers to conduct ongoing field explorations to document and date new evidence for all CSZ ruptures, as well as other earthquake sources that have previously caused significant ruptures, like the Seattle fault (Atwater 1999).
- There are several possible explanations for the discrepancy between predictions and observations of liquefaction at regional scale. These include: (i) there is an abundance of 1700 paleoliquefaction evidence that has not been discovered (e.g., in the Puget Lowland); (ii) component models used to predict site response and liquefaction are erroneous; (iii) the simulations overpredict PGAs in inland hubs such as Seattle, WA; and/or (iv) the 1700 CE rupture was not a full-margin rupture. The second is relatively unlikely given that the adopted techniques are widely used and well validated. It is true, however, that liquefaction models have been trained on relatively few case histories from earthquakes of M_w8 and above. These models have not been shown to be less accurate or more uncertain in such events, but the

possibility persists. The first is plausible but also perplexing, given that there are many locations in the Puget Lowland and Fraser River delta where paleoliquefaction has been discovered from earthquakes that are younger and/or older than 1700 CE. This indicates the strength of ground shaking in the Puget Lowland and Fraser River delta may have been lower than the simulations predict, given the lack of paleoliquefaction evidence despite relatively rigorous efforts to find it. In this regard, the fourth possibility could explain the lack of paleoliquefaction evidence in the Puget Lowland and Fraser River delta, given that smaller M_w ruptures have shorter duration of shaking and would likely produce lesser shaking intensities in certain regions of the CSZ. Furthermore, recent studies (Melgar, 2021) have shown the 1700 CSZ rupture need not be a full-margin rupture to produce the other types of paleoseismic evidence (e.g., subsidence and tsunami) that have been tied to the 1700 rupture.

- This study led to: (i) the first regional scale CSZ M_w9 liquefaction assessment using geotechnical data; and (ii) constraint of the intensity of ground shaking from the 1700 CE earthquake, to the degree permitted by current paleoliquefaction evidence. It was shown that even for the simulations that predict the weakest PGAs, the probability of ground failure is high in inland hubs such as Seattle, WA, Portland, OR, and Vancouver, BC. It was also shown that simulation csz-005, which predicts weaker shaking inland in the north (e.g., in Seattle, WA and Vancouver, BC) and stronger shaking inland in the south (e.g., Portland), was the more likely simulation, albeit this simulation is also relatively unlikely given the high probabilities of liquefaction manifestation computed at sites where none was observed. Collectively, these analyses indicate that 1700 CE ground motions may have been less intense in inland hubs such as Seattle, WA, Portland, OR, and Vancouver, BC than prevailing forecasts indicate. Ultimately, these conclusions will either be confirmed or revised as additional 1700 paleoliquefaction evidence is discovered.

Chapter 4. CONCLUSIONS

In regions of infrequent moderate-to-large earthquakes, historic earthquake catalogs are often insufficient to provide inputs to seismic-hazard analyzes (i.e., fault locations and magnitude-frequency relations) or to inform ground-motion predictions for certain seismic sources. In these regions, analysis of relic coseismic evidence is commonly used to infer information about the seismic hazard. However, while paleoseismic studies have been performed widely, all existing analysis techniques require *a prior* assumptions about the causative earthquake's location (i.e., rupture magnitude and ground motions can otherwise not be estimated). Accordingly, *Chapter* 2 of this dissertation presented a flexible inversion methodology that probabilistically constrains seismic source parameters from coseismic evidence. More specifically, this methodology results in: (i) a geospatial likelihood surface that constrains the rupture location; and (ii) a probability distribution of the rupture magnitude. To evaluate the methodology and glean lessons for forward use, it was tested on modern earthquakes with known seismic source models. The inversion methodology was first applied to liquefaction induced by three earthquakes in Canterbury, New Zealand, utilizing the extensive CPT-based liquefaction dataset resulting from these events. This was followed by further testing on 5 global earthquakes with fewer liquefaction case histories. Next, the methodology was tested on landslides from four earthquakes in California. These simulated analyses illustrate the proposed methodology's provocative potential, while also highlighting caveats and lessons for forward use. Most Importantly, the methodology remains conceptually valid for all earthquakes in all locations, so long as ground motions and coseismic response are predictable. Therefore, the proposed methodology has the potential to provide new insights in enigmatic seismic zones worldwide.

Moving beyond development and testing, *Chapter 3* summarizes the compilation and analysis of field data in the Charleston Seismic Zone and Cascadia Subduction Zone where seismic hazards are enigmatic. This included applying the likelihood methodology to coseismic evidence left behind by the 1886 CE Charleston, SC and 1700 CE Cascadia Subduction Zone earthquakes to constrain seismic parameters from these events. The most salient findings from these studies were:

- Application of the inversion methodology to the 1886 Charleston, SC earthquake finds the location and orientation of the 1886 fault rupture could not be confidently constrained by the available felt-intensity and liquefaction evidence. Furthermore, the analyses do not show strong support for the hypothesized Woodstock Fault, which has commonly been considered the source of the 1886 earthquake. However, this result may change if additional evidence were available for analysis, or if the component models used for analysis were less uncertain. Collectively, the analysis of the 1886 Charleston, SC earthquake largely support the M_w distribution adopted by the National Seismic Hazard Model Project, which ranges from M_w6.7-M_w7.5 with a median of M_w7.1. While analyses indicated that M_w < 6.7 ruptures have potential to produce the observed liquefaction response (particularly when the uncertainty of selecting an aging-correction model is considered), the felt-intensity evidence suggests a near-zero likelihood of such magnitudes. Conversely, both types of evidence suggest that M_w > 7.5 events have potential to produce the field evidence. While this conclusion hinges on which component models are adopted to predict soil-aging effects and felt-intensity, we fail to find conclusive evidence for outright rejecting some models in favor of others. In the absence of such evidence, M_w > 7.5 ruptures would merit consideration.

- All publicly available paleoliquefaction data in the U.S. Pacific Northwest and southeastern British Columbia were compiled and paleoliquefaction forecasted using a geospatial liquefaction model. The results of this study included:
    - Compilation of 185 paleoliquefaction study sites documented from 24 publications.
    - The first ever suite of site-adjusted M_w9 Cascadia Subduction Zone ground-motion predictions and forecasted maps of liquefaction impacts for the Cascadia Region using a geospatial liquefaction model. These forecasted maps include PGA, PGV, the probability of liquefaction, and liquefaction spatial extent.
    - Collectively, these maps can be used to guide field explorations by engineering geologists for new evidence, considering both: (i) where evidence is likely to be found; and (ii) whether evidence is likely to provide meaningful constraint of Cascadia Subduction Zone ground motion models. Of additional utility, this process resulted in the first ever suite of physics-based M_w9 Cascadia Subduction Zone ground-motion predictions that account for the response of near-surface sediment. In making these predictions, equivalent linear site-response analyses were used to modify the Frankel et al. (2018) and Wirth et al. (2018) simulations for local, near surface conditions. Together, the maps of expected shaking intensity and liquefaction may be useful in regional loss estimation, scenario planning, and science communication.
- Geotechnical liquefaction prediction models were tested in the U.S. Pacific Northwest on 24 CPT-based liquefaction case histories compiled from the 2001, Nisqually, Washington, earthquake. This study led to:
    - The 2nd largest compilation of CPT-based liquefaction case histories from a U.S. earthquake to date and are among the very few in literature associated with a subduction

zone earthquake. The Nisqually case histories provide a useful addition for model training and testing but must be combined with existing case histories and others not yet collected.

  - Testing 18 different geotechnical liquefaction prediction models on 24 Nisqually case histories. These prediction models revealed differences in model efficiency, albeit these differences were rarely statistically significant when considering finite-sample uncertainty. This conclusion alludes to the general inappropriateness of championing models based on individual earthquakes or otherwise small datasets, as is frequently done in the literature without statistical backing.
  - The dissemination of 24 CPT-based liquefaction case histories and their metadata to allow others to train and test new and existing liquefaction models.

- Finally, with the 18 liquefaction models tested, and the Pacific Northwest paleoliquefaction data compiled, the following analyses were performed: (i) the first regional scale Cascadia Subduction Zone M_w9 liquefaction assessment using geotechnical data; and (ii) constraint of the intensity of ground shaking from the 1700 CE earthquake, to the degree permitted by current paleoliquefaction evidence. The most salient findings from this study were:
  - Even for the simulations that predict the weakest peak ground accelerations, the probability of ground failure is high at some locations of inland hubs such as Seattle, WA, Portland, OR, and Vancouver, BC.
  - It was shown that simulation csz-005, which predicts weaker shaking inland in the north (e.g., in Seattle, WA and Vancouver, BC) and stronger shaking inland in the south (e.g., Portland), was the most likely simulation, albeit this simulation is also relatively unlikely given the high probabilities of liquefaction manifestation computed at sites

where none was observed. In that regard, all simulations predict far more liquefaction manifestation than has been observed. That is, paleoliquefaction evidence has not been constrained to the 1700 CE event in the Puget Lowland and Fraser River delta, although all simulations predict liquefaction should be present.

- Collectively, these analyses indicate the most plausible explanation is either:
  - There is an abundance of 1700 CE paleoliquefaction evidence that has not yet been discovered (e.g., in the Puget Lowland and Fraser River delta); or
  - The 1700 CE peak ground accelerations were less intense at inland hubs such as Seattle, WA, Portland, OR, and Vancouver, BC, than prevailing forecasts indicate, possibly suggesting that the 1700 CE earthquake was not a full-margin rupture.
- Ultimately, these conclusions will either be confirmed or revised as additional 1700 CE paleoliquefaction evidence is discovered.

BIBLIOGRAPHY

Architectural Institute Japan (2001). *Recommendations for design of building foundations*, 486 p.

Atwater, B.F. (compiler) (1994). Geology of Holocene liquefaction features along the lower Columbia River at Marsh, Brush, Price, Hunting, and Wallace Islands, Oregon and Washington. *U.S. Geol. Surv. Open-File Rept.* 94-209, 30 pp.

Atwater, B.F., (1999), Radiocarbon dating of a Seattle earthquake to A.D. 900–930, *Seismological Research Letters*, v. 70, p. 232.

Amick, D., G. Maurath, and R. Gelinas (1990). Characteristics of seismically induced liquefaction sites and features located in the vicinity of the 1886 Charleston, South Carolina earthquake, *Seismological Research Letters,* 61, no. 2, 117-211.

Archuleta RJ (1984) A Faulting Model for the 1979 Imperial Valley Earthquake. *Journal of Geophysical Research* 89(B6): 4559-4585.

Asimaki, D. and Shi, S. (2017). "SeismoSoil User Manual, v1.3" c 2014–2017, GeoQuake Research Group, California Institute of Technology.

Atkinson G.M., and S.I. Kaka (2007). Relationships between felt intensity and instrumental ground motion in the central United States and California, *Bull. Seismol. Soc. Am*. **97**, no. 2, 497-510, doi: 10.1785/0120060154

Atwater, B.F., Musumi-Rokkaku, S., Satake, K., Tsuji, Y., Ueda, K., and Yamaguchi, D.K. (2015), The orphan tsunami of 1700—Japanese clues to a parent earthquake in North America, 2nd ed.: Seattle, University of Washington Press, U.S. Geological Survey Professional Paper 1707, 135 p., doi: https://doi.org/10.3133/pp1707

Atwater, B.F., Stuiver, M., and Yamaguchi, D.K. (1991). "Radiocarbon test of earthquake magnitude at the Cascadia subduction zone." *Nature* 353: 156-158.

Bakun W.H., and M.G. Hopper (2004). Magnitudes and locations of the 1811-1812 New Madrid, Missouri, and the 1886 Charleston, South Carolina, earthquakes, *Bulletin of the Seismological Society of America*, **94**, no. 1, 64-75, doi: 10.1785/0120020122.

Bakun W.H., A.C. Johnston, and M.G. Hopper (2002). Modified Mercalli Intensities (MMI) for large earthquakes near New Madrid, Missouri, in 1811-1812 and near Charleston, South Carolina, in 1886, *U.S. Geol. Surv.* Open File Rep. 02-184.

Beavan J, Motagh M, Fielding EJ, Donnelly N, and Collett D (2012) Fault slip models of the 2010-2011 Canterbury, New Zealand, earthquakes from geodetic data and observations of postseismic ground deformation. *New Zealand Journal of Geology and Geophysics* 55(3): 207-221.

Boulanger RW and Idriss IM (2006) Liquefaction susceptibility criteria for silts and clays. *Journal of Geotechnical and Geoenvironmental Engineering* 132(11): 1413-1426.

Boulanger RW and Idriss IM (2014) *CPT and SPT Based Liquefaction Triggering Procedures*, Report No. UCD/CGM-14/01, Center for Geotechnical Modeling, University of California, Davis, CA.

Boulanger RW and Idriss IM (2016) CPT-based liquefaction triggering procedure. *Journal of Geotechnical and Geoenvironmental Engineering*, 142(2): 04015065-04015065.

Bourgeois, J. and Johnson, S.Y. (2001). "Geologic evidence of earthquakes at the Snohomish delta, Washington, in the past 1200 yr." *Geological Society of America Bulletin* 113(4), 482-494.

Bradley BA (2013) A New Zealand-Specific Pseudospectral Acceleration Ground-Motion Prediction Equation for Active Shallow Crustal Earthquakes Based on Foreign Models. *BSSA* 103(3): 1801-1822.

Bradley B.A. and Hughes M. (2012). "*Conditional Peak Ground Accelerations in the Canterbury Earthquakes for Conventional Liquefaction Assessment*." Technical Report Prepared for the Department of Building and Housing.

Brandenberg SJ, Zimmaro P, Stewart JP, Kwak DY, Franke KW, Moss RE, ... and Kramer SL (2020) Next-generation liquefaction database. *Earthquake Spectra*, 36(2): 939-959.

Bray JD, Sancio R, Kammerer AM, Merry S, Rodriguez-Marek A, Khazai B, ... and Nykamp M (2001) Some observations of the geotechnical aspects of the February 28, 2001, Nisqually earthquake in Olympia, South Seattle, and Tacoma, Washington. Report sponsored by NSF, PEER, UCB, University of Arizona, Washington State University, Shannon and Wilson Inc., and Leighton and Associates.

California Geological Survey (2014) Fault Activity Map of California. Scale not given. Available at: https://maps.conservation.ca.gov/cgs/fam/.

Caprio M., B. Tarigan, C.B. Worden, S. Wiemer, and D.J. Wald (2015). Ground motion to intensity conversion equations (GMICEs): A global relationship and evaluation of regional dependency, *Bull. Seismol. Soc. Am.* 105, no. 3, 1476-1490, doi: 10.1785/0120140286

Carena S, Suppe J (2002). Three-dimensional imaging of active structures using earthquake aftershocks: the Northridge thrust, California. Journal of Structural Geology 24:887-904

Chiou BS-J and Youngs RR (2014) Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. *EQS* 30(3): 1117-1153.

Clague, J.J., Naesgaard, E., and Nelson, A.R. (1997). "Age and significance of earthquake-induced liquefaction near Vancouver, British Columbia, Canada." Can. Geotech. J. 34: 53-62.

Clague, J.J., Naesgaard, E., and Sy, A. (1992). "Liquefaction features on the Fraser delta: evidence for prehistoric earthquakes?" *Canadian Journal of Earth Sciences* 29: 1734-1745.

Cox, R.T., Larsen, D., Forman, S.L., Woods, J., Morat, J., and Galluzzi, J., (2004). "Preliminary assessment of sand blows in the southern Mississippi embayment." *Bulletin of the Seismological Society of America* 93(3), 1125–1142.

Cramer, C.H. (2020). Updated GMICE for central and eastern North America extending to higher intensities, *Seimol. Res. Lett.* **91**, no. 6, 3518-3527, doi: 10.1785/0220200061

Cramer, C.H., and O.S. Boyd (2014). Why the New Madrid earthquakes are **M** 7-8 and the Charleston earthquake is ~**M** 7, *Bulletin of the Seismological Society of America*, **104**, no. 6, 2884-2903, doi: 10.1785/0120120257

Davis E. (2018). Evidence for liquefaction and flooding in the past 1,000 years along the Duwamish River, Seattle, Washington. *M.S. thesis*, University of Washington.

DeLong ER, DeLong DM and Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44: 837-845.

Durá-Gómez, I., and P. Talwani (2009a). Finding faults in the Charleston area, South Carolina: 1. seismological data. *Seismological Research Letters* **80**, no. 5, 883–900, doi: 10.1785/gssrl.80.5.883

Durá-Gómez, I., and P. Talwani (2009b). Finding faults in the Charleston area, South Carolina: 2. complementary data. *Seismological Research Letters* **80**, no. 5, 901–919, doi: 10.1785/gssrl.80.5.901

Dutton, C.E. (1889). The Charleston earthquake of August 31, 1886: *U.S. Geological Survey*, Ninth Annual Report 1887-88, 203-528.

Ellis, C., and de Alba, P. (1999). "Acceleration distribution and epicentral location of the 1755 'Cape Ann' earthquake from case histories of ground failure." *Seismological Research Letters* 70(6), 758-773.

Frankel, A., Wirth, E., Marafi, N., Vidale, J., and Stephenson, W. (2018). "Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust Based On 3D Simulations and Stochastic Synthetics (Part 1): Methodology and Overall Results." *Bulletin of the Seismological Society of America* 108 (5A): 2347-2369.

GeoEngineers, Inc. (2003) Summary Geotechnical Report Ground Improvement and Foundation Support Old Seattle Chocolate Factory Site, Seattle, Washington. July 22, 2003. Available from Washington State

Geyin M and Maurer BW (2020a) Fragility functions for liquefaction-induced ground failure. *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.

Geyin M, Baird AJ, and Maurer BW (2020b) Field assessment of liquefaction prediction models based on geotechnical vs. geospatial data, with lessons for each. *Earthquake Spectra* 36(3): 1386–1411.

Geyin M, Maurer BW, Bradley BA, Green RA, and van Ballegooy S (2021) CPT-based liquefaction case histories compiled from three earthquakes in Canterbury, New Zealand. *EQS:* 8755293021996367.

Gheibi, E., S. Gassman, and P. Talwani (2020). Regional assessment of prehistoric earthquake magnitudes in the South Carolina Coastal Plain, *Bulletin of Engineering Geology and the Environment*, **79**, no. 3, 1413-1427, doi: 10.1007/s10064-019-01627-7

Goldfinger, C., Nelson, C.H., Morey, A.E., Johnson, J.E., Patton, J.R., Karabanov, E., Gutiérrez-Pastor, J., Eriksson, A.T., Gràcia, E., Dunhill, G., Enkin, R.J., Dallimore, A., and Vallier, T., (2012). "Turbidite event history - methods and implications for Holocene paleoseismicity of the Cascadia subduction zone." *U.S. Geological Survey Professional Paper 1661–F*, 170 pg.

Goulet, C., Y. Bozorgnia, N. Abrahamson, N. Kuehn, L. Al Atik, R. Youngs, R.W. Graves, G.M. Atkinson (2018). Central and eastern North America ground-motion characterization (NGA-East Final Report, PEER Report Number 2018/08). Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California.

Green, R.A., Bommer, J.J., Rodriguez-Marek, A., Maurer, B.W., Stafford, P.J., Edwards, B., Kruiver, P.P, De Lange, G., and Van Elk, J. (2019). "Addressing limitations in existing 'simplified' liquefaction triggering evaluation procedures: application to induced seismicity in the Groningen gas field." *Bulletin of Earthquake Eng* 17(8), 4539-4557.

Green, R.A., Olson, S.M., and Obermeier, S.F., (2005). "Geotechnical analysis of paleoseismic shaking using liquefaction effects: field examples." *Engineering Geology* 76, 263–293.

Harmon, J., Y.M.A., Hashash, J.P. Stewart, E.M. Rathje, K.W. Campbell, W.J. Silva, and O. Ilhan (2019). Site Amplification Functions for Central and Eastern North America – Part II: Modular Simulation-Based Models, *Earthquake Spectra*, 35, no. 2, 815-847, doi: 10.1193/091117EQS179M

Hausler EA and Koelling M (2004) Performance of Improved Ground During the 2001 Nisqually Earthquake. International Conference on Case Histories in Geotechnical Engineering, New York, New York 18. https://scholarsmine.mst.edu/icchge/5icchge/session03/18

Hu, K., S.L. Gassman, and P. Talwani (2002). Magnitudes of prehistoric earthquakes in the South Carolina coastal plain from geotechnical data, *Seismological Research Letters*, **73**, no. 6, 979-991, doi: 10.1061/9780784479087.112

Hayati, H. and R.D. Andrus (2008). Liquefaction potential map of Charleston, South Carolina based on the 1886 earthquake, *Journal of Geotechnical and Geoenvironmental Engineering*, 134, no. 6, 815-828, doi: 10.1061/(ASCE)1090-0241(2008)134:6(815)

Idriss, I.M., and Boulanger, R.W. (2008). "*Soil liquefaction during earthquakes*." Monograph MNO-12 2008; Earthquake Engineering Research Institute, Oakland, CA, 261 pp.

Iwasaki T, Tatsuoka F, Tokida K, and Yasud S (1978) A practical method for assessing soil liquefaction potential based on case studies at various sites in Japan. *2nd Intl Conf. Microzonation*.

Jeschke, D.A., Eungard, D.W., Troost, K.G., Wisher, A.P. (2019). Subsurface Database of Washington State - GIS data. Washington Geological Survey Digital Data Series 11, version 2.1, previously released August 2017. http://www.dnr.wa.gov/publications /ger_portal_subsurface_database.zip.

Keefer, DK (1984) Landslides caused by earthquakes. Geol. Soc. Am. Bull. 95:406-421

Kelsey, H.M., Witter, R.C., Hemphill-Haley, E. (2002). "Plate-boundary earthquakes and tsunamis of the past 5500 yr, Sixes River estuary, southern Oregon. *GSA Bulletin* 114(3): 298-314.

Kuhn, G.G. (2005). "Paleoseismic features as indicators of earthquake hazards in North Coastal, San Diego County, California USA." *Engineering Geology* 80, 115-150.

Langridge RM, Ries WF, Litchfield NJ, Villamor P, Van Dissen RJ, Barrell DJA, Rattenbury MS, Heron DW, Haubrock S, Townsend DB, Lee JM, Berryman KR, Nicol A, Cox SC, and Stirling MW (2016) The New Zealand Active Faults Database. *New Zealand Journal of Geology and Geophysics* 59(1): 86-96.

Li, W., and Assimaki, D. (2010). "Site- and motion-dependent parametric uncertainty of site-response analyses in earthquake simulations." *BSSA*, 100(3): 954–968.

Marafi, N.A., Eberhard, M.O., Berman, J.W., Wirth, E.A., and Frankel, A.D. (2019). “Impacts of simulated M9 Cascadia subduction zone motions on idealized systems.” *Earthquake Spectra 35*(3), 1261-1287.

Marafi, N.A., Makdisi, A.J., Eberhard, M.O., & Berman, J.W. (2020). “Impacts of an M9 Cascadia Subduction Zone Earthquake and Seattle Basin on Performance of RC Core Wall Buildings.” *Journal of Structural Engineering* 146(2), 04019201.

Marafi, N. A., Grant, A., Maurer, B. W., Rateria, G., Eberhard, M. O., & Berman, J. W. (2021). “A generic soil velocity model that accounts for near-surface conditions and deeper geologic structure.” *Soil Dynamics and Earthquake Engineering 140*: 106461.

Martin, M.A., and Bourgeois, J. (2012). Vented sediments and tsunami deposits in the Puget Lowland, Washington – differentiating sedimentary processes. Sedimentology 59: 419-444.

Martin, J.R., II, and G.W. Clough (1994). Seismic parameters from liquefaction evidence, *J. Geotech. Engrg.*, 120, no. 8, 1345-1361, doi: 10.1061/(ASCE)0733-9410(1994)120:8(1345)

Maurer, B.W., Green, R.A. and Taylor, O.D.S. (2015). “Moving towards an improved index for assessing liquefaction hazard: lessons from historical data.” *Soils and Foundations* 55(4): 778-787.

Maurer BW, Green RA, van Ballegooy S and Wotherspoon L (2019) Development of region-specific soil behavior type index correlations for evaluating liquefaction hazard in Christchurch, New Zealand. *Soil Dynamics and Earthquake Engineering* 117: 96-105.

Melgar, D. (2021). Was the January 26th, 1700 Cascadia earthquake part of a rupture sequence? Journal of Geophysical Research: Solid Earth, 126, e2021JB021822. https://doi.org/10.1029/2021JB021822

Montalva G, Ruz F, Escribano D, Bastías N, Espinoza D, and Paredes F (2022) Chilean liquefaction case history database. *Earthquake Spectra* 38(3): 2260-2280. doi: 10.1177/87552930211070313

Morton DM (1971) *Seismically triggered landslides in the area above the San Fernando Valley, in The San Fernando, California, Earthquake of February 9, 1971*, U.S. Geological Survey Professional Paper 733, p. 99-109

Moss R.E.S., Seed R.B., Kayen R.E., Stewart J.P., Der Kiureghian A., Cetin K.O. (2006). “CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential.” *Journal of Geotechnical & Geoenvironmental Engineering*, 132(8), 1032-1051.

Murphy T (2011) FEMA: freaking over a fake quake? *Mother Jones*. available at: < http://www.motherjones.com/politics/2011/03/fema-earthquake-new-madrid/>.

Nowicki Jessee MA, Hamburger MW, Allstadt K, Wald DJ, Robeson SM, Tanyas H, Hearne M, Thompson EM (2018) A Global Empirical Model for Near-Real-Time Assessment of Seismically Induced Landslides. *Journal of Geophysical Research: Earth Surface* 123(8):1835-1859. https://doi.org/10.1029/2017JF004494

Nuttli, O.W., G.A. Bollinger, and R.B. Herrmann (1986). The 1886 Charleston, South Carolina, earthquake: A 1986 perspective, *U.S. Geological Survey*, **985**, doi: 10.3133/cir985

Obermeier, S.F., and Dickenson, S.E., (2000). "Liquefaction evidence for the strength of ground motions resulting from the late Holocene Cascadia subduction earthquakes, with emphasis on the event of 1700 A.D." *Bulletin of the Seismological Society of America* 90(4), 876–896.

Pacific Northwest Seismic Network (2001) Preliminary observations of ground deformation, Nisqually earthquake. Available at: https://assets.pnsn.org/notable/WEBDIR_01022818543p/maps/gd.jpg (Acc. Jan. 31, 2022).

Petersen MD, Frankel AD, Harmsen SC, Mueller CS, Haller KM, Wheeler RL, Wesson RL, Zeng Y, Boyd OS, Perkins DM, Luco N, Field EH, Wills CJ, Rukstales KS (2008) *Documentation for the 2008 update of the United States national seismic hazard maps.* USGS Open-File Report 2008-1128, 61 p.

Petersen MD, Moschetti MP, Powers PM, Mueller CS, Haller KM, Frankel AD, Zeng Y, Rezaeian S, Harmsen SC, Boyd OS, Field N, Chen R, Rukstales KS, Luco N, Wheeler RL, Williams RA and Olsen AH (2014) *Documentation for the 2014 update of the United States national seismic hazard maps.* USGS Open-File Report 2014–1091, 243 p.

Petersen, M.D., A.M. Shumway, P.M. Powers, C.S. Mueller, M.P. Moschetti, A.D. Frankel, … Y. Zeng (2020). The 2018 update of the US National Seismic Hazard Model: Overview of model and implications, *Earthquake Spectra*, **36**, no. 1, 5-41, doi: 10.1177/8755293019878199

Peterson, C.D., et al. (2014). "Large-Scale Fluidization Features from Late Holocene Coseismic Paleoliquefaction in the Willamette River Forearc Valley, Central Cascadia Subduction Zone, Oregon, USA." *Open Journal of Earthquake Research*, 3: 82-99.

Peterson C.D. and Madin, I.P. (1997). "Coseismic Paleoliquefaction Evidence in the Central Cascadia Margin, USA." *Oregon Geology* 59, 51-74.

Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.

Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.

Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615

Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040

Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.

Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra,* 10.1177/87552930231174244, EERI.

Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.

Rasanen, R.A. and Maurer, B.W. (2023). “Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating.” *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.

Rashidian, V., and Baise, L.G. (2020). “Regional efficacy of a global geospatial liquefaction model.” *Engineering Geology* 272: 105644.

Robertson PK and Wride CE (1998) Evaluating cyclic liquefaction potential using cone penetration test. *Canadian Geotechnical Journal* 35(3): 442-459.

Seattle Office of Emergency Management (2019) Seattle Hazard Identification and Vulnerability Analysis (SHIVA). April 10, 2019. Available from: https://www.seattle.gov/emergency-management/hazards, last accessed May 10, 2022.

Shi, J., and D. Asimaki (2017), “From stiffness to strength: Formulation and validation of a hybrid hyperbolic nonlinear soil model for site-response analyses.” *BSSA*, 107(3): 1336–1355.

Stein S (2008) Presentation to Arkansas legislative committee on economic development on earthquake safety provisions in building code. Little Rock, Arkansas; January 28, 2008. Available at: < http://www.earth.northwestern.edu/people/seth/research/testimony.html>.

Talwani, P., and Schaeffer, W.T., (2001). “Recurrence rates of large earthquakes in the South Carolina coastal plain based on paleoliquefaction data.” *Journal of Geophysical Research* 106, 6621–6642.

Tuttle, M.P., Dyer-Williams, K., and Barstow, N.L., (2002). “Paleoliquefaction study of the Clarendon-Lindon fault system, western New York state.” *Tectonophysics* 353, 263–286.

Tuttle, M.P., Schweig, E.S., Campbell, J., Thomas, P.M., Sims, J.D., and Lafferty, R.H., (2005). “Evidence for new Madrid earthquakes in A.D. 300 and 2350 B.C.” *Seismological Research Letters* 76, 489–501.

Updike RG, Brown WM, Johnson ML, Omdahl EM, Powers PS, Rhea S, Tarr AC (1996) USGS Response to an Urban Earthquake, Northridge '94. U.S. Geological Survey Open-File Report 96-263, 78 p

USGS (2017). “M 9.3 scenario earthquake – Cascadia megathrust – whole CSZ characteristic M branch.” *USGS National Earthquake Information Center*, available from < https://earthquake.usgs.gov/scenarios/eventpage/bssc2014cascadia_sub0_m9p34_se#executive>.

USGS (2022). Advanced National Seismic System (ANSS) Comprehensive Catalog of Earthquake Events and Products: Various. U.S. Geological Survey, Earthquake Hazards Program. doi: 10.5066/F7MS3QZH

Van Ballegooy, S., Malan, P., Lacrosse, V., Jacka, M.E., Cubrinovski, M., Bray, J.D., O’Rourke, T.D., Crawford, S.A., and Cowan, H. (2014). “Assessment of liquefaction-induced land damage for residential Christchurch.” *Earthquake Spectra*, 30(1): 31-55.

Vidale J, Atkinson G, Green R, Hetland E, Grant-Ludwig L, Mazzotti S, Nishenko S and Sykes L (2011) *Report of the independent expert panel on New Madrid Seismic Zone earthquake hazards as approved by NEPEC on April 16, 2011*. U.S. Geological Survey, 26 p.

Wald DJ, Helmberger DV and Heaton TH (1991) Rupture Model of the 1989 Loma Prieta Earthquake From the Inversion of Strong-Motion and Broadband Teleseismic Data. *BSSA* (5): 1540-1572.

Walsh, TJ (2001). Working a geologic disaster. Washington Geology 28(3): 6-18.

Wentworth CM, Zoback MD (1990). Structure of the Coalinga Area and Thrust Origin of the Earthquake. In: Rymer MJ, Ellsworth WL (eds.) The Coalinga, California, Earthquake of May 2, 1983, U.S. Geological Survey Professional Paper 1487, p 417

Wirth, E.A., Grant, A., Marafi, N.A., and Frankel, A.D. (2021). "Ensemble ShakeMaps for Magnitude 9 Earthquakes on the Cascadia Subduction Zone." *Seismological Research Letters*, 92(1)

Wirth, E.A., Frankel, A.D., Marafi, N., Vidale, J. E., and Stephenson, W. J. (2018). "Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust based on 3-D Simulations and Stochastic Synthetics (Part 2): Rupture Parameters and Variability." *Bulletin of the Seismological Society of America*, 108 (5A): 2370-2388.

Wong, I., J. Bouabid, W. Graf, C. Huyck, A. Porush, W. Silva, T. Siegel, G. Bureau, R. Eguchi, and J. Knight (2005). Potential losses in a repeat of the 1886 Charleston, South Carolina, earthquake, *Earthquake Spectra*, **21**, no. 4, 1157-1184, doi: 10.1193/1.2083907.

Worden, C.B., M.C. Gerstenberger, D.A. Rhoades, and D.J. Wald (2012). Probabilistic relationships between ground-motion parameters and Modified Mercalli Intensity in California, *Bull. Seismol. Soc. Am*. **102**, no. 1, 204-221, doi: 10.1785/0120110156.

Zhu, J., Baise, L.G., and Thompson, E.M. (2017). "An updated geospatial liquefaction model for global application." *Bulletin of the Seismological Society of America*, 107(3): 1365–1385.

APPENDIX A

Publication 1 (Conference Paper 1), highlighted below, is provided on the following pages in full (as a manuscript that has been accepted for publication).

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). “Seismic Source Parameters from Regional Paleoseismic Evidence.” *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). “Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence.” *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). “Probabilistic Seismic Source Inversion from Regional Landslide Evidence.” *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). “Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating.” *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). “Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest.” *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). “Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2).” *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). “Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance.” *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). “Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards.” *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.

[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

Seismic Source Parameters from Regional Paleoseismic Evidence

Ryan A. Rasanen, S.M.ASCE[1] and Brett W. Maurer, A.M.ASCE[2]

[a] Graduate Research Assistant, University of Washington, Seattle; rrasanen@uw.edu
[2] Assistant Professor, University of Washington, Seattle; bwmaurer@uw.edu

ABSTRACT

Historic earthquake catalogs are, in many regions, insufficient to constrain seismic source locations and magnitude-frequency relations, which are needed for seismic-hazard analyses. As a result, analyses of relic coseismic evidence, such as landslides, rockfall, and liquefaction, must often be used to infer information about the seismic hazard. However, while such studies are performed widely, they have tended to focus more on field identification, dating, and ascription of evidence to a seismic cause, with less effort focused on advancing the techniques for inverse-analyzing field evidence to elucidate the seismic-source model. Accordingly, this paper proposes a general inversion framework by which seismic source parameters are probabilistically constrained from paleoseismic evidence. Analyzing evidence at regional scale leads to: (i) a geospatial likelihood surface that constrains the causative rupture location; and (ii) a probability distribution of the rupture magnitude, each being direct inputs to probabilistic seismic hazard analyses. Simulated paleoseismic studies are performed on modern earthquakes with known seismic parameters. These simulations demonstrate the utility and efficacy of the proposed framework, which has the potential to provide new insights in enigmatic seismic zones worldwide.

INTRODUCTION

Because the return period of moderate-to-large earthquakes is often longer than the duration of recorded seismicity, earthquake catalogs may be insufficient to inform the inputs needed for probabilistic seismic-hazard analyses. In turn, the uncertain locations, magnitudes, and frequencies of fault ruptures influence the computed seismic hazard (Vidale et al. 2011), which is adopted by building codes and used by engineers and policy makers. Whether the uncertainty is biased toward greater-than-actual or lesser-than-actual seismic hazards, the consequences are significant. Either capital is spent superfluously on seismic design and rehabilitation (Stein 2008; Murphy 2011) or lives and infrastructure are at risk. For example, a M_W7.7 New Madrid earthquake is projected to cause economic losses of $300 billion - triple that of 2005 Hurricane Katrina (Elnashai et al. 2009).

In many regions of seismic-hazard uncertainty, earthquake scientists and engineers are thus resigned to study paleoseismic evidence (i.e., left by pre-historic and/or pre-instrumental earthquakes), which can be used to infer information about the causative ruptures. A complete study of such evidence (e.g., soil liquefaction, landslides, and rockfall, among other types) is three phased: (i) field identification and interpretation; (ii) dating; and (iii) constraint of the earthquake magnitude and/or ground motion under which the evidence field was formed. However, research to-date has focused more on field investigations and dating, with less effort focused on advancing the techniques for inverse-analyzing field evidence. The focus of this study is on the latter.

Notably, all existing techniques to inverse-analyze paleoseismic evidence are underpinned by the need to locate the earthquake that induced the evidence, yet none prescribe a quantitative approach to probabilistically constrain the rupture location, which is invariably uncertain. This in-turn leads to assumptions that could be inaccurate, leading to mischaracterizations of the causative ground motions and/or earthquake magnitude. Moreover, the joint probability of rupture location and magnitude, as inverted from field evidence, has never been investigated or quantified, despite this being the optimal input to a seismic hazard analysis. To date, all existing inverse-analysis techniques and studies have been fully or predominantly deterministic, the shortcomings of which are further discussed in Rasanen and Maurer (2021a; 2021b). Accordingly, this study presents a general framework for probabilistically constraining seismic source parameters from paleoseismic evidence. Specifically, this framework computes the likelihood of a rupture with some location, geometry, and magnitude producing a set of field observations. Repeating for a near-infinite number of possibilities results in a regional-scale quantification of possible source parameters. In the following, an overview of the framework is first given, followed by tests of the framework on two modern earthquakes (i.e., simulated paleoseismic studies are performed).

METHODOLOGY

The proposed framework is described in two parts. The first part presents a conceptual overview and brief hypothetical example. The second part provides implementation details needed to apply the framework and describes the validation tests that will be performed subsequently.

Overview of Proposed Framework

The proposed framework aims to probabilistically constrain paleoseismic source parameters from a near-infinite number of possibilities. This will be done by computing the likelihood that a rupture with given location, geometry, and magnitude would produce a set of field observations (sites with and without evidence), wherein various measurement and modelling uncertainties may be included. In general, the likelihood of a parameter having a value, given a set of observations, is the product of the probabilities of those observations, conditioned on the parameter value. Thus, the likelihood of a rupture having some location (L), geometry (G), and magnitude (M_w), given a set (x) of field observations at *N* different sites, is:

$$Likelihood\ (L, G, M_w|x) = P(X = x|L, G, M_w) = \prod_{i=1}^{N} P(X_i = x_i|L, G, M_w) \qquad (1)$$

Where $P(X_i = x_i|L, G, M_w)$ is the probability of the observation at site *i* (the observed presence or absence of field evidence) given a rupture with parameters *L*, *G*, and M_w. By repeating for a near-infinite number of possibilities, the actual rupture parameters are probabilistically constrained via the likelihood function (product of the probabilities of *N* observations), such that different combinations of *L*, *G*, and M_w will be found more and less likely to produce the observed evidence. In Eq. (1), the probability of a field observation, conditioned on all input variables, is computed at sites with and without field evidence using Eqs. 2a and 2b, respectively:

$$P(\text{Field Evid}|\text{EQK}: L, G, M_w) = \int_{IM} P(\text{Field Evid}|IM, M_w)\, f(IM|L, G, M_w) \cdot \mathrm{d}IM \qquad (2a)$$

$$P(\text{No Field Evid}|\text{EQK}: L, G, M_w) = 1 - \int_{IM} P(\text{Field Evid}|IM, M_w)\, f(IM|L, G, M_w) \cdot \mathrm{d}IM \qquad (2b)$$

Where $f(IM|L, G, M_w)$ is a probability density function (PDF) of an intensity measure (IM) computed by a ground motion model (GMM) considering site response, and conditioned on fault parameters L, G and M_w; and $P(\text{Field Evid}|IM, M_w)$ is the probability of observing field evidence at a site, given ground-motion parameters IM and M_w, as computed by a hazard model (e.g., a liquefaction, landslide, or rockfall prediction model). Expanding on Eqs. 2a and 2b, other uncertainties could similarly be incorporated (e.g., measurement uncertainties).

The framework's application is shown conceptually in Figure 1 considering four hypothetical field sites, where field evidence is observed at sites one and two, but not at sites three and four, and two possible source locations for the earthquake that produced the observations. In actual analyses, a near-infinite number of source locations could be considered. Figure 2 illustrates how the relative likelihoods of locations one and two are assessed. Shown in Figure 2a are the probabilities of individual field observations, given an earthquake at location one, as computed by Eq. 2 for varying M_w. At sites with field evidence, the probability of the field observation increases with increasing M_w, whereas at sites without field evidence, the opposite occurs. In Figure 2b, this process is repeated considering an earthquake at source location two. In Figure 2c, the likelihood of each source location is computed as a function of M_w, per Eq. 1 (the product of the black and gray curves). It can be seen that location one has a far greater peak likelihood, whereas an earthquake at location two is unlikely to produce the field observations, regardless of its M_w. Repeating this process at a near-infinite number of locations, a geospatial surface of likelihood is developed, constraining the location of fault-rupture, to the degree the evidence permits. Moreover, for any location considered, the likelihood distribution of earthquake magnitude is computed, as shown in Figure 2c. These outcomes (i.e., probability distributions of earthquake location and magnitude) are direct inputs to probabilistic seismic hazard analyses.

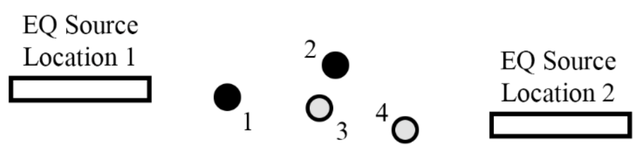


Figure 1. Hypothetical analysis of four field sites, where paleoseismic evidence was observed at sites one and two (black circles), but not at sites three and four (gray circles). Also shown are two possible sources for the earthquake producing these observations.

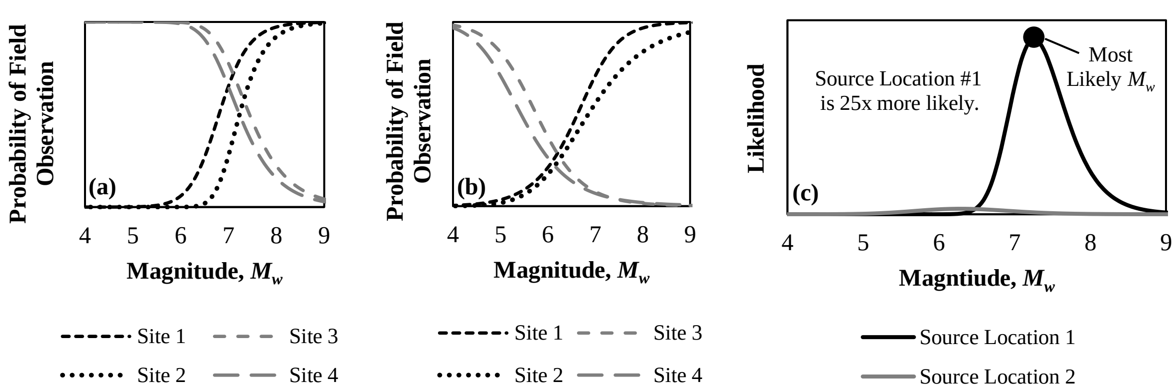


Figure 2. Conceptual illustration of the proposed approach to constraining the seismic source that produced the observations in Figure 1: (a) probabilities of field observations, given an event at location one of varying M_w; (b) probabilities of field observations, given an event at location two of varying M_w; (c) likelihood vs. M_w for locations one and two.

Implementation Details

The implementation of Eqs. 1 and 2 is next discussed in detail. This includes general concepts, as well as the specific methods and models adopted herein for application to two modern earthquakes: (i) the 2011 M_w6.3 Christchurch, New Zealand, earthquake, for which liquefaction evidence will be analyzed; and (ii) the 1971 M_w6.6 San Fernando, USA earthquake, for which landslide evidence will be analyzed. Owing to its modular form, the framework easily accommodates models other than those selected for use below. To implement the framework, an analyst must:

(1) Create an array of possible source locations. In this work, a 250 km by 250 km grid of surficial points, which will be treated as possible earthquake epicenters, was adopted and centered on the evidence field. Within the grid area (62,500 km^2 in this study), a finer point spacing gives greater spatial resolution while a coarser spacing allows for faster runtime. Balancing these ideals, a l km spacing was adopted for a 10,000 km^2 area at the center of the grid, while a 10 km spacing was adopted for the remainder of the 62,500 km^2 study area.

(2) Select *N* study sites where: (i) the presence or absence of field evidence is observed; and (ii) the required inputs to models for predicting the evidence are available (i.e., inputs to liquefaction or landslide models). Ideally, and as implemented in this work, an equal number of sites with and without field evidence is selected to limit sampling bias.

(3) Select a GMM. In this paper, the Bradley (2013) and Chiou and Youngs (2014) GMMs were adopted for the 2011 Christchurch and 1971 San Fernando, earthquakes, respectively.

(4) For each provisional seismic-source locale created in (1):

(5) For each provisional seismic-source M_w being considered (a range of M_w5 to M_w9 was used):

(6) For each of *N* study sites selected in (2), cycling from i = 1 to N:

(7) Compute the site-to-source distance(s) required by the GMM chosen in (3), as measured from study site *i* to the provisional seismic sources. In this work, the required metrics include Joyner-Boore distance (R_{JB}) and the closest distance to fault rupture (R_{RUP}). Since the seismic sources generated in (1) are epicenters, rather than 3-dimensional faults, the correlations of Scherbaum et al. (2004) were used to estimate R_{JB} from epicentral distance (R_{EPI}), assuming a strike-slip fault with dip (δ) of 90°. The Scherbaum et al. (2004) correlations, which are magnitude dependent, effectively convert the point sources from (1) into a statistical realization of a 3-dimensional fault. R_{RUP} was calculated per the Pythagorean theorem using the estimated R_{JB} and the depth to top-of-rupture (Z_{TOR}). Z_{TOR} was estimated using the correlation of Kaklamanos et al. (2011), which requires as inputs the: down-dip rupture width (W); hypocentral depth (Z_{HYP}); and δ. Following the recommendation of Kaklamanos et al. (2011), W and Z_{HYP} were respectively estimated per the methods of Wells and Coppersmith (1994) and Scherbaum et al. (2004).

(8) Using the GMM selected in (3) and site-to-source distances from (7), compute the PDF of expected *IM* at study site *i*, wherein the time-averaged shear wave velocity over the top 30 m (V_{S30}) is consistent with site *i*. In this paper, the fault source was assumed to have strike-slip parameters (δ = 90° and rake, λ = 0°) and V_{S30} was estimated per Heath et al. (2020). In general, the probability distribution of *IM* is assumed to be a lognormal random variable defined by the median and lognormal standard deviation, which are provided by a GMM.

(9) For each possible *IM* value at study site *i*, as computed in (8) for a given M_w and distance pair:

(10) Compute the probability of a field observation using Eq. 2a or 2b, depending on whether evidence was or was not observed at study site *i*. In this work, the probability of landsliding was computed using the Nowicki et al. (2018) model, which is a function of peak ground velocity (*PGV*). To compute the probability of liquefaction, the safety factor against liquefaction, FS_{liq}, was first computed throughout the soil profile using the Boulanger and Idriss (2014) CPT-based liquefaction triggering model, the inputs to which include peak ground acceleration (*PGA*) and $M_{w.}$ However, because liquefaction may trigger at depth but be inconsequential, or unobserved, at the ground surface, it is best to predict surficial manifestations of liquefaction (i.e., liquefaction vents and dikes at or near the surface). While several manifestation models are available, *LPI* was adopted in this paper and is defined as (Iwasaki et al. 1978):

$$LPI = \int_0^{20\,m} F\left(FS_{liq}\right) \times w(z)dz \tag{3}$$

Where $F\left(FS_{liq}\right)$ and $w(z)$ are functions that weight the respective influences of FS_{liq} and depth, z, on surface manifestation. Specifically, $F = 1 - FS_{liq}$ for $FS_{liq} \leq 1$ and $F = 0$ otherwise; $w(z) = 10 - 0.5z$, where z is the depth (m) below the ground surface. The probability of liquefaction manifestation was then computed using a fragility function of Geyin and Maurer (2020), which is conditioned on *LPI* and trained on all globally available liquefaction case histories. Using this function, the probability of surficial liquefaction manifestation is:

$$F_{MS}(LPI) = \Phi\left(\frac{ln(LPI) - \ln(\theta)}{\beta}\right) \tag{4}$$

Where $F_{MS}(LPI)$ is the probability of surface manifestation, Φ is the Gaussian cumulative distribution function, and θ and β are the distribution's median and logarithmic standard deviation, respectively. Geyin and Maurer (2020) separately trained functions using data from New Zealand and other global earthquakes. Accordingly, the New Zealand specific function (θ = 5.582, β = 1.537) was adopted for application to the Christchurch, New Zealand, earthquake. Repeating steps 6-10 for each provisional M_w being considered results in a probability of field observation versus M_w curve for each site, examples of which are shown in Figures 2a and 2b.

(11) Compute the likelihood of a provisional source locale, as a function of M_w, by multiplying the probabilities of field observations at all study sites (i.e., multiply the curves shown in Fig. 2a or 2b at each value of M_w). The result, an example of which is shown in Fig. 2c, is a PDF of the causative earthquake magnitude, *conditioned on a single seismic-source location*.

(12) Repeating steps 5-11 for all provisional seismic source locations created in (1) results in a likelihood distribution for each, as described in (11). By normalizing the peak likelihood at each source location by the peak likelihood among all locations, the relative likelihoods of source locations can be assessed, with the most likely location having a normalized value of one. The M_w corresponding to the peak likelihood (see Fig. 2c) is the median M_w likely to produce the field observations, given that an earthquake at the source location occurs.

(13) Finally, by aggregating PDFs from all potential source locations across the study area, an overall PDF of magnitude, *considering all possible source locations*, is produced. While a single source location will always be deemed "most likely," earthquakes at other source locations have potential to produce the evidence field, albeit with lesser likelihood. In general, this overall PDF will be wider (i.e., have greater uncertainty) than the PDF conditioned on the single, most likely source, especially when the source is not well constrained (e.g., due to limited field evidence).

SIMULATED APPLICATION

The proposed framework, as outlined above, provides probabilistic constraint on both the location and M_w of a seismic event considering paleoseismic evidence. To investigate the degree of successful constraint, simulated paleoseismic studies are performed on two modern events with known seismic sources. Inherent to these applications, it is assumed that the presence or absence of evidence would be accurately interpreted in a paleoseismic investigation, and moreover, that evidence would be properly attributed to a causative earthquake. These assumptions are not intended to diminish the importance of field interpretation and dating. The objective herein is to advance seismic-source inversion techniques, rather than to discuss or simulate the field investigations prerequisite for such analyses.

For each event the inverted probabilistic source location and magnitude will be compared to the actual, known parameters. However, no prior knowledge of the seismic source is incorporated into the analyses. In addition, because the two earthquakes are modern, the number of study sites available from reconnaissance (i.e., sites with and without observed evidence) far exceeds the number of field sites conceivable of a regional paleoseismic study. Accordingly, analyses are performed on a randomly selected sample of 100 study sites from each event (i.e., 50 sites with evidence and 50 sites without evidence). Because landslide inventories are comprised almost exclusively of positive observations, slopes within a given survey area are assumed not to have experienced slides if none are mapped. While this inevitably introduces uncertainty, it facilitates rapid implementation of the proposed framework and is not dissimilar from a paleoseismic study, wherein an absence of evidence may not be evidence of absence.

2011 Christchurch, New Zealand

The M_w6.3 Christchurch earthquake occurred on a blind, previously unknown fault beneath the southern environs of Christchurch, causing widespread damage and 185 fatalities (Quigley et al. 2016). CPT-based liquefaction case histories were obtained from the Geyin et al. (2021) inventory for analysis. An example of the framework's output is shown in Figure 3 for an analysis using 100 random study sites; the locations of these sites, and the surface projection of the 2011 fault rupture, are also shown. Mapped in Figure 3 are contours of back-calculated relative likelihood, wherein values near one and zero respectively denote source locations most and least likely to produce the observed evidence. As seen in Figure 3, the analysis identified a relatively small, well-constrained area just south of the evidence field as the likely source. This aligns with the actual rupture, such that the identified, most likely source falls within the surface projection of the faults that ruptured.

The median M_w conditioned on the most-likely source location was M_w6.4. Other M_w values could inevitably also produce the evidence field (e.g., if ground motions were higher or lower than the median expectation). Shown in Figure 4a is the full probability distribution of M_w, conditioned on the most likely source location, which is analogous to the conceptual PDFs shown in Figure 2c. The 95% confidence interval (CI) is M_w6.10- M_w6.65. Moreover, as evident in Figure 3, earthquakes elsewhere have potential to produce the evidence field, albeit with lesser likelihood. By aggregating individual PDFs from all locations across the 62,500 km^2 study grid, an overall PDF that includes source-location uncertainty is produced. This PDF, which is also shown in Figure 4a, has a median of M_w6.8 and 95% CI of M_w6.10- M_w7.90. Unless the causative fault can be constrained via on-fault (or some other) evidence, this latter, more uncertain PDF would be the more appropriate product of a paleoseismic analysis.

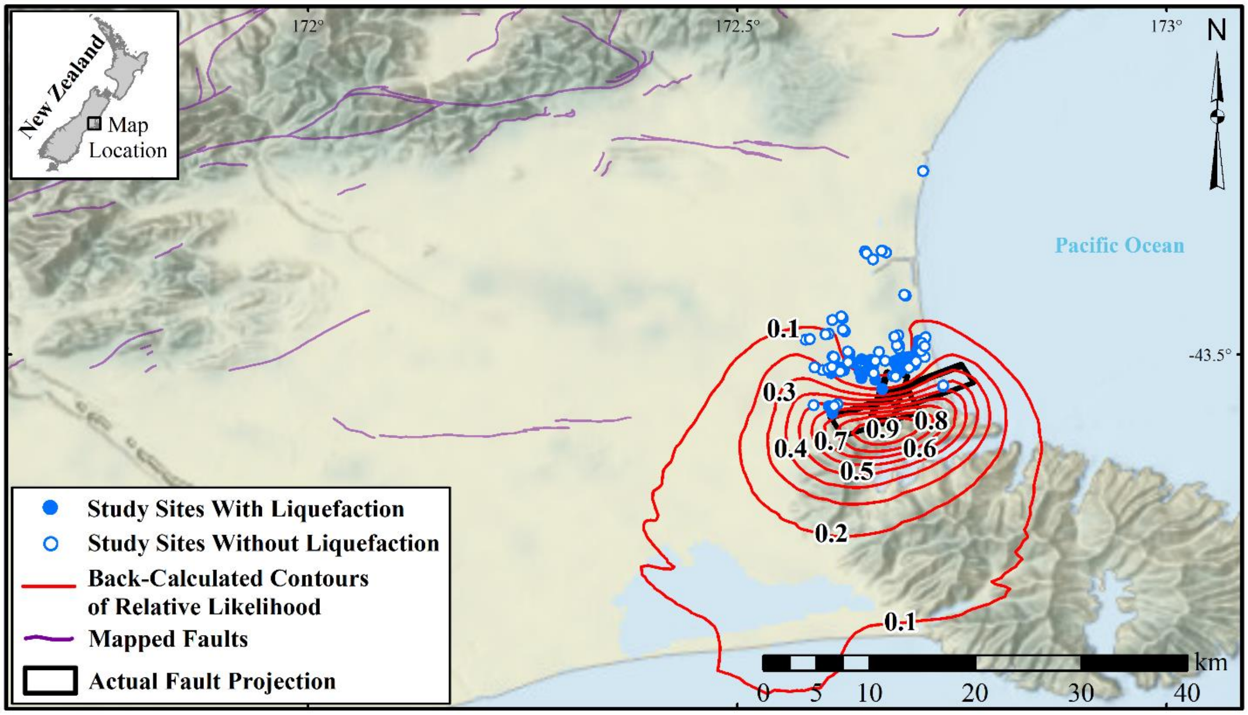


Figure 3. Representative analysis using 100 random study sites from the 2011 M_w6.3 Christchurch earthquake showing the relative likelihood contours of earthquake source locations. Black rectangles delineate the actual fault projection (Beavan et al. 2012); faint purple lines are other mapped faults in the region (Langridge et al. 2016).

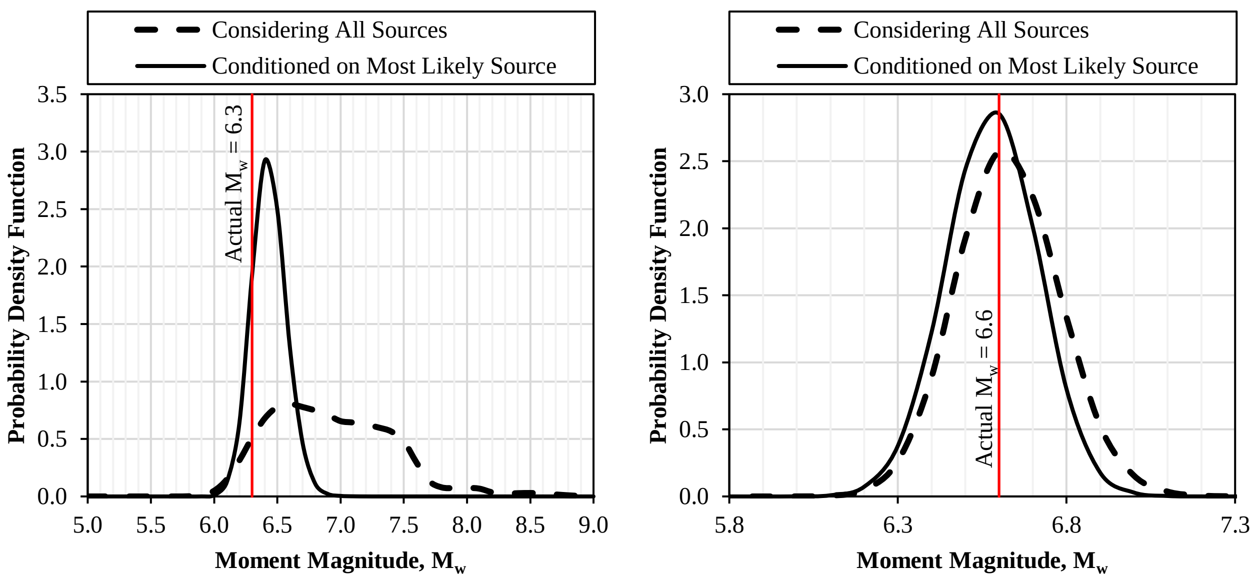


Figure 4. PDFs of earthquake magnitude inverted from (a) liquefaction evidence in the 2011 M_w6.3 Christchurch earthquake and (b) landslide evidence in the 1971 M_w6.6 San Fernando earthquake. Shown are PDFs conditioned on the most-likely source location as well as the aggregate PDFs considering source-location uncertainty.

1971 San Fernando, California

The 1971 M_w6.6 San Fernando, California earthquake occurred on the Sierra Madre Fault (Carena and Suppe 2002). An inventory of observed landslides was obtained from Morton (1971), as compiled by Schmitt et al. (2017), and is mapped in Figure 5. A representative example of the framework's output is shown in Figure 5 for an analysis using 100 randomly selected study sites; the surface projection of the 1971 fault rupture is also shown. Mapped in Figure 5 are contours of back-calculated relative likelihood, wherein values near one and zero respectively denote source locations most and least likely to produce the observed evidence. As seen in Figure 5, the analysis identified a relatively well-constrained area as the likely source. This aligns with the actual rupture, such that the identified most likely source falls within the surface projection of the fault that ruptured. The median M_w at the most-likely source given was M_w6.6 with a 95% CI of M_w6.25-M_w6.80, versus the actual M_w6.6. Similar to the 2011 Christchurch earthquake, this PDF – conditioned on the most-likely source location – and the overall PDF reflecting source-location uncertainty, are computed and shown in Figure 4b. The overall PDF considering source-location uncertainty also provides a median of M_w6.6 and a similar CI of M_w6.25- M_w6.90.

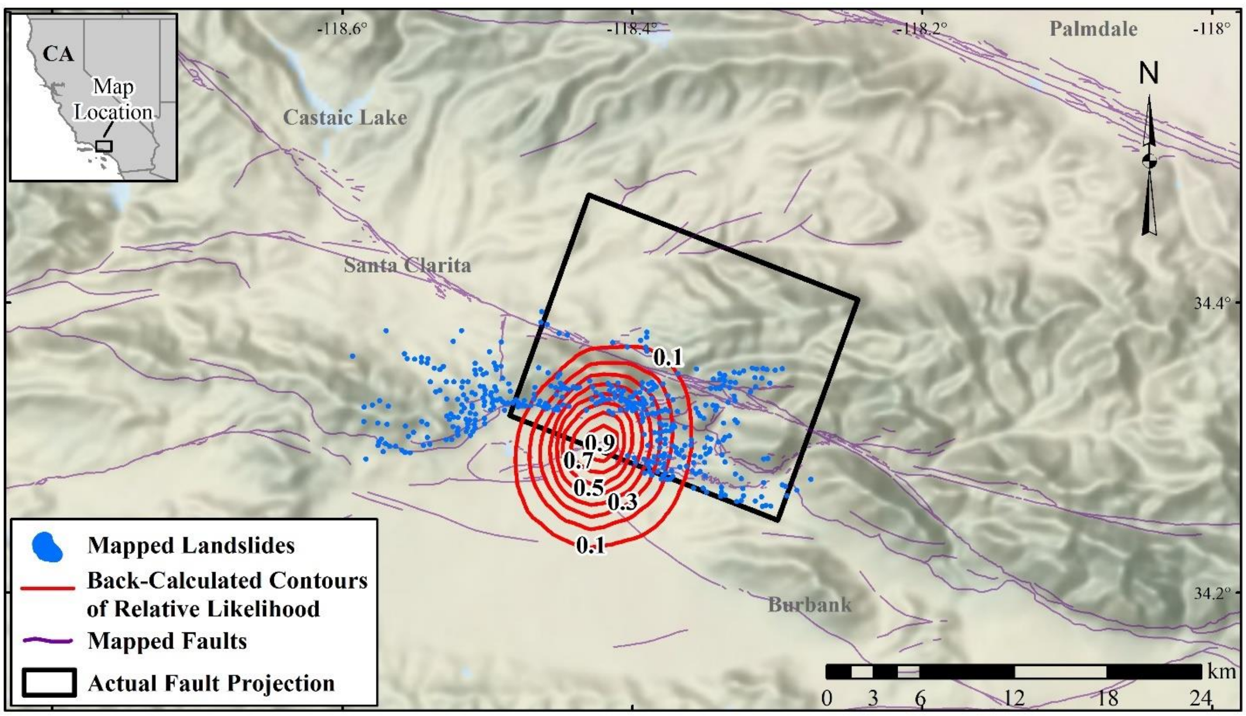


Figure 5. Representative analysis of 100 random study sites from the 1971 M_w6.6 San Fernando, USA earthquake showing the relative likelihood contours of earthquake source locations. Black rectangle is the actual fault projection (Carena and Suppe, 2002); faint purple lines are other mapped faults in the region (California Geological Survey 2014).

DISCUSSION AND CONCLUSIONS

Several types of paleoseismic evidence, such as paleoliquefaction and paleolandslides, are commonly used to infer information about the seismic hazard in regions where the recorded history of seismicity is short relative to the return periods of earthquakes. However, existing paleoseismic analysis techniques require assumptions about the location of the seismic source and do not consider its potential uncertainty. Accordingly, this paper introduced a framework to probabilistically constrain seismic source parameters using regional liquefaction and landslide evidence. This results in: (i) a geospatial likelihood surface that constrains the rupture location; and (ii) a probability distribution of the rupture magnitude. Simulated paleoseismic studies were performed on the 2011 Christchurch, New Zealand and 1971 San Fernando, USA earthquakes using liquefaction and landslide evidence, respectively. These simulated studies, which quantify the uncertainties of seismic-source magnitude and location, demonstrate both the utility and efficacy of the proposed framework for use in paleoseismic analysis.

In this paper, earthquake sources were modeled as epicenters and converted to statistical fault realizations. However, 3-D source models could be used with minimal modification. That is, the framework could quantify the likelihoods of faults with different strike, depth, length, etc. of producing an evidence field, provided the fault geometries result in different ground motions at study sites. As a framework that relies on predicting multiple phenomena in series (e.g., ground-motion attenuation, site response, liquefaction and landslides), inaccuracies are likely whenever the response of one or more phenomena differs greatly from expectations. However, as a modular framework, performance should improve over time as better component models are continuously developed (e.g., improved liquefaction models). While paleoseismic evidence has been investigated widely, the evidence has arguably not been exploited fully in the context of a fully probabilistic framework that quantifies the true seismic-source uncertainty. Towards that end, the proposed framework has the potential to provide new insights in seismic zones worldwide.

REFERENCES

Beavan J., Motagh M., Fielding E.J., Donnelly N., and Collett D. (2012). “Fault slip models of the 2010-2011 Canterbury, New Zealand, earthquakes from geodetic data and observations of postseismic ground deformation.” *New Zealand J. of Geol. & Geophysics* 55(3): 207-221.

Boulanger, R.W. and Idriss, I.M. (2014). *CPT and SPT based liquefaction triggering procedures*. Report No. UCD/CGM.-14/01, Center for Geotech. Modelling, CE Eng., UC Davis, CA.

Bradley B.A. (2013). “A New Zealand-Specific Pseudospectral Acceleration Ground-Motion Prediction Equation for Active Shallow Crustal Earthquakes Based on Foreign Models.” *BSSA* 103(3): 1801-1822. doi: 10.1785/0120120021.

California Geological Survey (2014). Fault Activity Map of California. Scale not given. Accessed April 9, 2021, at: https://maps.conservation.ca.gov/cgs/fam/

Carena S, Suppe J (2002). “Three-dimensional imaging of active structures using earthquake aftershocks: the Northridge thrust, California.” *Journal of Structural Geology* 24:887-904.

Chiou B.S.-J. and Youngs R.R. (2014). “Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra.” *Earthquake Spectra* 30(3): 1117-1153. doi: 10.1193/072813EQS219M.

Elnashai A.S., Cleveland L.J., Jefferson T., and Harrald J. (2009). *Impact of earthquakes on the Central USA*. Mid-America Earthquake Center Report 09-03, University of Illinois at Urbana Champaign, Urbana, IL.

Geyin M. and Maurer B.W. (2020). “Fragility functions for liquefaction-induced ground failure.” *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.

Geyin M., Maurer B.W., Bradley B.A., Green R.A., and van Ballegooy S. (2021). “CPT-based liquefaction case histories compiled from three earthquakes in Canterbury, New Zealand.” *Earthquake Spectra*. In Review.

Iwasaki T., Tatsuoka F., Tokida K., and Yasud S. (1978). “A practical method for assessing soil liquefaction potential based on case studies at various sites in Japan.” *2nd Intl Conf. Microzonation*.

Kaklamanos J., Baise L.G., & Boore D.M. (2011). “Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice.” *Earthquake Spectra* 27(4): 1219-1235.

Langridge R.M., Ries W.F., Litchfield N.J., Villamor P., Van Dissen R.J., Barrell D.J.A., Rattenbury M.S., Heron D.W., Haubrock S., Townsend D.B., Lee J.M., Berryman K.R., Nicol A., Cox S.C., and Stirling M.W. (2016). “The New Zealand Active Faults Database.” *New Zealand Journal of Geology and Geophysics* 59(1): 86-96.

Morton D.M. (1971). “Seismically triggered landslides in the area above the San Fernando Valley, in The San Fernando, California, Earthquake of February 9, 1971.” *U.S. Geological Survey* Professional Paper 733, p. 99-109

Murphy T. (2011). FEMA: freaking over a fake quake? *Mother Jones*. available at: < http://www.motherjones.com/politics/2011/03/fema-earthquake-new-madrid/>.

Nowicki Jessee M.A., Hamburger M.W., Allstadt K., Wald D.J., Robeson S.M., Tanyas H., Hearne M., Thompson E.M. (2018). “A Global Empirical Model for Near-Real-Time Assessment of Seismically Induced Landslides.” *J. Geophys. Res. Earth Surf.* 123(8):1835-1859.

Quigley M.C., Hughes M.W., Bradley B.A., van Ballegooy S., Reid C., Morgenroth J., Horton T., Duffy B. and Pettinga J. (2016). “The 2010-2011 Canterbury earthquake sequence: Environmental effects, seismic triggering thresholds, geologic legacy.” *Tectonophysics* 672-673: 228-274.

Rasanen R.A. and Maurer, B.W. (2021a). “Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence.” *Earthquake Spectra* (In Review).

Rasanen R.A. and Maurer, B.W. (2021b). “Probabilistic Seismic Source Inversion from Regional Landslide Evidence.” *Landslides* (In Review).

Scherbaum F., Schmedes J. and Cotton F. (2004). “On the Conversion of Source-to-Site Distance Measures for Extended Earthquake Source Models.” *BSSA* 94(3): 1053-1069.

Schmitt R.G., Tanyas H., Nowicki Jessee M.A., Zhu J., Biegel K.M., Allstadt K.E., Jibson R.W., Thompson E.M., van Westen C.J., Sato H.P., Wald D.J., Godt J.W., Gorum T., Xu C., Rathje E.M., Knudsen K.L. (2017). *An Open Repository of Earthquake-triggered Ground Failure Inventories.* USGS data release collection, accessed March 1, 2021, at https://doi.org/10.5066/F7H70DB4.

Stein S. (2008). *Presentation to Arkansas legislative committee on economic development on earthquake safety provisions in building code*. Little Rock, Arkansas; January 28, 2008. Available at: < http://www.earth.northwestern.edu/people/seth/research/testimony.html>.

Vidale J., Atkinson G., Green R., Hetland E., Grant-Ludwig L., Mazzotti S., Nishenko S. and Sykes L. (2011). *Report of the independent expert panel on New Madrid Seismic Zone earthquake hazards as approved by NEPEC on April 16, 2011*. USGS, 26 p.

APPENDIX B

Publication 2 (Journal Paper 1), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra*, *In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.
[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence

Ryan A. Rasanen[a] and Brett W. Maurer[b]

Abstract: In regions of infrequent moderate-to-large earthquakes, historic earthquake catalogs are often insufficient to provide inputs to seismic-hazard analyses (i.e., fault locations and magnitude-frequency relations) or to inform ground-motion predictions for certain seismic sources. In these regions, analysis of relic coseismic evidence, such as paleoliquefaction, is commonly used to infer information about the seismic hazard. However, while paleoliquefaction studies have been performed widely, all existing analysis techniques require *a priori* assumptions about the causative earthquake's location (i.e., rupture magnitude and ground motions can otherwise not be estimated). This may lead to inaccurate assumptions in some settings, and by corollary, erroneous results. Accordingly, this paper proposes an inversion framework to probabilistically constrain seismic-source parameters from paleoliquefaction. Analyzing evidence at regional scale leads to: (i) a geospatial likelihood surface that constrains the rupture location; and (ii) a probability distribution of the rupture magnitude, wherein source-location uncertainty is explicitly considered. Simulated paleoliquefaction studies are performed on earthquakes with known parameters. These examples demonstrate the framework's potential, even in cases of limited field evidence, as well as important caveats and lessons for forward use. The proposed framework has the potential to provide new insights in enigmatic seismic zones worldwide.

1. Introduction

In many seismic zones, the return period of moderate-to-large earthquakes is longer than the duration of recorded seismicity (e.g., the New Madrid Seismic Zone, South Carolina Coastal Plain, and Cascadia Subduction Zone, among others in the United States). As a result, historic earthquake catalogs may be insufficient to inform the inputs requisite for seismic-hazard analyses (i.e., the locations, magnitudes, and frequencies of fault ruptures). The uncertainties of these parameters are more than an academic curiosity, given that they influence the computed seismic hazard (e.g., the US National Seismic Hazard Map) (Vidale et al., 2011), which is adopted by building codes and used by engineers and policy makers. If the uncertainty is biased toward greater-than-actual hazards, capital is spent superfluously on seismic design, construction, and rehabilitation of infrastructure, hindering the economy (Stein, 2008; Murphy, 2011). Conversely, if seismic hazards are underestimated, lives,

[a] Graduate Research Assistant, University of Washington, Seattle; rrasanen@uw.edu
[b] Assistant Professor, University of Washington, Seattle; bwmaurer@uw.edu

infrastructure, and property are at risk. For example, a M_w7.7 New Madrid earthquake is projected to extensively damage Memphis and St. Louis, resulting in 80,000 injuries and fatalities, and economic losses of $300 billion - triple the cost of the 2005 Hurricane Katrina (Elnashai et al., 2009).

In addition to being affected by this uncertainty, many seismic zones have another commonality: pre-historic and/or pre-instrumental earthquakes induced soil liquefaction (or "paleoliquefaction" in such cases) that was later used to infer information about the causative ruptures. As an example, paleoliquefaction has been studied in ten regions of the Central-Eastern United States, as shown in Fig. 1, with the results used to inform seismic hazard analyses for nuclear facilities (Nuclear Regulatory Commission, 2012). Unsurprisingly, relic liquefaction also influences the US National Seismic Hazard Maps (Petersen et al., 2008; 2014). However, paleoliquefaction research to-date has focused more on field investigations, with less effort focused on advancing the techniques for inverse-analyzing field evidence. Both components are important, and the focus of this study is on the latter.

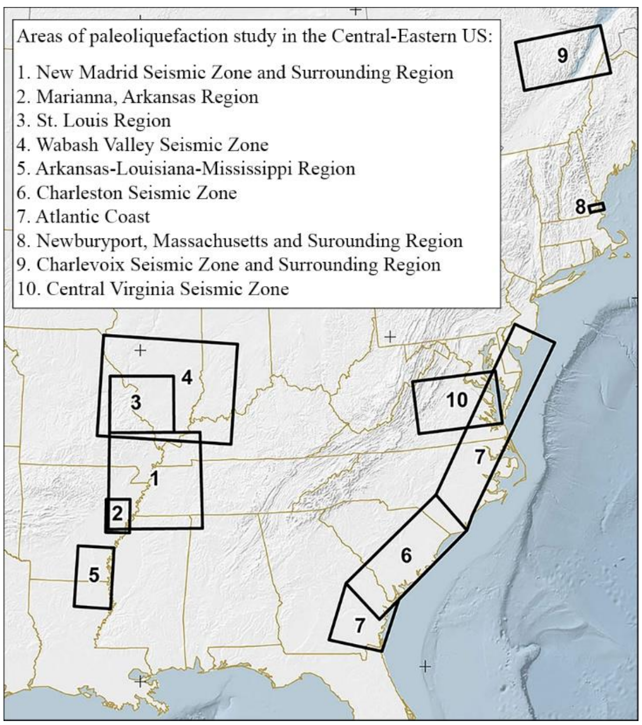


Fig. 1. Locations of regional paleoliquefaction datasets in the Central-Eastern US Seismic Source Characterization Project, which were used to develop inputs to probabilistic seismic hazard analyses for nuclear facilities in the region (adapted from Tuttle and Hartleb, 2012).

Notably, existing paleoliquefaction analysis techniques are underpinned by the need to locate the earthquake that induced the observed liquefaction. This can lead to inaccurate assumptions in some settings (e.g., where faults are blind or historically quiescent), and potentially, erroneous characterizations of the seismic hazard. Moreover, existing techniques do not explicitly consider the influence of source-location uncertainty, which can be significant, on the uncertainty of magnitude estimates derived from liquefaction. Accordingly, this study proposes and tests a framework for constraining seismic source parameters from relic liquefaction in an objective, fully probabilistic manner – a capability previously unavailable. Analyzing regional evidence, the likelihood of a rupture with given location, geometry, and magnitude producing a set of field observations is computed, wherein various measurement and modelling uncertainties may be accounted for. Repeating for a near-infinite number of possibilities results in a regional-scale uncertainty quantification of the source location and magnitude, with the degree of successful constraint dependent on factors to be investigated via tests on earthquakes with known parameters. That is, simulated paleoliquefaction studies will be performed to assess the framework's efficacy. In the following, existing paleoliquefaction analytics are briefly summarized, to include discussion of the need to geolocate an earthquake's source prior to analysis. An overview of the framework proposed herein is then given, followed by specific implementation details. Lastly, tests of the framework on eight modern earthquakes are presented, from which important lessons and caveats for forward use are gleaned.

1.1 Existing Paleoliquefaction Analytics

Paleoliquefaction studies can provide information about the dates, locations, and magnitudes of paleoearthquakes, as well as the resultant, regionally distributed ground-motion intensities. A complete study of paleoliquefaction is three phased (e.g., Tuttle 2001): (i) field identification and interpretation; (ii) dating; and (iii) constraint of the earthquake magnitude and/or ground motion under which it formed. The reader is referred to the overviews of field interpretation by Obermeier et al. (2001; 2005), and to the studies of Sims (1973), Obermeier and Dickenson (2000), Cox et al. (2007), Tuttle et al. (2002; 2005; 2019b), Tuttle and Atkinson (2010), Giona Bucci et al. (2018a; 2018b), and Villamor et al. (2016) for specific case studies. In addition, Sims and Garvin (1995), Quigley et al. (2013), Bastin et al. (2016), Giona Bucci et al. (2017), Tuttle et al. (2017), and Maurer et al. (2019) discuss field interpretation specific to spatiotemporally clustered earthquakes. Once identified, paleoliquefaction features may be dated via radiocarbon, optically stimulated luminescence, archeological context, stratigraphic context, or soil development indicators, such as weathering and biologic activity, among others. A comprehensive overview of these methods is provided by Tuttle et al. (2019a).

The techniques by which earthquake magnitude and/or shaking intensity are quantitatively constrained are generally called back- or inverse-analysis methods. While several such methods have been used, the two most common and credible to-date are the "magnitude-bound" method (e.g., Kuribayashi and Tatsuoka, 1975; Ambraseys, 1988; Olson et al., 2005a; Maurer et al., 2015) and the "site-specific geotechnical analysis," or for brevity, the "site-specific" method (e.g., Obermeier and Dickenson, 2000; Olson et al., 2005b; Rodriguez-Marek and Ciani, 2008; Green et al., 2005, 2014). These methods, and their limitations and uncertainties, have been discussed by many authors (e.g., Tuttle, 2001; Green et al., 2005; Olson et al., 2005; Tuttle et al., 2019a).

The magnitude-bound method uses a correlation relating earthquake magnitude to the site-to-source distance of the most distal site of liquefaction (typically using Joyner-Boore or epicentral distance). Developed from observations in modern earthquakes, these correlations traditionally use empirical data from variable geologic-tectonic settings. As an example, Fig. 2 presents several such correlations. The position of these curves is inherently a function of seismic source traits (e.g., rupture mechanism), transmission characteristics (e.g., ground motion attenuation and site effects), and liquefaction susceptibility (e.g., relative density and saturation). Because these factors vary spatially, region-specific correlations can provide more accurate estimates than those developed from global data (Obermeier et al., 2001; Olson et al., 2005a, 2005b; Maurer et al., 2015a). As evident from Fig 2., application of this method results in constraint of a paleoearthquake's magnitude, wherein the maximum site-to-source distance of liquefaction is the only input. Knowing or estimating the source location is thus critical, as is the most distal feature, with each generally being uncertain.

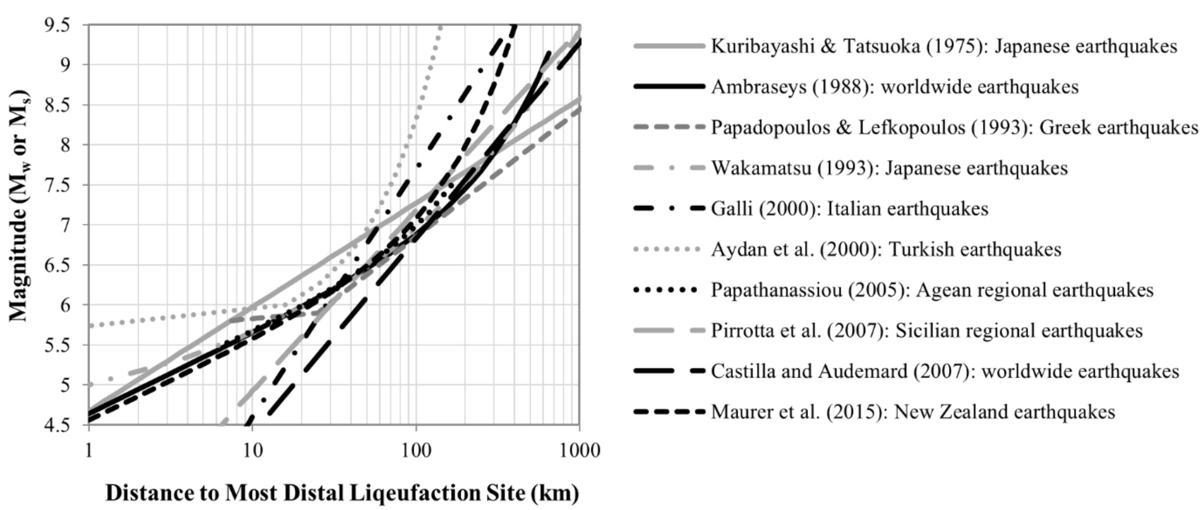


Fig. 2. Magnitude-bound curves for various geographic and tectonic settings, where site-to-source distance is quantified in terms of epicentral distance.

The second, more technical site-specific method uses liquefaction prediction models based on in-situ test measurements. In conventional hazard analyses, wherein the seismic loading is given, these models (e.g., Boulanger and Idriss, 2014; Green et al., 2019) predict future liquefaction. In inverse analyses, wherein the outcome is given (i.e., liquefaction did or did not occur), the models are used in reverse to assess the seismic loading that would, and would not, produce the observed outcome. While implementations have varied with time and place (e.g., Martin and Clough, 1990; Hayati and Andrus, 2008; Holzer et al., 2015), the site-specific method can: (i) constrain combinations of earthquake magnitude and ground-motion intensity likely to produce outcomes at individual sites; and (ii) by analyzing regional evidence, give an estimate of the causative earthquake magnitude.

Recent iterations of this method (e.g., Olson et al., 2005b) have utilized the so-called "stress-based" model for liquefaction triggering (Seed and Idriss, 1971), which was first developed for standard-penetration-test data. The model has had many updates, with variants developed for the cone-penetration-test (CPT), among others. Regardless of the variant, the factor of safety against liquefaction (FS_{liq}) is computed as the ratio of the cyclic resistance of the soil (CRR) to the earthquake induced cyclic stress ($CSR_{M=7.5,\sigma'=1atm}$). The CSR subscripts indicate it is computed for an M_w7.5 earthquake and vertical effective stress (σ'_{vo}) of 1 atm, reflecting the influence of shaking duration and overburden stress. To inverse-analyze paleoliquefaction, the stratum with depth-thickness-density combination consistent with the observed liquefaction response may be assumed to have FS_{liq} = 1.0:

$$FS_{liq} = \frac{CRR}{CSR_{M=7.5,\,\sigma'_v=1\,atm}} = 1.0 \tag{1}$$

By expanding the general form of the stress-based liquefaction triggering model (in this case, for CPT data), the minimum peak ground acceleration (PGA) required to induce liquefaction is:

$$PGA = CRR(q_{c1Ncs}) \times MSF(M_w) \times K_\sigma \times \frac{g \times \sigma'_{vo}}{0.65 \times \sigma_v \times r_d} \tag{2}$$

where PGA = peak horizontal ground acceleration; q_{c1Ncs} = CPT resistance, normalized for overburden pressure and adjusted for fines content; MSF = magnitude scaling factor to adjust for shaking duration; M_w = moment magnitude; K_σ = dimensionless factor incorporating the effect of overburden pressure; g = acceleration due to gravity; σ'_{vo}= vertical effective stress; σ_v = vertical total stress; and r_d = dimensionless factor accounting for the non-rigid response of the soil column, where the exact form of many variables depends on which version of the model is used (e.g., Green et al., 2019 versus Boulanger and Idriss, 2014). As shown in Fig. 3a, Eq. (2) separates combinations of PGA and M_w sufficient to induce liquefaction from combinations that are insufficient. As there are infinitely many such combinations, a ground-motion model (GMM) is used to determine credible combinations for a given site, as shown in Fig. 3b (dashed line), where the GMM predicts PGA as a function of M_w and site-to-source distance (R), among other factors. As indicated in Fig. 3b, the portion of this line plotting above the boundary defined by Eq. (2) corresponds to credible PGA-M_w combinations that could induce liquefaction at the site. The intersection of the GMM with the boundary-line (i.e., $FS_{liq} = 1$) thus defines the lower-bound PGA-M_w combination. In other words, the minimum M_w expected to induce liquefaction. Repeating the procedure at sites that did not liquefy results in a maximum M_w (i.e., if M_w had been any greater, liquefaction would have been induced). Finally, to obtain a best-estimate of the paleoearthquake's M_w, a regional assessment of PGA-M_w constraints is performed by coalescing analyses from sites with and without paleoliquefaction.

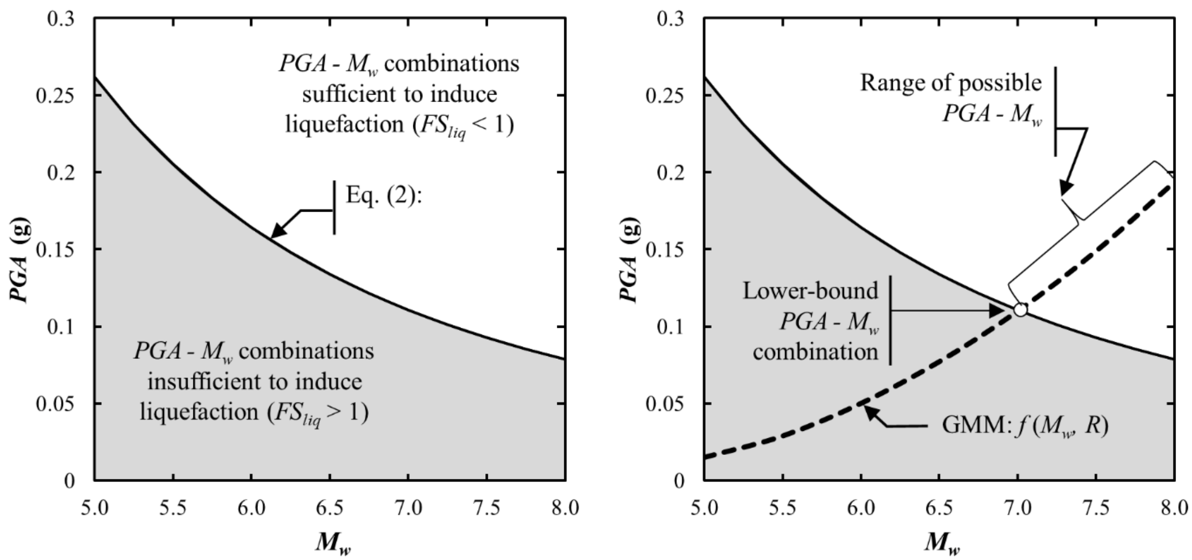


Fig. 3. (a) PGA – M_w combinations required to induce liquefaction for a hypothetical site; (b) Determination of credible lower-bound PGA – M_w combination (adapted from Green et al., 2005).

As evident from the preceding outline, site-to-source distance must be known to perform the site-specific method, as it presently exists. By corollary, and like the magnitude-bound method, the source location must be known to constrain seismic parameters. In addition, while proposed variants of these inverse-analysis methods incorporate uncertainty (Olson et al., 2013; Maurer et al., 2015a; Green et al., 2017), these variants still require the seismic source to be located prior to analysis. The accurate geolocation of a paleoearthquake's source thus underpins all existing methods, which can lead to inaccurate assumptions (e.g., where fault locations are unknown, or where prospective causative faults are plentiful). In such cases, the "energy center", or centroid of strongest shaking (similar to the epicenter) has been assumed to be the center of either the liquefaction field, or the region of most severe liquefaction (Obermeier et al., 2001; Green et al., 2005). However, this assumption is credible only if the subsurface is uniform across the region, which is rarely true. The depth, thickness, density, and susceptibility of deposits vary spatially, as do site response effects and the depth of ground water. As a result, the most severe liquefaction is often induced far from the earthquake's source. For example, in the 2010 Darfield, New Zealand, earthquake, the epicentral region was largely free of liquefaction, but widespread and severe liquefaction was observed more than 40 km away (Maurer et al., 2014a). The potential difficulty of accurate interpretation (i.e., if this were a hypothetical paleoseismic study) is depicted in Fig. 4, wherein select sites with and without liquefaction in the Darfield earthquake are mapped, along with several potential causative faults (among infinitely many, when considering unknown blind faults). Using existing techniques, the source might be assumed to reside at the center of the liquefaction field, rather than near the actual source (i.e., the Greendale fault) over 30 km to the west. Given this error, existing paleoliquefaction analytics could lead some to greatly miscalculate the Darfield earthquake's location and magnitude (Maurer et al., 2014b).

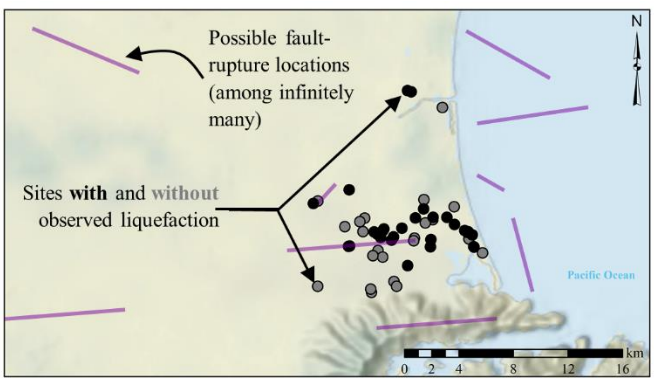


Fig. 4. Select sites with and without liquefaction (black and gray circles, respectively) during the 2010 Darfield, New Zealand, earthquake and potential causative faults. Can you guess the rupture location? Guessing incorrectly will significantly alter inferred earthquake parameters. (Ans: the fault most SW).

2. Methodology

The proposed framework is described in two parts. The first presents a conceptual overview and simple hypothetical application. The second provides implementation details to assist analysts in applying the framework, and to describe the validation tests that will be performed subsequently.

2.1 Overview of Proposed Framework

Analyzing regional paleoliquefaction evidence, the proposed framework aims to probabilistically constrain the causative rupture location and magnitude from a near-infinite number of possibilities. This will be accomplished by assessing the likelihood that a rupture with given location, geometry, and magnitude would produce a series of field observations (sites with and without liquefaction), wherein uncertainties inherent to ground motion- and liquefaction modelling are considered. In general, the likelihood of a parameter having a value, given a set of observations, is the product of the probabilities of those observations, conditioned on the parameter value. Thus, the likelihood of a rupture having some location (L), geometry (G), and magnitude (M_w), given a set (*x*) of field observations at *N* different sites, can be computed as:

$$Likelihood\ (L, G, M_w|x) = P(X = x|L, G, M_w)\ =\ \prod_{i=1}^{N}\ P(X_i = x_i|L, G, M_w) \quad (3)$$

Where $P(X_i = x_i|L, G, M_w)$ is the probability of the observation at site *i* (liquefaction or no liquefaction) given an earthquake with parameters *L*, *G*, and M_w. By repeating for a near-infinite number of possibilities, the actual rupture parameters are probabilistically constrained via the likelihood function (product of the probabilities of *N* observations), such that different combinations of *L*, *G*, and M_w will be found more and less likely to produce the observed field evidence. In Eq. (3), the probability of a field observation, conditioned on all input variables, is computed at sites with and without manifestations of liquefaction using Eqs. 4a and 4b, respectively:

$$P(\text{Liquefaction}|\text{EQK}: L, G, M_w) = \int_{PGA} P(\text{Liquefaction}|PGA, M_w)\, f(PGA|L, G, M_w) \cdot \text{d}PGA \quad (4a)$$

$$P(\text{No Liquefaction}|\text{EQK}: L, G, M_w) = 1 - \int_{PGA} P(\text{Liquefaction}|PGA, M_w)\, f(PGA|L, G, M_w) \cdot \text{d}PGA \quad (4b)$$

Where $f(PGA|L, G, M_w)$ is a probability density function (PDF) computed by a ground motion model (GMM) considering site response, and conditioned on fault parameters *L*, *G* and M_w; and $P(\text{Liquefaction}|PGA, M_w)$ is the probability of observing liquefaction at a site, given ground-motion parameters *PGA* and M_w, as computed by a model of liquefaction triggering or, more ideally, liquefaction surface manifestation (embedded in which is a triggering model). As an example, Geyin and Maurer (2020a) proposed fragility functions conditioned on three different liquefaction manifestation models (e.g., the liquefaction potential index, *LPI*, proposed by Iwasaki et al., 1978) computed using six different liquefaction triggering models (e.g., Green et al., 2019) such that users can select from, or average, 18 total functions. Expanding on Eqs. 4a and 4b, other uncertainties (i.e., other than *PGA*) could similarly be incorporated into the analysis (e.g., measurement uncertainties).

The framework's application is demonstrated conceptually in Fig. 5 considering four hypothetical field sites, where liquefaction is observed at sites one and two, but not at sites three and four, and two possible source locations for the earthquake that produced the observations. In actual analyses, a near-infinite number of source locations may be considered. Fig. 6 illustrates how the relative likelihoods of locations one and two are assessed. Shown in Fig. 6a are the probabilities of individual field observations, given an earthquake at location one, as computed by Eq. 4 for varying M_w. At sites with liquefaction, the probability of the field observation increases with increasing M_w, whereas at sites without liquefaction, the opposite occurs. In Fig. 6b, this process is repeated considering an earthquake at source location two. In Fig. 6c, the likelihood of each source location is computed as a function of M_w, per Eq. 3 (the product of the black and gray curves). It can be seen that location one has a far greater peak likelihood, whereas an earthquake at location two is unlikely to produce the field observations, regardless of its M_w. Repeating this process at a near-infinite number of locations, a geospatial surface of peak likelihood can be developed, potentially constraining the location of fault-rupture. Moreover, for any location considered, a likelihood distribution of earthquake magnitude is computed, as shown in Fig. 6c. These outcomes (i.e., probability distributions of earthquake location and magnitude) are direct inputs to probabilistic seismic hazard analyses.

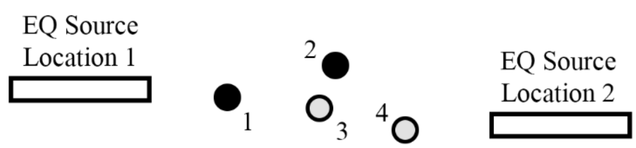


Fig. 5. Hypothetical paleoliquefaction analysis consisting of four field sites, where liquefaction was observed at sites one and two (black circles), but not at sites three and four (gray circles). Also shown are two possible sources for the earthquake that produced these observations.

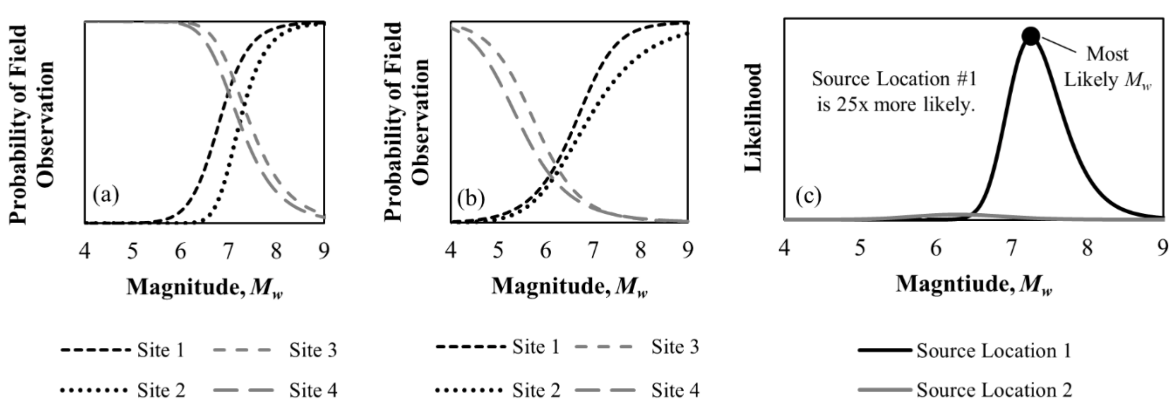


Fig. 6. Conceptual illustration of the proposed framework for locating the source of the earthquake that produced the field observations depicted in Fig. 5: (a) probabilities of field observations, given an earthquake at location one of varying M_w; (b) probabilities of field observations, given an earthquake at location two of varying M_w; (c) likelihood vs. M_w for source locations one and two.

2.2 Implementation Details

The implementation of Eqs. 3 and 4 is next discussed in detail. This includes general concepts, as well as the specific methods and models adopted herein for application to eight modern earthquakes. While the details of these events will be provided later, the framework will be applied in China, Japan, Taiwan, New Zealand, and the United States, covering a variety of seismologic and geologic settings. Owing to its modular form, the framework is easily modified for alternative models once developed. To implement the framework, an analyst must:

(1) Create an array of possible source locations. In this paper, a 250 km by 250 km grid of surficial points, which will be treated as possible earthquake epicenters, was adopted and centered on the liquefaction field. Within the grid area (62,500 km^2 in this study), a finer point spacing gives greater spatial resolution while a coarser spacing allows for faster runtime. Balancing these ideals, a l km spacing was adopted for a 10,000 km^2 area at the center of the grid, while a 10 km spacing was adopted for the remainder of the 62,500 km^2 study area. In the event that points in the coarsely spaced region are identified as likely sources, the grid can be recast at higher spatial resolution (i.e., to further constrain the likely source location, if possible).

(2) Select N study sites where: (i) the presence or absence of liquefaction manifestation was observed; and (ii) in-situ geotechnical testing was performed. Ideally, and as implemented in this work, an equal number of sites with and without manifestations is selected to limit sampling bias. While a larger N is generally advantageous, its influence on the accuracy of results will be investigated subsequently. In this paper, sites were characterized by the CPT, and as such, CPT-based liquefaction models will be utilized. Other tests for which liquefaction models are available can also be used.

(3) Select a regionally appropriate GMM. In this paper, the Bradley (2013) GMM was adopted for events in New Zealand while the Chiou and Youngs (2014) GMM was adopted for all others.

(4) For each provisional seismic-source locale created in (1):

(5) For each provisional seismic-source M_w being considered (a range of M_w5 to M_w9 was used in this paper, in increments of 0.1M_w):

(6) For each of N study sites selected in (2), cycling from i = 1 to N:

(7) Compute the site-to-source distance(s) required by the GMM chosen in (3), as measured from study site i to the provisional source locations. In this paper, the required metrics include Joyner-Boore distance (R_{JB}) and the closest distance to fault rupture (R_{RUP}). Since the seismic sources generated in (1) are epicenters, rather than 3-dimensional faults, the correlations of Scherbaum et al. (2004) were used to estimate R_{JB} from epicentral distance (R_{EPI}), assuming a strike-slip fault with dip (δ) of 90°.

The Scherbaum et al. (2004) correlations, which are magnitude dependent, effectively convert the point sources from (1) into a statistical realization of a 3-dimensional fault. R_{RUP} was calculated per the Pythagorean theorem using estimates of R_{JB} (computed as in the above) and the depth to top-of-rupture (Z_{TOR}). Z_{TOR} was estimated using the correlation of Kaklamanos et al. (2011), which requires as inputs the: down-dip rupture width (W); hypocentral depth (Z_{HYP}); and δ. Following the recommendation of Kaklamanos et al. (2011), W and Z_{HYP} were respectively estimated per the methods of Wells and Coppersmith (1994) and Scherbaum et al. (2004). It should be noted that the Scherbaum et al. (2004) R_{JB} correlation was used beyond the $5.0 < M_w < 7.5$ parameter space of the data used to train it. As a result, some adjustments to physically indefensible values were needed at very small and large M_w: (i) if $R_{JB} > R_{EPI}$, $R_{JB} = R_{EPI}$; (ii) if $R_{JB} < 0$, $R_{JB} = 0$.

(8) Using the GMM selected in (3) and site-to-source distances from (7), compute the PDF of expected *PGA* at study site *i*, wherein the time-averaged shear wave velocity over the top 30 m (V_{S30}) is consistent with site *i*. In this paper, the fault source was assumed to have strike-slip parameters (δ = 90° and rake, λ = 0°) (assuming the actual parameters are unknown in a paleoseismic study) and V_{S30} was estimated by the method of Heath et al. (2020). In general, the probability distribution of *PGA* is assumed to be a lognormal random variable defined by the median and lognormal standard deviation, which are provided by a GMM. In this paper, *PGA* predictions beyond +/- 3 standard deviations of the median are truncated, as is typical in ground-motion modeling (since such values may be physically impossible) and the PDF is scaled such that the area underneath it is one.

(9) For each possible *PGA* value at study site *i*, as computed in (8) for a given M_w and site-to-source distance pair (a *PGA* increment of 0.05 g was adopted in this paper):

(10) Compute the safety factor against liquefaction, FS_{liq}, throughout the profile using a liquefaction triggering model, the inputs to which include *PGA* and M_w. In this paper, the Boulanger and Idriss (2014) CPT-based model was adopted. However, because liquefaction may trigger at depth but be inconsequential, or unobserved, at the ground surface, it is generally more appropriate to predict surficial manifestations of liquefaction (i.e., liquefaction vents and dikes at or near the surface), since such evidence is generally observed in the field, rather than evidence at depths well below the ground surface at the time of liquefaction. While several manifestation models are available, *LPI* was adopted in this paper and is defined as (Iwasaki et al., 1978):

$$LPI = \int_0^{20\,m} F(FS_{liq}) \times w(z)dz \tag{5}$$

Where $F(FS_{liq})$ and $w(z)$ are functions that weight the respective influences of FS_{liq} and depth, z, on surface manifestation. Specifically, $F = 1 - FS_{liq}$ for $FS_{liq} \leq 1$ and $F = 0$ otherwise; $w(z) = 10 - 0.5z$, where z is the depth (m) below the ground surface. *LPI* thus assumes that manifestation depends on the thickness of all liquefied strata in the upper 20 m, their proximity to the surface, and the amount by which FS_{liq} in each stratum is less than 1.0. Given this definition, *LPI* can range from 0 for sites with no expected liquefaction to 100 for sites with FS_{liq} = 0 over the entire 20-m depth.

(11) Following from (10), compute the probability of liquefaction manifestation using a function conditioned on the adopted manifestation model (Geyin and Maurer (2020a) provide a variety of functions), then multiply by the probability of the given *PGA* value occurring, as shown in Eq. 4. In this paper, the fragility function of Geyin and Maurer (2020a), which relates the probability of surficial liquefaction manifestation to *LPI*, was adopted and is defined as:

$$F_{MS}(LPI) = \Phi\left(\frac{ln(LPI) - \ln(\theta)}{\beta}\right) \tag{6}$$

Where $F_{MS}(LPI)$ is the probability of surface manifestation, Φ is the Gaussian cumulative distribution function, and θ and β are the distribution's median and logarithmic standard deviation, respectively. Geyin and Maurer (2020a) separately trained functions using data from New Zealand and other global earthquakes. Accordingly, the New Zealand specific function (θ = 5.582, β = 1.537) was adopted for application to earthquakes in New Zealand, whereas the global function (θ = 7.508, β = 1.335) was adopted for all others.

(12) Compute the probability of field observation using Eq. 4a or 4b, depending on whether manifestations were or were not observed at study site *i*. Completing these equations (i.e., by summing over all *PGA* values) gives the probability of the field observation at site *i* for a given seismic source location and M_w. Repeating steps 6-12 for all considered M_w results in a probability of field observation versus M_w curve for each site, examples of which are in Figs. 6a and 6b.

(13) Compute the likelihood of a provisional source locale, as a function of M_w, by multiplying the probabilities of field observations at all study sites (i.e., multiply the curves shown in Fig. 6a or 6b at each value of M_w). The result, an example of which is shown in Fig. 6c, is a PDF of the causative earthquake magnitude, conditioned on a single seismic-source location.

(14) Repeating steps 5-13 for all provisional seismic source locations created in (1) results in a likelihood distribution for each, as described in (13).

(15) Following from (14), normalize the peak likelihood at each source location by the peak likelihood among all locations. This allows for the relative likelihoods of source locations to be assessed, with

the most likely location having a normalized value of one. Create geospatial surfaces or contours to identify this location. The M_w distribution here (see Fig. 6c) is in-turn the PDF of the inverted rupture M_w, conditioned on the most likely source location.

(16) Finally, by aggregating PDFs from all potential source locations across the study area, an overall PDF of M_w, considering all possible source locations, is produced. While a single source location will always be "most likely," earthquakes at multiple other source locations have potential to produce the field of evidence, albeit with lesser likelihood. In general, this overall PDF will be wider (i.e., have greater uncertainty) than the PDF conditioned on the single most likely source, especially when the source zone is not well constrained (e.g., due to limited field evidence).

3. Simulated Application

The proposed framework can potentially constrain both the location and M_w of a seismic source considering only liquefaction evidence, given that geotechnical data is available. To investigate the degree of successful constraint and glean lessons for forward use, simulated paleoliquefaction studies are next performed on eight earthquakes with known seismic sources. Inherent to these applications, it is assumed that the presence or absence of liquefaction would be accurately interpreted in a paleoliquefaction investigation, and moreover, that liquefaction effects would be properly attributed to their respective, causative earthquakes. These assumptions are not intended to diminish the importance or trivialize the difficulty of field interpretation; the objective herein is to advance inverse-analysis techniques, rather than to discuss or simulate the field work prerequisite for such analyses. As an example, it assumed that liquefaction did not manifest if none was observed. While absence of evidence is not necessarily evidence of absence, the need nonetheless exists, where prudent, to judge sites as "negative." Similarly, temporal changes over long geologic timescales may introduce challenges and uncertainties not present in the modern earthquakes on which the framework is demonstrated herein.

The framework will first be applied to three earthquakes in Canterbury, New Zealand. Here, a uniquely large number of liquefaction case histories provides the opportunity to investigate the framework's performance in the context of finite-sample uncertainty (i.e., the relationship between performance and quantity of field data). Five other global events will then be studied, wherein the quantity of field data varies widely. In all events, the inverted source location and magnitude will be compared to the actual, known parameters. The framework's geolocation error will be quantified as the distance from the identified, most likely point source to the surface projection of fault rupture (akin to R_{JB}). However, no prior knowledge of the seismic source model is incorporated into the analyses.

3.1 *New Zealand Events*

Earthquakes occurring in Canterbury, New Zealand, from 2010 to 2016 resulted in an unprecedented quantity of liquefaction case histories (e.g., Geyin et al. 2020b, 2021). Three of these events will be studied, including the: 2010 M_w7.1 Darfield; 2011 M_w6.3 Christchurch; and 2016 M_w5.7 Christchurch earthquakes. Notably, these events all occurred on previously unrecognized faults near a large city, highlighting the importance of paleoliquefaction studies. Approximately 5,000 case histories, or locations where liquefaction response was observed and CPT testing was performed, were compiled from each event by Geyin et al. (2021). Given that this far exceeds the number of field sites conceivable in a paleoliquefaction study, analyses will be run on random samples ranging from 10 to 400 total sites (50% with manifestation and 50% without). Multiple analyses of each sample size will be performed to quantify finite-sample uncertainty. While this uncertainty may differ across events, it nonetheless could provide insights that guide future paleoliquefaction studies (e.g., the number of study sites generally needed to reduce finite-sample uncertainty to desirable limits).

3.1.1 *2011 M_w6.3 Christchurch Earthquake, New Zealand*

The M_w6.3 Christchurch earthquake occurred on a blind, previously unknown fault beneath the southern environs of Christchurch, causing widespread damage and 185 fatalities (Quigley et al., 2016). A total of 4,846 CPT-based liquefaction case histories (3,942 sites with manifestation and 904 without) were obtained from the Geyin et al. (2021) compilation for analysis. An example of the framework's output is shown in Figs. 7a and 7b for an analysis using 100 randomly selected study sites; their locations, and the surface projection of the 2011 fault rupture, are also shown. Mapped in Fig. 7a are contours of back-calculated relative likelihood, wherein values near one and zero respectively denote source locations most and least likely to produce the observed evidence. Mapped in Fig. 7b are M_w contours, indicating the median M_w likely to produce the observed evidence, conditioned on an earthquake at that location occurring.

As seen in Fig. 7a, the analysis identified a relatively small, well-constrained area just south of the evidence field as the likely source. This aligns with the actual rupture, such that the identified, most likely source falls within the surface projection of the faults that ruptured (i.e., geolocation error = 0). In addition to these faults (unknown at the time of the event), other known faults in the region are mapped. Notable seismic sources east, north, and west of the evidence field (e.g., the Alpine, Kaiapoi, Porter's Pass, and Greendale Faults) are all unlikely to produce the field observations. In addition, as shown in Fig. 7b, the median magnitude corresponding to the most-likely source location was ~M_w6.4, very close to the actual M_w6.3. Because M_w contours tend to predictably map as concentric circles, as in Fig. 7b, these contours are provided in Appendix A for the remainder of events.

In addition to the median M_w, other M_w values at the most-likely source location could inevitably produce the evidence field (e.g., if ground motions were higher or lower than the median expectation). Shown in Fig. 8a is the complete cumulative distribution function (CDF) of M_w at the most likely source location, which is analogous to the conceptual distributions shown in Fig. 6c. The 95% confidence interval (CI) is M_w6.10-M_w6.65. Moreover, as mapped in Fig. 7a, earthquakes elsewhere have potential to produce the evidence field, albeit with lesser likelihood. By aggregating individual CDFs from all locations across the 62,500 km^2 study grid, an overall CDF that includes source-location uncertainty is produced. This CDF, which is also shown in Fig. 8a, has a median of M_w6.8 and 95% CI of M_w6.10-M_w7.90. The cause of this "skew" toward large M_w is readily apparent from Fig. 7a, wherein likelihood contours are asymmetric about the most-likely source, skewing to the south where larger M_w events at greater distances have some (albeit low) potential to produce the evidence field. Unless the causative fault can be constrained via on-fault or other geologic evidence, the latter, more uncertain CDF would be the more appropriate product of the analysis. While not guaranteed, it is generally the case that inverted source parameters are more accurate and less uncertain when conditioned on a single, most-likely source (assuming its location is accurate). In other words, constraint on source location should better constrain the causative rupture M_w. The likelihood contours mapped in Fig. 7a should therefore be considered in the context of other seismologic and geologic evidence (e.g., maps of known and potential seismic sources; interpretation of liquefaction sites lacking geotechnical testing; and consideration of other co-seismic evidence, if present). In this way, analysts might arrive at a "preferred" estimate of the source location and M_w by employing a more holistic approach.

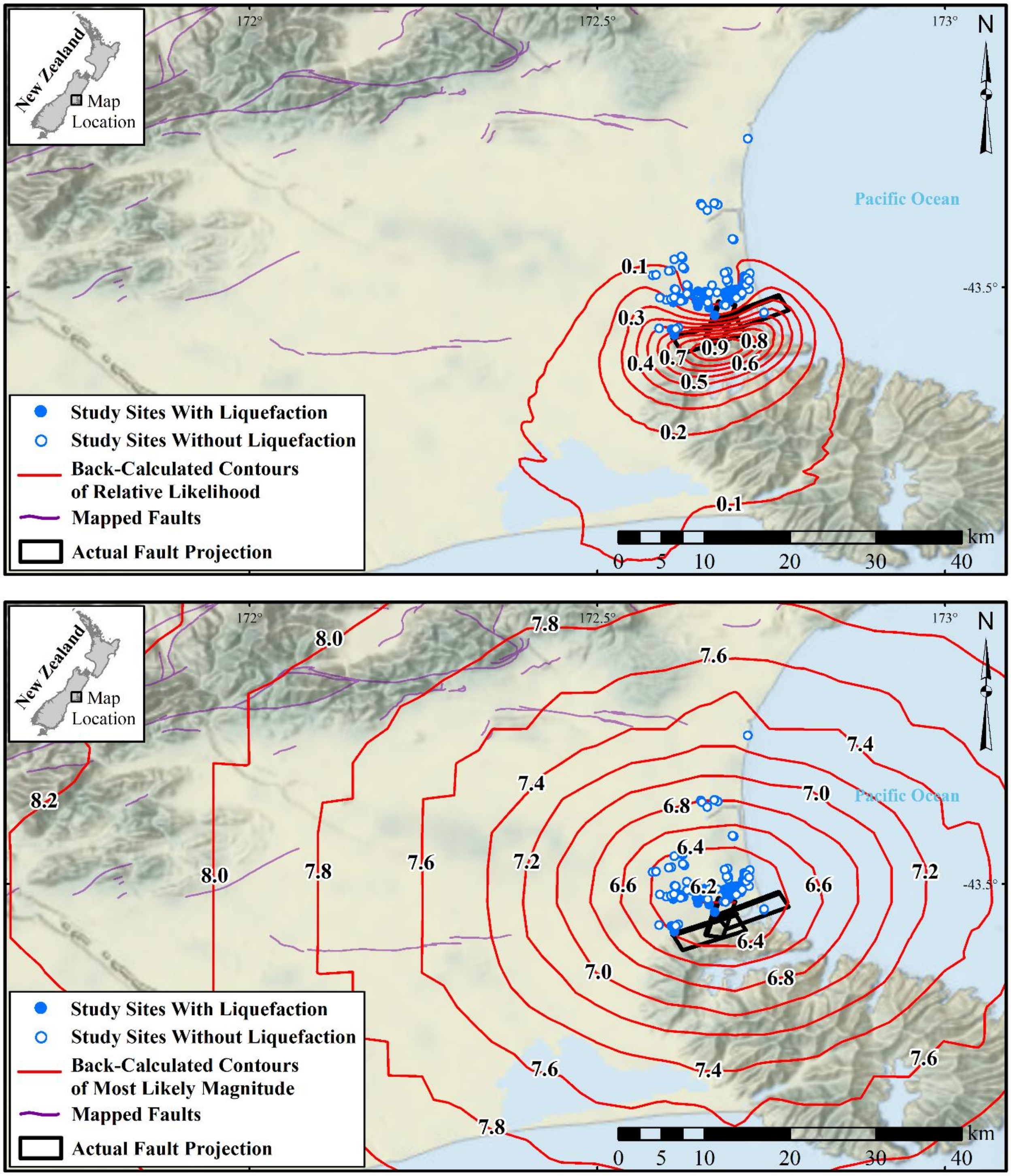


Fig. 7. Representative analysis using 100 study sites from the 2011 M_w6.3 Christchurch earthquake: (a) contours showing the relative likelihood of earthquake source locations; and (b) corresponding M_w most likely to produce the field evidence, if the earthquake occurred at a given location. Black rectangles delineate the actual fault projection (Beavan et al., 2012); faint purple lines are other mapped faults in the region (Langridge et al., 2016).

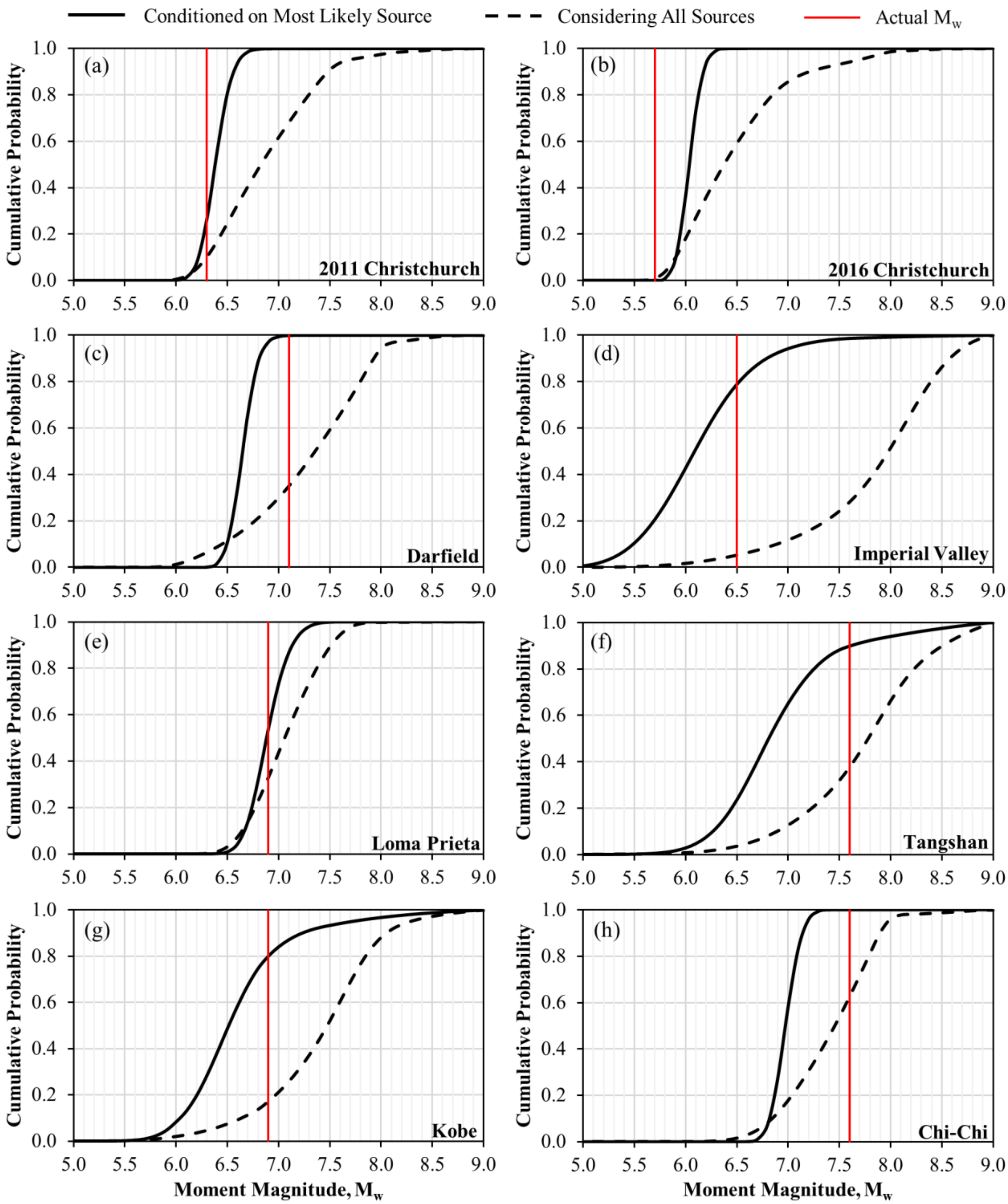


Fig. 8. CDFs of earthquake magnitude inverted from regional liquefaction evidence in the: (a) 2011 M_w6.3 Christchurch; (b) 2016 M_w5.7 Christchurch; (c) 2010 M_w7.1 Darfield; (d) 1979 M_w6.5 Imperial Valley; (e) 1989 M_w6.9 Loma Prieta; (f) 1976 M_w7.6 Tangshan; (g) 1995 M_w6.9 Kobe; and (h) 1999 M_w7.6 Chi-Chi earthquakes. Shown are the CDFs conditioned on the most-likely source location (solid lines) as well as the overall CDFs considering source-location uncertainty (dashed lines).

While the location and magnitude of the 2011 Christchurch earthquake were accurately inverted, the analysis in Fig. 7 utilized 100 sites among nearly 5000 available. To evaluate the finite-sample uncertainty of these results as a function of dataset size, analyses were performed on random samples of varying size. The results are shown in Fig. 9 for both products in Fig. 7 (i.e., geolocation error and median M_w, conditioned on the most likely location). While these results do improve slightly as the quantity of data grows, accurate results are possible with very few sites, albeit with large finite-sample uncertainty. With respect to geolocation, the 1σ finite-sample uncertainty declines from ~3 km studying 10 sites to ~2.15 km for 100 sites. With respect to the median M_w, the 1σ uncertainty declines from ~0.4M_w studying 10 sites to ~0.2M_w for 100 sites. While these trends may differ between events, they nonetheless provide insights for future paleoliquefaction studies and indicate that outcomes significantly different from those in Fig. 7 are relatively rare, regardless of which 100 sites are studied.

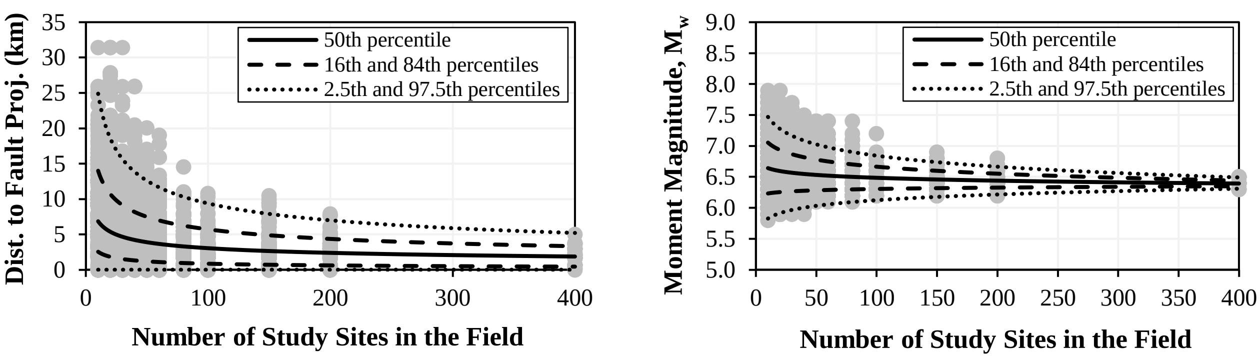


Fig. 9. Finite-sample uncertainty of the inverted (a) most likely source location; and (b) median M_w of the 2011 Christchurch earthquake as a function of the number study sites analyzed.

3.1.2 2016 Christchurch (Valentine's Day) Earthquake, New Zealand

The M_w5.7 Valentine's Day Earthquake occurred just offshore of Christchurch and induced liquefaction in the city's eastern environs (Kaiser et al., 2016). A total of 4,933 liquefaction case histories (155 sites with manifestations and 4,778 without) were obtained from the Geyin et al. (2021) compilation for analysis. A representative output is shown in Fig. 10 for an analysis using 100 random study sites; the locations of these sites, along with the surface projection of the 2016 fault rupture, are also shown. Mapped in Fig. 10 are contours denoting source locations most and least likely to produce the observed evidence. The analysis again identified a relatively well-constrained area as the likely source, this time just east of the evidence field. It can also be seen that the source with greatest likelihood is within the actual fault projection (geolocation error = 0). The median M_w here (See Appendix A for map) was M_w6.0 with a 95% CI of M_w5.80-M_w6.25, versus the actual M_w5.7. As with the 2011 earthquake, this CDF - conditioned on the most-likely source location - and the overall CDF including source-location uncertainty, are computed and shown in Fig. 8b. Like the 2011 event, the latter, overall CDF is more uncertain, but regardless of which CDF is consulted, a M_w5.7 event is found unlikely to produce the evidence field. This discrepancy, despite accurate geolocation, could be due to the point representation of faults. While the points are converted to a statistical fault-realization that assumes some orientation and depth, this realization may differ from the true source. As a result, the computed median M_w could differ from actual, even when the location is accurately predicted. Of course, the inversion framework also relies on accurate prediction of ground-motion attenuation, site response, liquefaction triggering, and liquefaction manifestation. Discrepancies are thus likely whenever these phenomena differ from expectations.

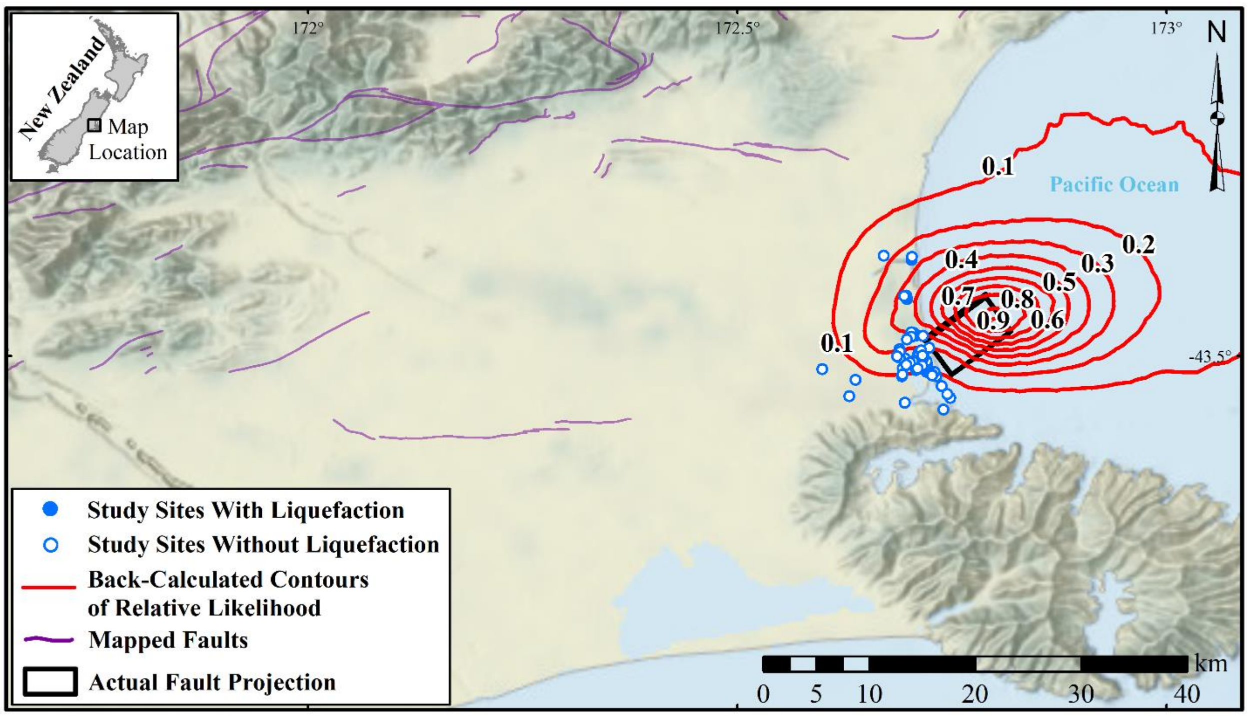


Fig. 10. Representative analysis of 100 study sites from the 2016 M_w5.7 Christchurch earthquake: contours showing the relative likelihood of source locations. Black rectangle = actual fault projection (Kaiser et al., 2016); faint purple lines = other mapped faults (Langridge et al., 2016).

To assess the finite-sample uncertainty of results, multiple analyses were again performed on samples of varying size, the results of which are shown in Fig. 11. Like the 2011 Christchurch earthquake, accurate results were possible with few sites, albeit with greater finite-sample uncertainty. With respect to geolocation, this 1σ uncertainty declines from ~4.1 km studying 10 sites to ~0.8 km for 100 sites. As for the median M_w, this 1σ uncertainty declines from ~0.2M_w studying 10 sites to

~0.1M_w for 100 sites. Thus, while exceptions do exist, most analyses using 100 study sites result in outcomes similar to those in Fig. 10.

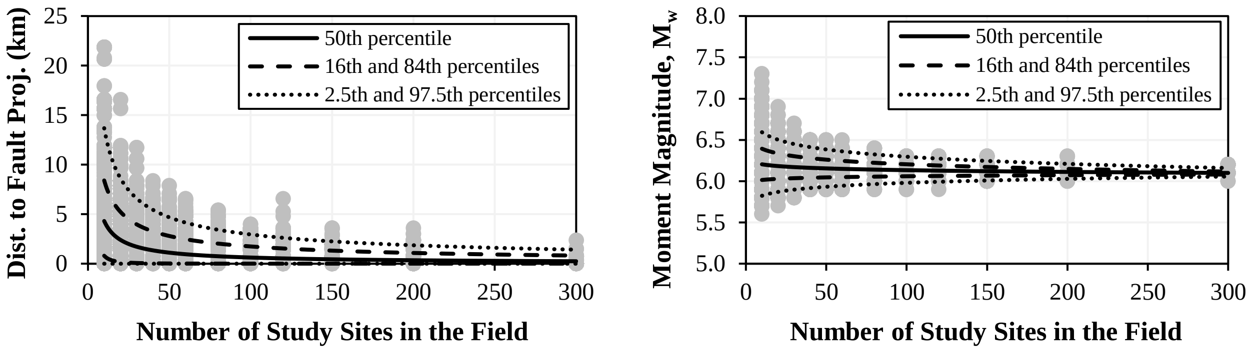


Fig. 11. Finite-sample uncertainty of the inverted (a) most likely source location; and (b) median M_w of the 2016 Christchurch earthquake as a function of the number study sites analyzed.

3.1.3 2010 Darfield Earthquake, New Zealand

The M_w7.1 Darfield earthquake occurred ~35 km west of Christchurch on the previously unknown Greendale Fault (Quigley et al., 2016). A total of 5,417 liquefaction case histories (1,133 sites with manifestations and 4,284 without) were obtained from Geyin et al. (2021). A representative output is shown in Fig. 12 for an analysis of 100 study sites; their locations, along with the surface projections of the faults that ruptured, are also shown. In contrast to the events discussed previously, the source is not well constrained (i.e., a large area has relatively high likelihood). In addition, while the analysis correctly predicted the general direction of the source relative to the evidence field, the identified, most likely source is ~20 km from the actual fault. The corresponding median M_w here (map in Appendix A; CDF in Fig. 8c) was M_w6.65 with a 95% CI of M_w6.4-M_w6.9, versus the actual M_w7.1. As in the above, multiple analyses were performed on samples of varying size, the results of which are in Fig. 13. Two observations from these analyses are: (i) very few analyses accurately predicted the location or magnitude, regardless of how many study sites were analyzed; and (ii) relative to the 2011 and 2016 events, the most likely location has large finite-sample uncertainty. That is, it varies widely depending on which data is analyzed. Studying 10 sites, for example, the 1σ uncertainty is ~12.3 km (versus ~4.5 km in the previous events). Returning to Fig. 8c, the overall M_w CDF considering source location uncertainty has a median of M_w7.35 and 95% CI of M_w6.1-M_w8.1. Thus, in this case, the overall CDF provides a more accurate characterization of the rupture magnitude, whereas as that conditioned on a single, most-likely source underpredicts the magnitude, owing its erroneous location.

While there are several possible causes of less accurate (and more uncertain) results, this may be attributable in the Darfield event to lower rates of ground-motion attenuation at greater site-to-source distances, given that none of the compiled study sites are in the epicentral area. Similar to the geolocation of epicenters using seismometers, which relies on seismic waves having different arrival times, the proposed inversion framework relies on waves having different arrival amplitudes. As amplitudes become more similar at two sites, it becomes more difficult to determine the direction from which motions are propagating, and thus more difficult to locate the seismic source. Plotted in Fig. 14, as a function of R_{RUP}, are *PGAs* recorded in the Darfield earthquake, compared against predictions by the Bradley (2013) GMM. It can be seen that recorded *PGA*s attenuated more rapidly at short site-to-source distances than in the far field, as predicted by the GMM. This results in *PGA*s being relatively uniform across the study sites, which are located at R_{RUP} distances of ~20-35 km. Thus, for events where the evidence field is far from the seismic source, or wherever differences in *PGA* are subtle among sites, the difficulty of analysis increases. Nonetheless, had the analysis been based on a single source at the centroid of liquefaction severity (as one might assume in a paleoliquefaction study), the median magnitude estimate would be ~M_w6.1 (Appendix A). In this regard, the proposed framework provides results that are useful even if imperfect.

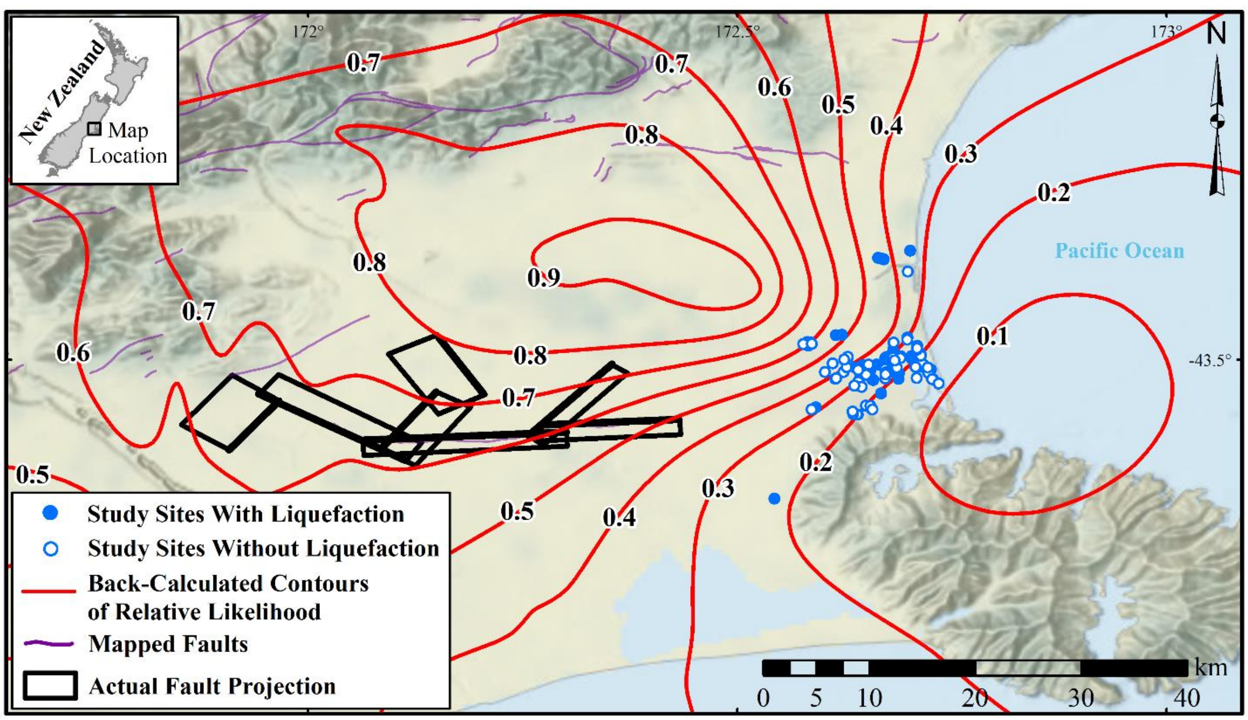


Fig. 12. Representative analysis of 100 study sites from the 2010 M_w7.1 Darfield earthquake: contours showing the relative likelihood of source locations. Black rectangles = actual fault projection (Beavan et al., 2012); faint purple lines = other mapped faults (Langridge et al., 2016).

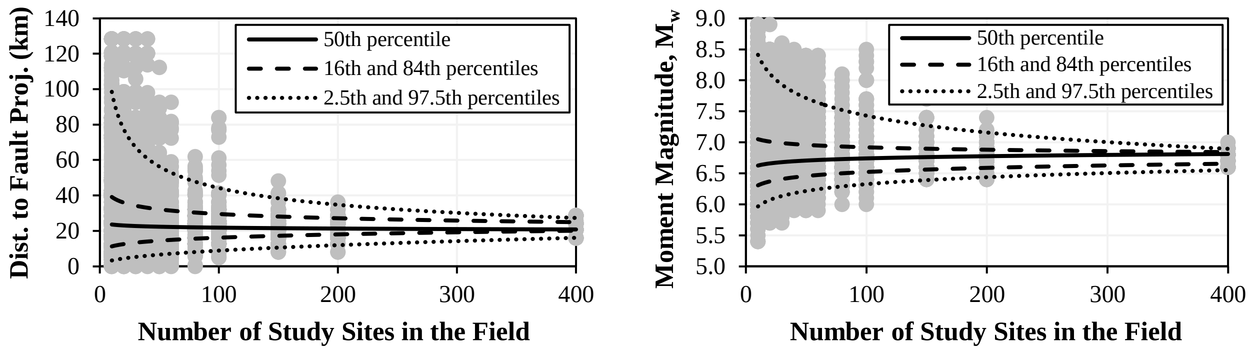


Fig. 13. Finite-sample uncertainty of the inverted (a) most likely source location; and (b) median M_w of the 2010 Darfield earthquake as a function of the number of study sites analyzed.

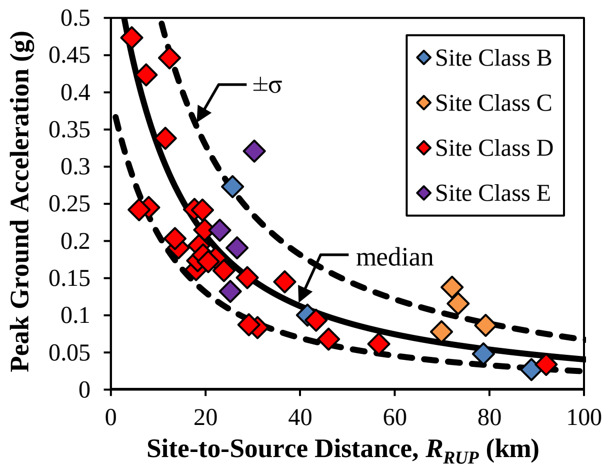


Fig. 14. Recorded *PGA*s from the 2010 Darfield earthquake, plotted as a function of R_{RUP} and color coded based on the station's site class, compared against the Bradley (2013) GMM for site class D.

3.2 Global Events

Liquefaction case-history data was obtained from Geyin et al. (2020b) for five additional earthquakes in four countries, wherein the quantity of data varies, but is far less than the New Zealand earthquakes (varying from 6 to 46 study sites). For these events, all available data was selected and a single analysis was performed using an equal number of sites with and without liquefaction manifestations.

3.2.1 1979 Imperial Valley, United States

The 1979 M_w6.5 Imperial Valley, California earthquake occurred on the Imperial and Brawley Faults (Archuleta, 1984). Six liquefaction case histories were obtained from Geyin et al. (2020b) and analyzed, the results of which are in Figs. 8c and 15. The inversion identified a relatively small but weakly constrained source area, with the most likely source falling near the fault projection (geolocation error = 2 km). The corresponding median magnitude here was M_w6.1 (mapped in

Appendix A) with a 95% CI of M_w5.2-M_w7.3, versus the actual M_w6.5. The Imperial Valley event thus demonstrates that seismic source parameters may be accurately predicted even when few study sites are available. However, as evident from Fig. 15, earthquakes located elsewhere throughout the study area have appreciable potential to produce the field observations. As a result, the overall CDF considering source-location uncertainty has a median of M_w8.0 and 95% CI of M_w6.1-M_w8.8. The potential benefit of coupling the results in Fig. 15 with additional geologic evidence and interpretation is apparent, given the large uncertainty that ensues when the source location is weakly constrained.

3.2.2 1989 Loma Prieta, United States

The 1989 M_w6.9 Loma Prieta, California earthquake occurred on the San Andreas Fault (Wald et al., 1991). Forty-six liquefaction case histories were obtained for analysis, the results of which are mapped in Fig. 16a. In this case, the identified, most likely source was ~30 km north of the actual rupture; moreover, the median magnitude here was M_w7.9 (Appendix A), an overestimate of M_w1.0. These mispredictions are attributable to unusual site-response effects in the San Francisco Bay Area, which are variable and difficult to predict without site-specific ground motion analyses. In particular, the northernmost study sites (e.g., in the environs of Alameda and Richmond) experienced *PGA*s far exceeding predictions by the Chiou and Youngs (2014) GMM (wherein Heath et al. (2020) V_{S30} values were used). That is, liquefaction was not expected at these sites given GMM predictions of the actual rupture, causing the inverted source to be larger in magnitude and more northward than actual. To confirm, a revised analysis simulating a more rigorous seismic site-characterization was performed. First, *PGA*s observed at each study site were estimated from USGS ShakeMap (Wald et al., 2005). Second, *PGA*s were predicted at each study site using the GMM, given known source parameters and rock conditions (V_{S30} = 700 m/s). Third, a site-specific amplification factor (AF), PGA_{Soil}/PGA_{Rock} was computed at each study site. The inversion framework was then reapplied using the site-specific AF in conjunction with GMM predictions on rock. By way of this approach, the site response component of ground-motion prediction was accounted for by the AF, rather than by empirical coefficients within the GMM. This reanalysis, shown in Fig. 16b, resulted in a geolocation error of 0 km. The median M_w at the most likely source location was M_w6.9 with 95% CI of M_w6.6-M_w7.3, as shown in Fig. 8e. The overall CDF considering source-location uncertainty resulted in a similarly accurate but slightly wider distribution, having a median of M_w7.05 and 95% CI of M_w6.5-M_w7.7. This demonstrates the potential benefit of performing site-specific ground motion analyses in paleoliquefaction studies (i.e., *Vs* profiling and site-response analysis), rather than relying solely on empirical GMMs.

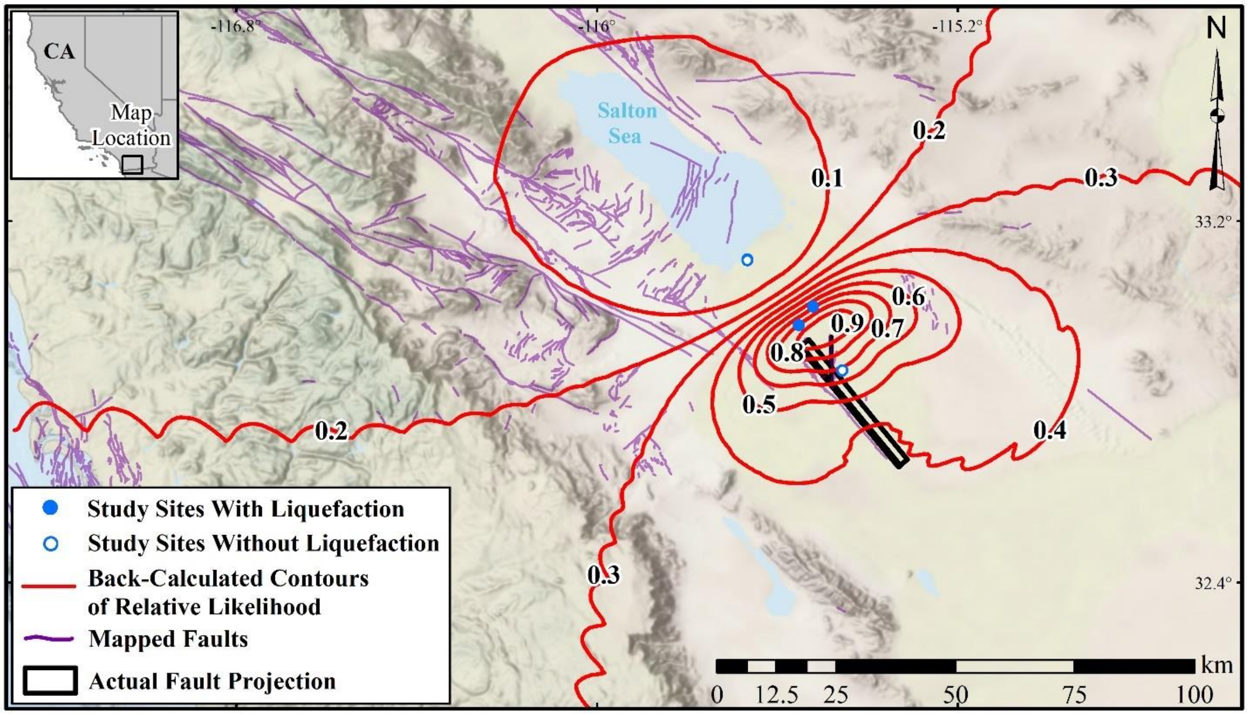


Fig. 15. Analysis of six study sites from the 1979 Imperial Valley, California earthquake: contours showing the relative likelihood of source locations. Black rectangles = actual fault projection (Archuleta, 1984); faint purple lines = other mapped faults (California Geological Survey, 2014).

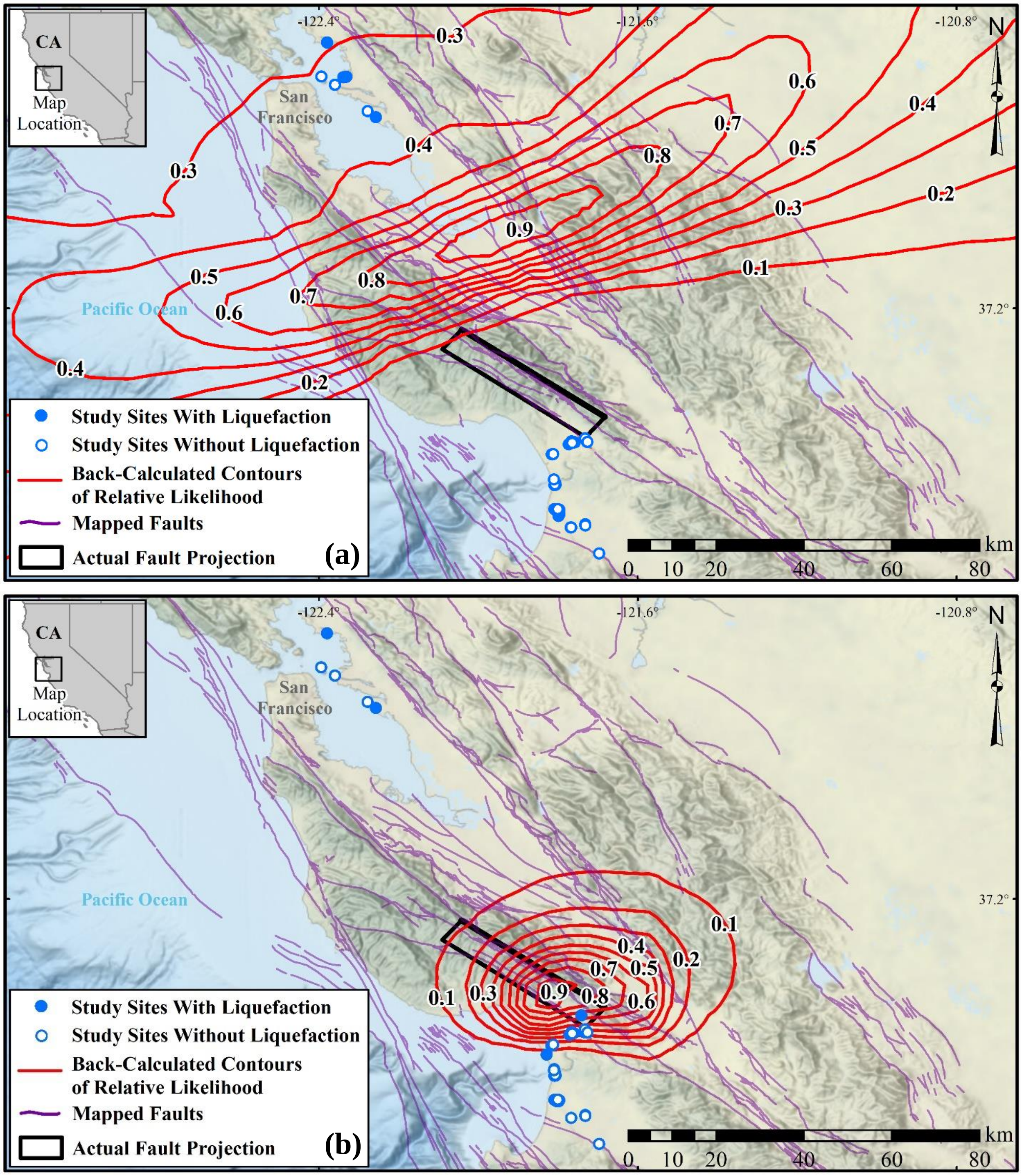


Fig. 16. Analysis of 46 study sites from the 1989 Loma Prieta, California earthquake: (a) contours showing the relative likelihood of source locations; (b) contours showing the relative likelihood of source locations, after inclusion of site-specific ground motion amplification factors, as discussed in the text. Black rectangles = actual fault projection (Wald et al., 1991); faint purple lines = other mapped faults (California Geological Survey, 2014).

3.2.3 1976 Tangshan, China

The 1976 M_w7.6 Tangshan, China earthquake occurred on the previously unknown two-segment Tangshan Fault (Huang and Yeh, 1997). Six liquefaction case histories were obtained and analyzed, the results of which are mapped in Fig. 17. The inversion identified a weakly constrained source area aligning with the actual fault projections (geolocation error = 0 km). As shown by the CDF in Fig. 8f, the median M_w here was M_w6.8 with 95% CI of M_w6.0-M_w8.5. Like the Imperial Valley earthquake, the Tangshan event demonstrates that a small number of study sites is likely to constrain the source in a manner that is weak, even if accurate. Considering source-location uncertainty (CDF in Fig. 8f), the median M_w was M_w7.8 with 95% CI of M_w6.4-M_w8.8, versus the actual M_w7.6. This large uncertainty should not be interpreted as a failing of the methodology, but rather, as an objective quantification of the actual uncertainty, given the limited quantity of field data.

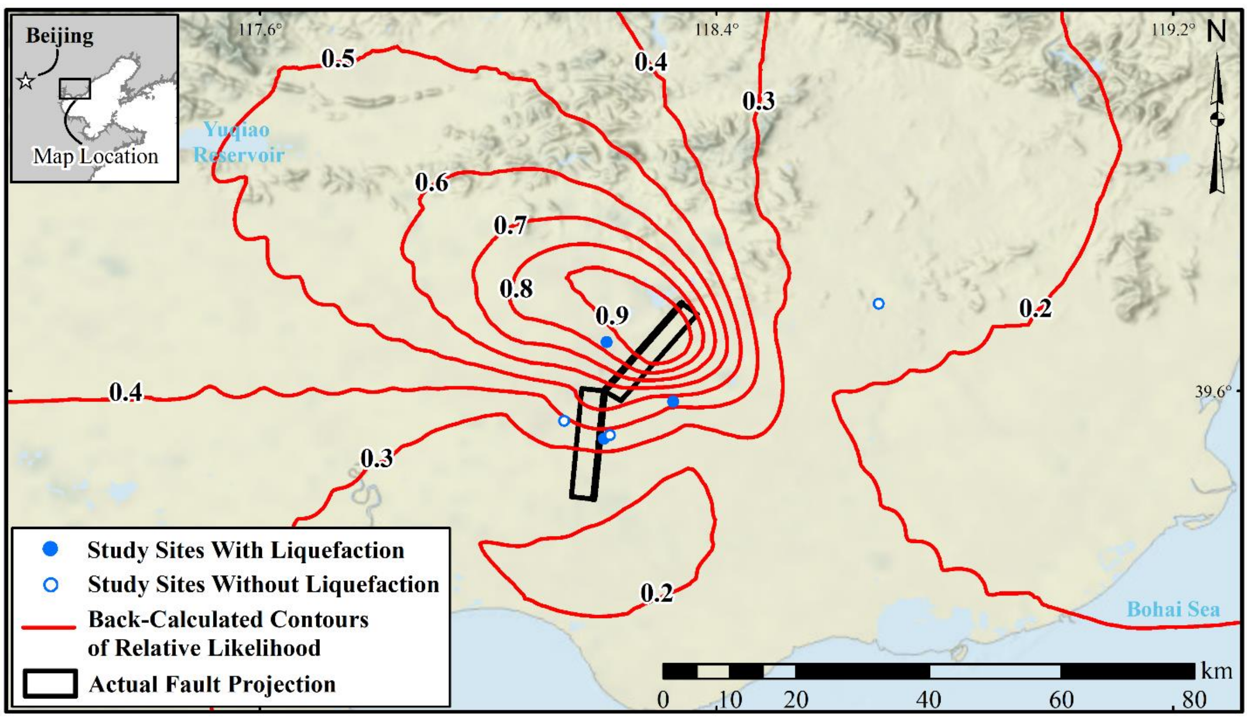


Fig. 17. Analysis of six sites from the 1976 Tangshan, China earthquake: contours showing the relative likelihood of source locations. Black rectangles = actual fault projections (Huang and Yeh, 1997).

3.2.4 1995 Kobe, Japan

The 1995 M_w6.9 Kobe, Japan, earthquake occurred on the Nojima, Suma, and Suwayama Faults (Wald, 1995). Fourteen liquefaction case histories were analyzed, the results of which are in Figs. 8g and 18. The identified, most likely source was in Osaka Bay, 16 km southeast of the nearest fault projection. The corresponding median M_w was M_w6.5 with 95% CI of M_w5.8-M_w8.1, as shown by the CDF in Fig.

8g, versus the actual M_w6.9. While not grossly erroneous, these results highlight the importance of compiling study sites that are spatially distributed, increasing the odds that they have variable site-to-source distances and azimuths and in-turn, experienced a range of ground-motion intensities (required for accurate geolocation). As mapped in Fig. 18, the 14 study sites are in similar proximity to the fault rupture and thus, likely experienced similar *PGA*s in the Kobe earthquake, complicating geolocation. Had sites from the southern shores of Osaka Bay been available for analysis (where liquefaction was not observed) the likely source would potentially shift north, closer to actual.

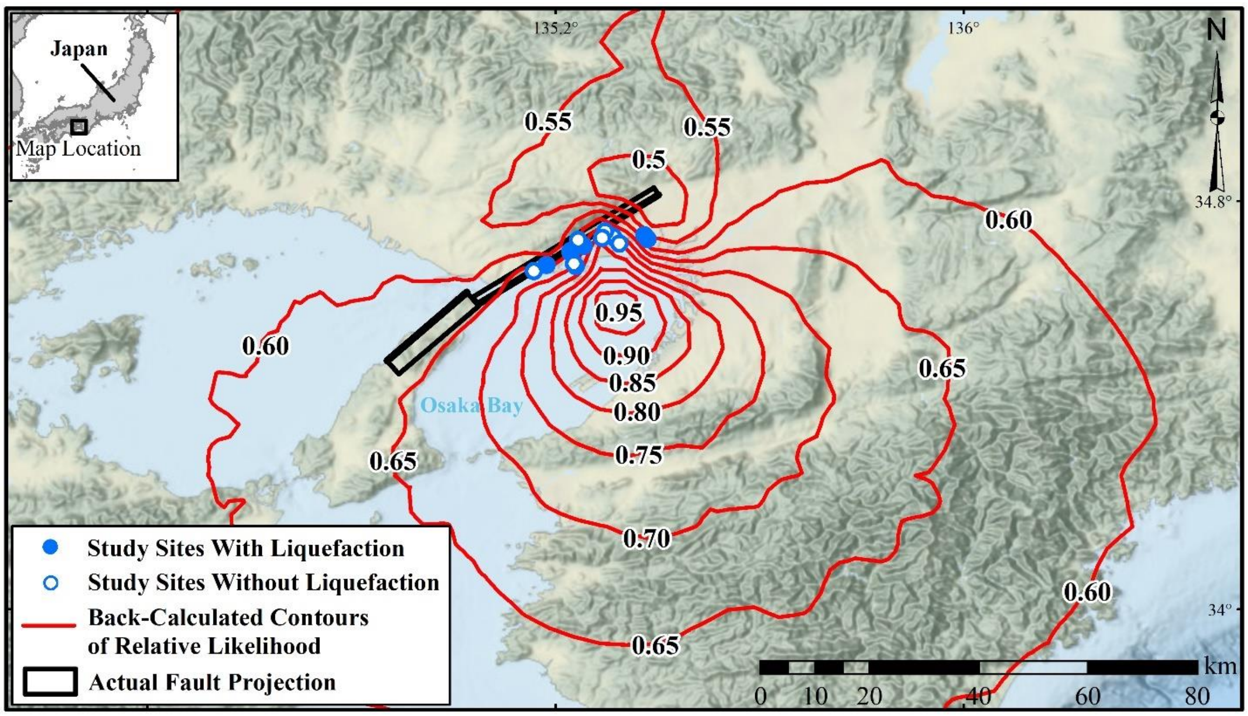


Fig. 18. Analysis of 14 study sites from the 1995 Kobe, Japan earthquake: contours showing the relative likelihood of source locations. Black rectangles = actual fault projection (Wald, 1995).

3.2.5 1999 Chi-Chi, Taiwan

The 1999 M_w7.6 Chi-Chi, Taiwan, earthquake occurred on the Chelungpu Fault (Chi et al. 2001). Twenty-eight liquefaction case histories were analyzed, the results of which are in Figs. 8g and 19. The inversion correctly located the fault rupture, yet the corresponding median M_w here was M_w5.9, an underestimate of M_w1.7. Upon investigation, the erroneous magnitude is partly attributable to ground motion and liquefaction model inaccuracies caused, in part, by atypical profiles of highly interbedded silty soils. To confirm, site-specific amplification factors were first estimated at each study site following the method previously applied in Loma Prieta, wherein predicted ground-motions are scaled to match those observed/recorded, mimicking that which might be achieved via site-specific ground-motion analysis. Second, because *LPI* is prone to overpredict liquefaction manifestations on profiles highly interbedded with silts and clays (e.g., Maurer et al., 2015b), a Chi-Chi specific fragility function was developed. Silty, interbedded profiles are common among liquefaction case histories published from Chi-Chi, where the observed liquefaction was less than predicted by *LPI* (Juang et al., 2008). The inversion framework's underestimated magnitude (i.e., M_w5.9) was thus partly because observed liquefaction was less than expected. Utilizing the methodology of Geyin and Maurer (2020a), and analyzing the Chi-Chi case histories compiled therein, a Chi-Chi specific fragility function was developed and is given by Eq. 6 using $\theta = 11.629$ and $\beta = 1.806$. This θ, which is the *LPI* value with 50% probability of manifestation, is significantly higher than that proposed by Geyin and Maurer (2020a) for global case histories, indicating that liquefaction was less likely in Chi-Chi at a given *LPI*, as compared to other global events. Of course, the development of a region-specific liquefaction model is generally not practical, but was undertaken here to confirm the cause of the underpredicted M_w. The inversion framework was reapplied using these adjusted components, with the results shown in the Fig. 19 inset and as CDFs in Fig. 8h. This resulted in a geolocation error of 0 km and overall M_w CDF with median of M_w7.45 and 95% CI of M_w6.6-M_w8.1, versus the actual M_w7.6.

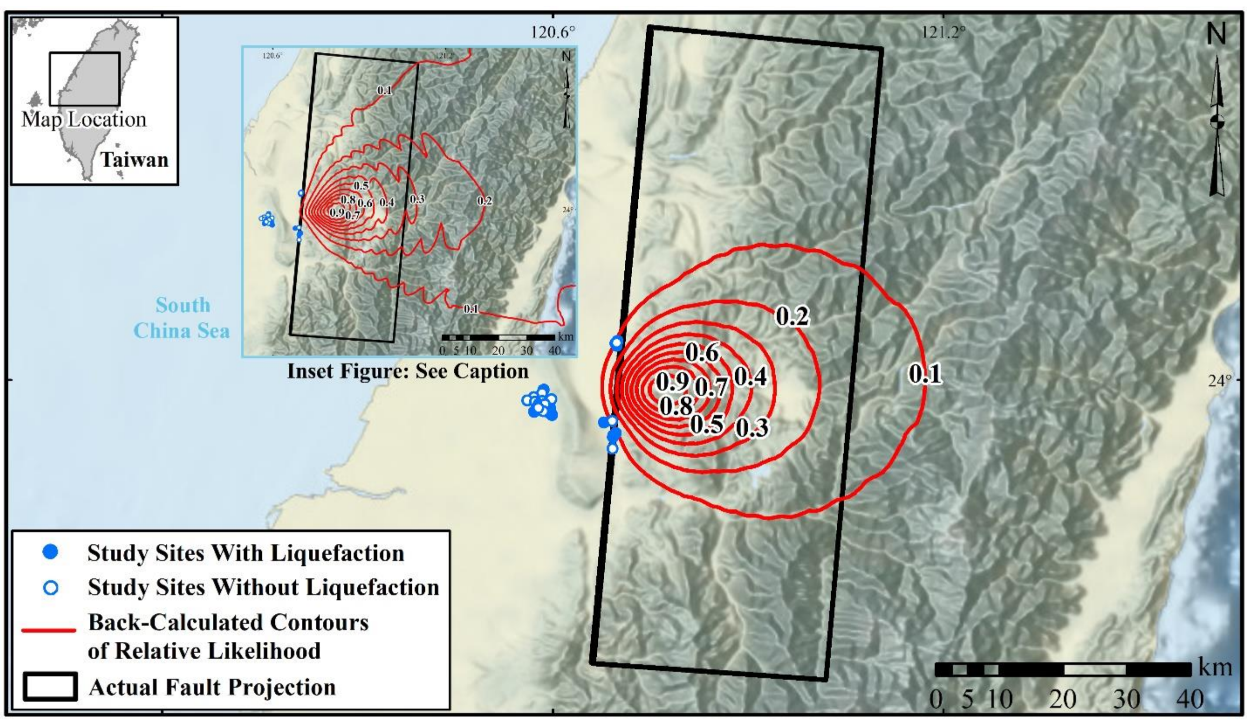


Fig. 19. Analysis of 28 study sites from the 1999 Chi-Chi, Taiwan, earthquake: contours showing the relative likelihood of source locations. The inset figure shows results after inclusion of region-specific liquefaction fragility functions and site-specific ground motion amplification factors. Black rectangles = actual fault projection (Chi et al. 2001).

4. Discussion and Conclusions

In regions of infrequent but potentially damaging seismicity, analysis of relic coseismic evidence, such as paleoliquefaction, is commonly used to infer information about the seismic hazard. However, existing paleoliquefaction analysis techniques require *a priori* assumptions about the causative earthquake's location, potentially leading to erroneous results. Accordingly, this paper proposed an inversion framework to probabilistically constrain seismic parameters from paleoliquefaction evidence, wherein source-location uncertainty is explicitly quantified. The framework is general and modular in nature, meaning it is transferrable to other types of evidence (e.g., landslides) and accommodating of alternate component models. In this regard, the latest regionally appropriate models should be used to predict ground motions and liquefaction, which in the future will invariably differ from those used here. Simulated paleoliquefaction studies were performed on modern earthquakes with known parameters, wherein the quantity of field data varied, but is conceivable of paleoliquefaction investigations. The most salient results of these simulations are summarized in Table 1.

Table 1. Summary of simulated paleoliquefaction analyses for eight modern earthquakes.

Earthquake	Study Sites	Geolocation Error (km)	Inverted M_w, Considering Most Likely Source[2]	Inverted M_w, Considering All Sources[2]	Actual M_w
1976 Tangshan, China	6	0	6.8 (6.0-8.5)	7.8 (6.4-8.8)	7.6
1979 Imperial Valley, USA	6	0	6.1 (5.2-7.3)	8.0 (6.1-8.8)	6.5
1989 Loma Prieta, USA	46	29	7.8 (7.6-8.1)	8.0 (7.5-8.85)	6.9
1989 Loma Prieta, USA Revisit[1]	46	0	6.9 (6.6-7.3)	7.05 (6.5-7.7)	6.9
1995 Kobe, Japan	14	16	6.5 (5.8-8.1)	7.5 (6.1-8.5)	6.9
1999 Chi-Chi, Taiwan	28	0	5.9 (5.6-6.3)	6.6 (5.6-7.6)	7.6
1999 Chi-Chi, Taiwan Revisit[1]	28	0	7.0 (6.75-7.2)	7.45 (6.6-8.1)	7.6
2010 Darfield, New Zealand	100	20	6.65 (6.4-6.9)	7.35 (6.1-8.1)	7.1
2011 Christchurch, New Zealand	100	0	6.4 (6.1-6.65)	6.8 (6.1-7.9)	6.3
2016 Christchurch, New Zealand	100	0	6.0 (5.8-6.25)	6.35 (5.7-7.9)	5.7

[1]Events reanalyzed to identify and correct sources of error, as discussed in the text.
[2]The values listed are the median magnitude and (in parentheses) its 95% confidence interval.

These examples demonstrated the framework's potential, even in cases of limited field data (e.g., 1976 Tangshan and 1979 Imperial Valley events), as well as important caveats and lessons for forward use. As a framework that relies on predicting multiple phenomena in series (e.g., ground-motion attenuation, site response, liquefaction triggering, and liquefaction manifestation), inaccuracies are likely whenever the response of one or more phenomena differs greatly from expectations. The 1989 Loma Prieta and 1999 Chi-Chi events, for example, demonstrated how shortcomings in component models can lead to inaccurate seismic-source inversions. In this regard, more rigorous seismic site-characterization and ground motion study (i.e., *Vs* profiling and wave propagation-based site-response analysis) may help to mitigate GMM shortcomings. Moreover, as a modular framework, performance should improve over time as better component models are continuously developed (e.g., improved liquefaction triggering and manifestation models). Because the framework relies on seismic waves arriving to different sites with different amplitudes, the difficulty of accurate analysis increases as *PGA* values become similar among study sites, whether due to sites having similar site-to-source distance (1995 Kobe) or all being relatively distant from fault rupture, where *PGA* attenuation rates are lower (2010 Darfield). As a result, it is best practice to compile spatially distributed study sites, since multiple sites in proximity may provide little, if any, additional constraint on source parameters. Nonetheless, events in which the liquefaction field is far from the seismic source may be more difficult to analyze. While it was shown that accurate results are possible even with limited data, analyses from the 2010, 2011, and 2016 Canterbury events provided insights into possible finite-sample uncertainties. As an example, the likelihood of accurate results generally improved significantly when increasing from 10 total study sites to 30, 40, or 50, beyond which diminishing returns were observed. In addition, the applications showed the potential benefits of on-fault or other geologic evidence and interpretation toward further constraint of the source location, which in general reduces the M_w uncertainty. Considering all events studied herein, the average 95% CI range decreased from ~2M_w considering source location uncertainty to ~1M_w when conditioned on a single source. In either case (i.e., whether a causative fault is strongly suspected or not), the inversion framework can provide an objective and transparent quantification of uncertainty, given a set of field observations. In this regard, the framework inherently requires geotechnical test data at study sites, which may be difficult and expensive to acquire. In such cases, geologic interpretation can fill gaps in, and inform, the geotechnical analyses.

While provisional earthquake sources were treated as epicenters (and then converted to statistical fault realizations), the proposed framework could directly assess the plausibility of 3-dimensional faults with minimal modification. By analyzing planes in lieu of epicenters, it may be possible to constrain not only the rupture location, but also its geometry and depth, provided that different fault geometries and depths result in different ground motions at study sites. The inversion framework, as implemented herein, did not attempt to quantify every uncertainty in a paleoliquefaction study, but it did demonstrate how significant, quantifiable uncertainties may be accounted for (i.e., those of ground motions and liquefaction response); additional uncertainties, if known (e.g., those of geologic field interpretation or in-situ test data), may be similarly incorporated via the total probability theorem. While paleoliquefaction evidence has been analyzed widely, with findings influencing national seismic-hazard maps, paleoliquefaction research has to-date focused more on field investigations and less on advancing the techniques for inverse-analyzing field evidence. In this regard, the proposed framework has the potential to provide new insights in enigmatic seismic zones worldwide.

5. Acknowledgments

We gratefully acknowledge Jack Baker at Stanford University for providing a code implementation of the Chiou and Youngs (2014) GMM (https://web.stanford.edu/~bakerjw/GMMs.html).

6. Data Availability

The Canterbury case-history dataset analyzed herein is available in digital format from NEHRI DesignSafe: https://doi.org/10.17603/ds2-tygh-ht91.

7. Supplemental Material

Appendix A, containing contour maps of the median M_w for each event, is available online.

8. Funding

The presented study is based on work supported by the National Science Foundation (NSF) under Grant No. CMMI-1751216, by the NSF Graduate Research Fellowship Program under Grant No. DGE-1762114, and by the University of Washington Royalty Research Fund (RRF). However, any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF or RRF.

9. References

Ambraseys NN (1988) Engineering seismology. *Earthquake Engineering and Structural Dynamics* 17: 1-105.

Archuleta RJ (1984) A Faulting Model for the 1979 Imperial Valley Earthquake. *Journal of Geophysical Research* 89(B6): 4559-4585.

Aydan O, Ulusay R, Kumsar H, and Tuncay E (2000) Site investigation and engineering evaluation of the Duzce-Bolu earthquake of November 12, 1999. *Turkish Earthquake Foundation, Istanbul*. Report No. TDV/DR 09-51, 307 p.

Bastin S, Bassett K, Quigley MC, Maurer BW, Green RA, Bradley BA, and Jacobson D (2016) Late Holocene liquefaction at sites of contemporary liquefaction during the 2010-2011 Canterbury Earthquake Sequence. *BSSA* 106(3): 881-903.

Beavan J, Motagh M, Fielding EJ, Donnelly N, and Collett D (2012) Fault slip models of the 2010-2011 Canterbury, New Zealand, earthquakes from geodetic data and observations of postseismic ground deformation. *New Zealand Journal of Geology and Geophysics* 55(3): 207-221.

Boulanger RW and Idriss IM (2014) *CPT and SPT Based Liquefaction Triggering Procedures*, Report No. UCD/CGM-14/01, Center for Geotechnical Modeling, University of California, Davis, CA.

Bradley BA (2013) A New Zealand-Specific Pseudospectral Acceleration Ground-Motion Prediction Equation for Active Shallow Crustal Earthquakes Based on Foreign Models. *BSSA* 103(3): 1801-1822.

California Geological Survey (2014) Fault Activity Map of California. Scale not given. Available at: https://maps.conservation.ca.gov/cgs/fam/.

Castilla RA and Audemard FA (2007) Sand blows as a potential tool for magnitude estimation of pre-instrumental earthquakes. *Journal of Seismology* 11: 473-487.

Chi W, Dreger D, and Kaverina A (2001) Finite-Source Modeling of the 1999 Taiwan (Chi-Chi) Earthquake Derived from a Dense Strong-Motion Network. *BSSA* 91(5): 1144-1157.

Chiou BS-J and Youngs RR (2014) Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. *EQS* 30(3): 1117-1153.

Cox RT, Hill AA, Larsen D, Holzer T, Forman SL, Noce T, Gardner C, and Morat J (2007) Seismotectonic implications of sand blows in the southern Mississippi Embayment. *Engineering* Geology 89: 278-299.

Elnashai AS, Cleveland LJ, Jefferson T, and Harrald J (2009) *Impact of earthquakes on the Central USA*. Mid-America Earthquake Center Report 09-03, University of Illinois at Urbana Champaign, Urbana, IL.

Galli P (2000) New empirical relationships between magnitude and distance for liquefaction. *Tectonophysics* 324: 169-187.

Geyin M and Maurer BW (2020a) Fragility functions for liquefaction-induced ground failure. *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.

Geyin M, Baird AJ, and Maurer BW (2020b) Field assessment of liquefaction prediction models based on geotechnical vs. geospatial data, with lessons for each. *Earthquake Spectra* 36(3): 1386–1411.

Geyin M, Maurer BW, Bradley BA, Green RA, and van Ballegooy S (2021) CPT-based liquefaction case histories compiled from three earthquakes in Canterbury, New Zealand. *EQS:* 8755293021996367.

Giona Bucci M, Almond P, Villamor P, Ries W, Smith C, and Tuttle M (2017), When the earth blisters: exploring recurrent liquefaction features in the coastal system of Christchurch, New Zealand: *Terra Nova* (29): 162-172.

Giona Bucci M, Almond P, Villamor P, Tuttle M, Stringer M, Smith C, Ries W, Bourgeois J, Loame R, Howarth J, Watson M, (2018) Controls on patterns of liquefaction in a coastal dune environment. *Sedimentary Geology* 337: 17-33.

Giona Bucci M, Villamor P, Almond P, Tuttle M, Stringer M, Ries W, Smith C, Hodge M, and Watson M (2018) Associations between sediment architecture and liquefaction susceptibility in fluvial settings: the 2010-20111 Canterbury Earthquake Sequence, New Zealand. *Engineering Geology* 237: 181-197.

Green RA, Maurer BW, Bradley BA, Wotherspoon L, and Cubrinovski M (2014) *Implications from liquefaction observations in New Zealand for interpreting paleoliquefaction data in the central eastern United States*. U.S. Geological Society Technical Report G12AP20002, 97pp.

Green RA, Obermeier SF, Olson SM (2005) Engineering geologic and geotechnical analysis of paleoseismic shaking using liquefaction effects: field examples. *Engineering Geology* 76: 263-293.

Green RA, Maurer BW, and Haskell A (2017) *Development of probabilistic magnitude-bound curves for the New Madrid Seismic Zone (NMSZ) for paleoliquefaction studies*. U.S. Geological Society Technical Report G14AP00046, 102pp.

Green RA, Bommer JJ, Rodriguez-Marek A, Maurer BW, Stafford PJ, Edwards B, Kruiver PP, De Lange G and Van Elk J (2019) Addressing limitations in existing 'simplified' liquefaction triggering evaluation procedures: application to induced seismicity in the Groningen gas field. *Bulletin of Earthquake Engineering* 17(8): 4539-4557.

Hayati H. and Andrus RD (2008) Liquefaction potential map of Charleston, South Carolina base on the 1886 earthquake. *Journal of Geotechnical and Geoenvironmental Engineering* 134(6): 815-828.

Heath DC, Wald DJ, Worden CB, Thompson EM, & Smoczyk GM (2020) A global hybrid VS 30 map with a topographic slope–based default and regional map insets. *EQS* 8755293020911137.

Holzer TL, Noce TE, and Bennett MJ (2015) Strong ground motion inferred from liquefaction caused by the 1811-1812 New Madrid, Missouri, Earthquakes. *BSSA* 105(5): 2589-2603.

Huang B and Yeh YT (1977) The Fault Ruptures of the 1976 Tangshan Earthquake Sequence Inferred from Coseismic Crustal Deformation. *BSSA* 87(4), 1046-1057.

Iwasaki T, Tatsuoka F, Tokida K, and Yasud S (1978) A practical method for assessing soil liquefaction potential based on case studies at various sites in Japan. *2nd Intl Conf. Microzonation*.

Juang CH, Liu CN, Chen CH, Hwang JH, & Lu CC (2008) Calibration of liquefaction potential index: A re-visit focusing on a new CPTU model. *Engineering Geology 102*(1-2): 19-30.

Kaiser A, Holden C, Hamling I, Hreinsdottir S, Horspool N, Massey C, ... & Benites R (2016) The 2016 valentine's day M_w5.7 Christchurch earthquake: preliminary report. In: *Proceedings of the New Zealand Society for Earthquake Engineering Annual Conference*, Christchurch, NZ, pp. 1-3.

Kaklamanos J, Baise LG, & Boore DM (2011) Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice. *EQS* 27(4): 1219-1235.

Kuribayahsi E and Tatsuoka F (1975) Brief review of liquefaction during earthquakes in Japan. *Soils and Foundations* 15(4): 81-92.

Langridge RM, Ries WF, Litchfield NJ, Villamor P, Van Dissen RJ, Barrell DJA, Rattenbury MS, Heron DW, Haubrock S, Townsend DB, Lee JM, Berryman KR, Nicol A, Cox SC, and Stirling MW (2016) The New Zealand Active Faults Database. *New Zealand Journal of Geology and Geophysics* 59(1): 86-96.

Martin JR and Clough GW (1990) *Implications from a geotechnical investigation of liquefaction phenomena associated with seismic events in the Charleston, SC area*. Report to USGS Grant No. 14-08-001-G-1348, Virginia Tech, Blacksburg, VA, 414 p.

Maurer BW, Green RA, Cubrinovski M and Bradley BA (2014a) Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand. *Journal of Geotechnical and Geoenvironmental Engineering* 140(7): 04014032.

Maurer BW, Green RA, Bradley BA and Cubrinovski M (2014b) Evaluating the efficacy of paleoliquefaction analysis techniques using modern analogs. *Geological Society of America Abstracts with Programs* 46(3): 23; GSA Southeastern Section 63rd Annual Meeting.

Maurer BW, Green RA, Quigley MC and Bastin S (2015a) Development of magnitude-bound relations for paleoliquefaction analyses: New Zealand case study. *Engineering Geology* 197: 253-266.

Maurer BW, Green RA, Cubrinovski M and Bradley BA (2015b) Fines-content effects on liquefaction hazard evaluation for infrastructure during the 2010-2011 Canterbury, New Zealand earthquake sequence. *Soil Dynamics and Earthquake Engineering* 76, 58-68.

Maurer BW, Green RA, Wotherspoon LM & Bastin S (2019) The Stratigraphy of Compound Sand Blows at Sites of Recurrent Liquefaction: Implications for Paleoseismicity Studies. *EQS* 35(3), 1421-1440.

Murphy T (2011) FEMA: freaking over a fake quake? *Mother Jones*. available at: < http://www.motherjones.com/politics/2011/03/fema-earthquake-new-madrid/>.

Nuclear Regulatory Commission (NRC) (2012) *Central and Eastern United States Seismic Source Characterization for Nuclear Facilities*. Tech.Rep.NUREG-2115, NRC, Rockville, MD.

Obermeier S (1989) *The New Madrid earthquakes: an engineering geologic interpretation of relict liquefaction features*. U.S. Geological Survey Professional Paper 1336-B, 114 pg.

Obermeier SF, Olson SM and Green RA (2005). Field occurrences of liquefaction-induced features: a primer for engineering and geologic analysis of paleoseismic shaking. *Engineering Geology* 76: 209-234.

Obermeier SF, Pond EC, Olsen SM with contributions by Green RA, Mitchell JK, and Stark TD (2001). *Paleoliquefaction studies in continental settings: geologic and geotechnical factors in interpretations and back-analysis*. U.S. Geological Survey Open-File Report 01–029

Obermeier SF and Dickenson SE (2000). Liquefaction evidence for the strength of ground motions resulting from late Holocene Cascadia subduction earthquakes, with emphasis on the event of 1700 A.D. *BSSA* 90(4): 876-896.

Olson SM, Green RA and Obermeier SF (2005a) Revised magnitude bound relation for the Wabash valley seismic zone of the central United States. *Seismological Research Letters* 76(6): 756–771.

Olson SM, Green RA and Obermeier SF (2005b) Geotechnical analysis of paleoseismic shaking using liquefaction features: a major updating. *Engineering Geology* 76: 235–261.

Olson SM, Wen YK, Song J, Johnson CI and Muhammad K (2013) *Quantifying uncertainties in paleoliquefaction studies.* U.S. Geological Society Technical Report 06HQGR0013, 47pp.

Papadopoulos GA and Lefkopoulos G (1993) Magnitude-distance relations for liquefaction in soil from earthquakes. *Bulletin of the Seismological Society of America* 83(3): 925-938.

Papathanassiou G, Pavlides S, Christaras B and Pitilakis K (2005) Liquefaction case histories and empirical relations of earthquake magnitude versus distance from the broader Aegean region. *Journal of Geodynamics* 40: 257-278.

Petersen MD, Frankel AD, Harmsen SC, Mueller CS, Haller KM, Wheeler RL, Wesson RL, Zeng Y, Boyd OS, Perkins DM, Luco N, Field EH, Wills CJ, Rukstales KS (2008) *Documentation for the 2008 update of the United States national seismic hazard maps.* USGS Open-File Report 2008-1128, 61 p.

Petersen MD, Moschetti MP, Powers PM, Mueller CS, Haller KM, Frankel AD, Zeng Y, Rezaeian S, Harmsen SC, Boyd OS, Field N, Chen R, Rukstales KS, Luco N, Wheeler RL, Williams RA and Olsen AH (2014) *Documentation for the 2014 update of the United States national seismic hazard maps.* USGS Open-File Report 2014–1091, 243 p.

Pirrotta C, Barbano MS, Guarnieri P and Gerardi F (2007) A new dataset and empirical relationships between magnitude/intensity and epicentral distance for liquefaction in central-eastern Sicily. *Annals of Geophysics* 50(6): 763-774.

Quigley MC, Bastin S, Bradley BA (2013) Recurrent liquefaction in Christchurch, New Zealand, during the Canterbury earthquake sequence. *Geology* 41: 419–422.

Quigley MC, Hughes MW, Bradley BA, van Ballegooy S, Reid C, Morgenroth J, Horton T, Duffy B and Pettinga J (2016) The 2010-2011 Canterbury earthquake sequence: Environmental effects, seismic triggering thresholds, geologic legacy. *Tectonophysics* 672-673: 228-274.

Rodriguez-Marek A, and Ciani D (2008) Probabilistic methodology for the analysis of paleoliquefaction features. *Engineering Geology* 96(3): 159-172.

Scherbaum F, Schmedes J and Cotton F (2004) On the Conversion of Source-to-Site Distance Measures for Extended Earthquake Source Models. *BSSA* 94(3): 1053-1069.

Seed HB and Idriss IM (1971) Simplified procedure for evaluating soil liquefaction potential. *Journal of Soil Mechanics and Foundations Division* ASCE, 107(SM9): 1249-1274.

Sims JD (1973) Earthquake-induced structures in sediment of Van Norman Lake, San Fernando, California, *Science*, 182, 161-163, doi:10.1126/science.182.4108.161.

Sims JD & Garvin CD (1995) Recurrent liquefaction induced by the 1989 Loma Prieta earthquake and 1990 and 1991 aftershocks: implications for paleoseismicity studies. *BSSA* 85(1): 51-65.

Stein S (2008) Presentation to Arkansas legislative committee on economic development on earthquake safety provisions in building code. Little Rock, Arkansas; January 28, 2008. Available at: < http://www.earth.northwestern.edu/people/seth/research/testimony.html>.

Tuttle MP (2001) The use of liquefaction features in paleoseismology: lessons learned in the New Madrid seismic zone, central United States. *Journal of Seismology* 5: 361-380.

Tuttle MP and Atkinson GM (2010) Localization of large earthquakes in the Charlevoix seismic zone, Quebec, Canada during the past 10,000 years. *Seismological Research Letters* 81: 140-147.

Tuttle MP, Sims JD, Dyer-Williams K, Lafferty RHIII and Schweig ESIII (2000) *Dating of Liquefaction Features in the New Madrid Seismic Zone: U.S. Nuclear Regulatory Commission*, NUREG/GR-0018, 42 pp.

Tuttle M and Hartleb R (2012) *CEUS Paleo-liquefaction Database, Uncertainties Associated with Paleo-liquefaction Data, and Guidance for Seismic Source Characterization: Central and Eastern United States Seismic Source Characterization for Nuclear Facilities, Appendix E*, NUREG-2115, U.S. Nuclear Regulatory Commission (NRC), Rockville, MD.

Tuttle MP, Villamor P, Almond P, Bastin S, Giona Bucci M, Langridge R, ... & Hardwick CM (2017) Liquefaction induced during the 2010–2011 Canterbury, New Zealand, earthquake sequence and lessons learned for the study of paleoliquefaction features. *Seismological Research Letters* 88(5): 1403-1414.

Tuttle MP, Schweig ES, Campbell J, Thomas PM, Sims JD and Lafferty RH (2005) Evidence for New Madrid earthquakes in A.D. 300 and 2350 B.C. *Seismological Research Letters* 76: 489-501.

Tuttle MP, Schweig ES, Sims JD, Lafferty RH, Wolf LW and Haynes ML (2002) The earthquake potential of the New Madrid Seismic Zone. *BSSA* 92(6): 2080-2089.

Tuttle MP, Hartleb R, Wolf L, and Mayne PW (2019a) Paleoliquefaction studies and the evaluation of seismic hazard, *Geosciences*, 9(7), 61 p, doi:10.3390/geosciences9070311.

Tuttle MP, Wolf LW, Starr ME, Villamor P, Mayne PW, Lafferty RH, et al. (2019b) Evidence for large New Madrid earthquakes about A.D. 0 and B.C. 1050, Central United States, *Seismological Research Letters* 90(3): 1393-1406.

Vidale J, Atkinson G, Green R, Hetland E, Grant-Ludwig L, Mazzotti S, Nishenko S and Sykes L (2011) *Report of the independent expert panel on New Madrid Seismic Zone earthquake hazards as approved by NEPEC on April 16, 2011*. U.S. Geological Survey, 26 p.

Villamor P, Almond P, Tuttle MP, Giona-Bucci M, Langridge RM, Clark K, ... & Howarth J. (2016) Liquefaction features produced by the 2010–2011 Canterbury earthquake sequence in southwest Christchurch, New Zealand, and preliminary assessment of paleoliquefaction features. *BSSA 106*(4): 1747-1771.

Wakamatsu K (1993) *History of soil liquefaction in Japan and assessment of liquefaction potential based on geomorphology*. A Thesis in the Department of Architecture Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Engineering, Waseda University, Tokyo, Japan.

Wald DJ (1995) A Preliminary Dislocation Model for the 1995 Kobe (Hyogo-ken Nanbu), Japan, Earthquake Determined from Strong Motion and Teleseismic Waveforms. *Seismological Research Letters* 66(4): 22-28.

Wald DJ, Helmberger DV and Heaton TH (1991) Rupture Model of the 1989 Loma Prieta Earthquake From the Inversion of Strong-Motion and Broadband Teleseismic Data. *BSSA* (5): 1540-1572.

Wald DJ, Worden BC, Quitoriano V & Pankow KL (2005) *ShakeMap manual: technical manual, user's guide, and software guide* (No. 12-A1).

Wells DL and Coppersmith KJ (1994) New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, And Surface Displacement. *BSSA* 84(4): 974-1002.

APPENDIX C

Publication 3 (Journal Paper 2), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.

[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

Probabilistic Seismic Source Inversion from Regional Landslide Evidence

Ryan A. Rasanen[c] and Brett W. Maurer[d]

Abstract: In regions of infrequent but potentially damaging seismicity, modern earthquake inventories may be insufficient to provide inputs to seismic-hazard analyses (i.e., fault locations and magnitude-frequency relations). As a result, analysis of paleoseismic evidence, such as coseismic landsliding, is commonly used to help elucidate the seismic record, thereby reducing seismic-hazard uncertainty. However, while paleolandslides have been investigated widely, existing inverse-analysis techniques (i.e., to constrain the causative earthquake magnitude and/or ground motions) have several shortcomings. Namely, they: (i) require the location of the causative earthquake to be known; (ii) provide only a lower-bound estimate of seismic parameters; and (iii) are deterministic in nature. Accordingly, this paper proposes a flexible inversion framework that probabilistically constrains seismic-source parameters from regional paleolandslide evidence. The outputs of this framework are: (i) a geospatial likelihood surface that constrains the location of fault rupture; and (ii) a probability distribution of the rupture magnitude. Simulated paleolandslide studies are performed on modern earthquakes with known parameters. These examples demonstrate the framework's provocative potential as well as important lessons for implementation. The proposed framework has the potential to extract new insights from relic landslide evidence in seismic zones worldwide.

1. Introduction

In many seismic zones, the return period of potentially damaging earthquakes is longer than the duration of seismic observation. As a result, earthquake catalogs may be insufficient to inform the inputs requisite for seismic-hazard analyses (i.e., the locations, magnitudes, and frequencies of fault ruptures). The uncertainties of these parameters are of significant interest, given that they influence the computed seismic hazard adopted by building codes and used by policy makers (e.g., Vidale et al., 2011). In such cases, paleoseismic relics (i.e., artifacts from pre-historic or pre-instrumental earthquakes) are commonly studied to elucidate the seismic hazard. This includes turbidites (e.g., Goldfinger et al. 2012), tsunami deposits (e.g., Peters et al. 2007), dendrochronology (e.g., Atwater et al. 1991), soil liquefaction (e.g., Maurer et al. 2015), microfossils (e.g., Engelhart et al. 2013), geochemical markers (e.g., O'Donnell et al. 2017), and seafloor morphology (e.g., Watt et al. 2017). Among the many types of such evidence that have been analyzed, paleolandslides are particularly valuable for their ability to "record" the intensity of prior ground motions. This distinction is owed to the fact models exist for predicting landslides as a function of ground motions (or at a minimum, ground-motion intensity measures, *IM*s). In forward hazard analyses, wherein the seismic loading is given, these models are widely used to predict future landslides. In paleoseismic studies, wherein the outcome is given (e.g., landslides were or were not observed), the models can be inverted to constrain the ground motions that likely would, and would not, produce the observation (e.g., Yagoda-Biran et al. 2010; Wang 2020). Although the date of an earthquake, and thus the recurrence rate, can be derived from many types of evidence, few have a quantifiable relationship with ground motion parameters (e.g., Rasanen et al. 2021). In this respect, paleolandslides are more than just a proxy of past shaking, given that they may quantitatively constrain the motions experienced.

While paleoseismic studies have significant influence on computed seismic hazards in some regions (e.g., Petersen et al. 2008; 2014), paleolandslide research has to-date generally focused more on field investigations (i.e., identification, interpretation, and dating of landslides) and less on advancing the inversion procedures to constrain the causative seismic parameters. All phases of investigation are important, but the focus of this study is on the latter. Notably, existing paleolandslide analysis methods suffer from up to three shortcomings. *First*, existing methods are underpinned by the need to locate the earthquake that induced the observed field of evidence (e.g., earthquake magnitude cannot otherwise be constrained). This requires assumptions in many settings (e.g., where faults are blind or historically quiescent), and in turn, may lead to erroneous characterizations of the seismic hazard. *Second*, existing

[c] Graduate Research Assistant, University of Washington, Seattle; rrasanen@uw.edu
[d] Assistant Professor, University of Washington, Seattle; bwmaurer@uw.edu

methods have traditionally analyzed slopes *with* observed landslides (i.e., positive observations) but not slopes *without* observed landslides (i.e., negative observations). As a result, derivative estimates of shaking intensity and/or earthquake magnitude are inherently lower-bound estimates, even if this is not explicitly stated. *Third*, existing methods have been deterministic in nature, whereas paleolandslides involve multiple uncertainties that should be accounted for probabilistically to arrive at a transparent and useful characterization of the seismic hazard.

Accordingly, the objective of this paper is to formulate and test a probabilistic framework for inverting seismic-source parameters from ancient landslides. Analyzing regional evidence, this framework computes the likelihood of a rupture with given location, geometry, and magnitude producing a set of field observations, wherein various measurement and modelling uncertainties may be included. Repeating for a near-infinite number of possible faults results in a regional-scale understanding of the likely source parameters. The degree of successful constraint will be investigated using tests on modern earthquakes with known parameters. That is, simulated paleolandslide studies will be performed to assess the framework's efficacy. In the following, existing paleolandslide analytics are briefly summarized, to include discussion of the aforementioned shortcomings. An overview of the framework proposed herein is then given, followed by specific implementation details. Lastly, tests of the framework on four modern earthquakes are presented, from which important lessons and caveats for forward use are gleaned.

1.1 Existing Paleolandslide Analytics and Their Shortcomings

Paleolandslide studies can potentially provide information about the dates, locations, and magnitudes of paleoearthquakes, as well as the resultant, regionally distributed ground motions. A complete study of paleolandslides is three phased: (i) field identification and interpretation; (ii) dating; and (iii) constraint of the earthquake magnitude and/or ground motion under which the field of evidence was formed. Interpreting whether a paleolandslide is of seismic origin is inherently challenging. Crozier (1992), for example, proposed six criteria for this purpose, as based on observations in New Zealand. In general, the most important markers of seismic origin are: (i) multiple spatiotemporally clustered landslides for which (ii) static slope-stability analyses predict stability even in saturated conditions; and where (iii) other paleoseismic evidence of the same age is also found. For overviews of field identification and interpretation, the reader is referred to Jibson (1996) and Clague (2014), which may be complimented with specific case studies that highlight the use of remote-sensing technology (e.g., Sutinen et al. 2014). Once identified, paleolandslides may be dated via radiocarbon (e.g., Stout 1977; Ojala et al. 2018); optically stimulated luminescence (e.g., Katz et al. 2011); dendrochronology (e.g., Hupp et al. 1987; Struble et al. 2020), cosmic ray exposure (e.g., Le Roux et al. 2009; Mackey and Quigley 2014), and several relative dating approaches (e.g., lichenometry – Bull et al. 1994), among others. Ideally, multiple dating methods are used to reduce uncertainty (Johnson 1987). A comprehensive overview of dating methods is given by Jibson (1996) and Panek (2015). Panek (2015) also provides a list of regions with well-established landslide chronologies. These regional chronologies – excellent examples of which include Chen et al. (2012) and Martino et al. (2014) – form both the basis of paleoseismic research and the landslide inventories required to train and test improved landslide prediction models. In some cases, these chronologies may provide more recent analog events that aide in the interpretation of older paleoearthquakes.

Following field interpretation, dating, and inventorying, the techniques by which the causative earthquake magnitude and/or shaking intensity are quantitatively constrained are generally called back- or inverse-analysis methods. While several such methods have been used, we refer to the two most common and credible to-date as the "magnitude-bound" method (e.g., Keefer 1984; Rodriguez et al. 1999; Papadopoulos and Plessa 2000; Delgado et al. 2011) and the "site-specific stability analysis," or for brevity, the "site-specific" method (e.g., Jibson and Keefer 1993; Strasser et al. 2006; Yagoda-Biran et al. 2010; Wang 2020; Junquera-Torrado et al. 2021).

The magnitude-bound method uses a correlation relating earthquake magnitude to the site-to-source distance of the most distal landslide observation (typically in terms of epicentral distance). Variants of this method relate earthquake magnitude to the area affected by landslides (e.g., Tanyas et al. 2018). Developed from observations in modern earthquakes, these correlations traditionally use empirical data from variable geologic-tectonic settings. As an example, Fig. 1 presents the magnitude-bound correlation developed by Keefer (1984) using data from 40 global earthquakes that occurred AD 1811-1980. This curve can be used to estimate the lower-bound magnitude required to induce landslides at a given site-to-source distance. Subsequent authors (e.g., Rodriguez et al. 1999; Delgado et al. 2011) have generally corroborated the Keefer (1984) curves for more recent earthquakes. The shape of any such curve is a function of seismic source, path, and site effects (e.g., rupture mechanism, ground motion attenuation, and site effects), as well as landslide susceptibility (e.g., strength parameters, slope angle, and water table location). Because these factors vary spatially, region-specific correlations may provide more accurate estimates than those developed from global data (e.g., Papadopoulos and Plessa 2000; Chen et al. 2012; Martino et al. 2014; Comerci et al. 2015). As evident from Fig. 1, application of this method results in constraint of a paleoearthquake's magnitude, wherein the maximum site-to-source distance of landsliding is the only required input. Knowing or estimating the source location is thus critical, yet the magnitude-bound method provides little means to locate it. Additionally, as a lower-bound method, the actual magnitude is potentially much larger than predicted. But, despite the uncertainties inherent to magnitude-bound curves, probabilistic correlations have yet to be developed. Is it thus unclear what a "lower bound" prediction *is* in statistical terms.

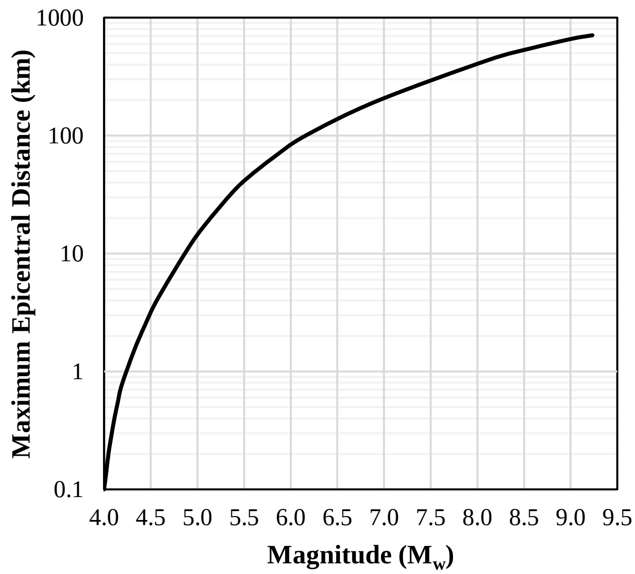


Fig. 1. Magnitude-bound curve proposed by Keefer (1984) for world-wide earthquakes, where site-to-source distance is quantified in terms of maximum epicentral distance to disrupted landslides.

The second, more technical site-specific method uses a dynamic slope-stability analysis based on some combination of geospatial, geologic, and geotechnical measurements. Several approaches of differing sophistication are available to predict dynamic slope stability, depending on the degree to which site-specific profiling and measurements are available. These include empirical (e.g., Nowicki et al. 2018), limit equilibrium (e.g., Morgenstern and Price 1965), sliding block (e.g., Jibson 1993), and non-linear dynamic models (e.g., Boulanger 2019). In a paleoseismic analysis, wherein the outcome is given (i.e., landslides did or did not occur), any of these models can be used to determine ground motions that likely would, and would not, produce the observed outcome. While implementations have varied greatly with time and place (e.g., Jibson and Keefer 1993; Strasser et al. 2006; Yagoda-Biran et al. 2010; Meunier et al. 2013; Wang 2020; Junquera-Torrado et al. 2021), the site-specific method can: (i) constrain ground-motion parameters likely to produce outcomes at individual sites; and (ii) by analyzing regional evidence, give an estimate of the causative earthquake magnitude. If, for example, a slope-stability model was adopted that characterized seismic demand in terms of earthquake magnitude (M_w) and some intensity measure (*IM*), then a limit-state condition could, for a given slope, be computed to separate combinations of M_w and *IM* sufficient to induce landsliding from combinations that are insufficient. This is shown conceptually in Fig. 2a. Because there are infinitely many such combinations, a ground-motion model (GMM) is used to determine credible combinations for a given site, as shown in Fig. 2b (dashed line), where the GMM predicts the *IM* as a function of M_w and site-to-source distance (*R*), among other factors. As indicated in Fig. 2b, the portion of this line plotting above the defined limit state corresponds to credible M_w-*IM* combinations that could induce a landslide at the site, given a rupture at some distance *R*. The intersection of the GMM with the limit state thus defines the minimum M_w expected to induce a landslide.

As evident from Fig. 2, site-to-source distance must be known to perform the site-specific method, as it presently exists. This can lead to inaccurate assumptions, and by corollary, erroneous results (e.g., where fault locations are unknown, or where prospective causative faults are plentiful). In addition, implementations of this method (e.g., Jibson and Keefer 1993; Strasser et al. 2006; Yagoda-Biran et al. 2010; Wang 2020; Junquera-Torrado et al. 2021) have ignored negative observations (i.e., slopes without landslides), the analysis of which would result in an upper-bound M_w-*IM* combination (i.e., had the seismic loading been larger, a landslide would be expected). Moreover, existing implementations have not considered the uncertainties inherent to ground-motion and landslide prediction. These implementations thus have the same shortcoming as the magnitude-bound method in that they provide a deterministic, lower-bound constraint on the causative earthquake magnitude.

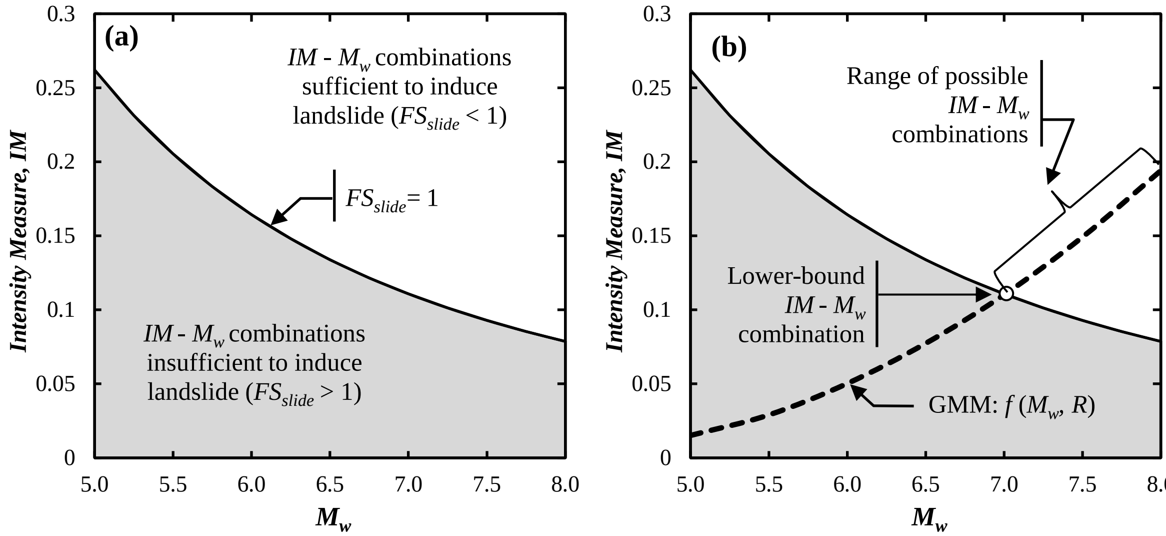


Fig. 2. (a) *IM* – M_w combinations required to induce a landslide for a hypothetical site; (b) Determination of credible lower-bound *IM* – M_w combination. GMM = ground motion model.

2. Methodology

The proposed framework is described in two parts. The first presents a conceptual overview and simple hypothetical application. The second provides implementation details to assist analysts in applying the framework, and to describe the validation tests that will be performed subsequently.

2.1 Overview of Proposed Framework

Analyzing regional paleolandslide evidence, the proposed framework aims to probabilistically constrain the causative seismic source from a near-infinite number of possibilities. This will be accomplished by assessing the likelihood that a rupture with given location, geometry, and magnitude would produce a series of field observations (sites with and without landslides) wherein uncertainties inherent to ground motion and landslide modelling may be considered. In general, the likelihood of a parameter having a value, given a set of observations, is the product of the probabilities of those observations, conditioned on the parameter value. The likelihood of a seismic source having some location (L), geometry (G), and magnitude (M_w), given a set (x) of field observations at *N* different sites, can thus be computed as:

$$Likelihood\ (L, G, M_w|x) = P(X = x|L, G, M_w) = \prod_{i=1}^{N} P(X_i = x_i|L, G, M_w) \quad (1)$$

Where $P(X_i = x_i | L, G, M_w)$ is the probability of the observation at site i (landslide or no landslide) given an earthquake with parameters L, G, and M_w. By repeating for a near-infinite number of possibilities, the most likely, actual rupture parameters are probabilistically constrained via the likelihood function (product of the probabilities of N observations), such that different combinations of L, G, and M_w will be found more and less likely to produce the observed field evidence. In Eq. (1), the probability of a field observation, conditioned on all input variables, is computed at sites with and without landslides using Eqs. 2a and 2b, respectively:

$$P(\text{Landslide}|\text{EQK}: L, G, M_w) = \int_{IM} P(\text{Landslide}|IM)\, f(IM|L, G, M_w) \cdot \mathrm{d}IM \tag{2a}$$

$$P(\text{No Landslide}|\text{EQK}: L, G, M_w) = 1 - \int_{IM} P(\text{Landslide}|IM)\, f(IM|L, G, M_w) \cdot \mathrm{d}IM \tag{2b}$$

Where $f(IM|L, G, M_w)$ is a probability density function (PDF) computed by a ground motion model (GMM) considering site response, and conditioned on fault parameters L, G and M_w; and $P(\text{Landslide}|IM)$ is the probability of observing a landslide, given ground-motion intensity measure IM, as computed by a landslide prediction model.

The proposed framework is demonstrated conceptually in Fig. 3 considering four hypothetical field sites, where a landslide is observed at site numbers one and two, but not at site numbers three and four, and two possible source locations for the earthquake that produced the observations. In actual analyses, a near-infinite number of source locations may be considered. Fig. 4 illustrates how the relative likelihoods of locations one and two are assessed. Shown in Fig. 4a are the probabilities of individual field observations, given an earthquake at location one, as computed by Eq. 2 for varying M_w. At sites with a landslide, the probability of the field observation increases with increasing M_w, whereas at sites without a landslide, the opposite occurs. In Fig. 4b, this process is repeated considering an earthquake at source location two. In Fig. 4c, the likelihood of each source location is computed as a function of M_w, per Eq. 1 (the product of the black and gray curves). It can be seen that location one has a far greater peak likelihood, whereas an earthquake at location two is unlikely to produce the field observations, regardless of its M_w. Repeating this process at a near-infinite number of locations, a geospatial surface of peak likelihood can be developed, potentially constraining the location of fault-rupture. Moreover, for any location considered, the likelihood distribution of earthquake magnitude is computed, as shown in Fig. 4c. This can be interpreted as a PDF of the causative earthquake magnitude, conditioned on a seismic-source location. By aggregating PDFs from potential source locations across the study area, an overall PDF of earthquake magnitude, considering all possible source locations, is produced. In situations where the source location is well constrained, the overall PDF is similar to the PDF conditioned on a single, most likely source location. In contrast, when the source location is less well constrained (e.g., due to limited field evidence), the overall PDF is wider than that conditioned on the single most-likely location. Collectively, these products (i.e., probability distributions of rupture location and magnitude) are inputs to regional and national probabilistic seismic hazard analyses.

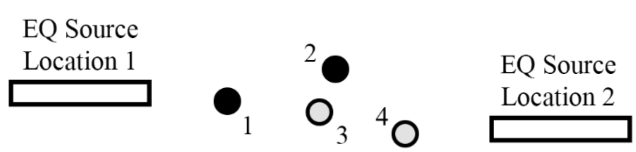


Fig. 3. Hypothetical paleolandslide analysis consisting of four field sites, where a landslide was observed at sites one and two (black circles), but not at sites three and four (gray circles). Also shown are two possible sources for the earthquake that produced these observations.

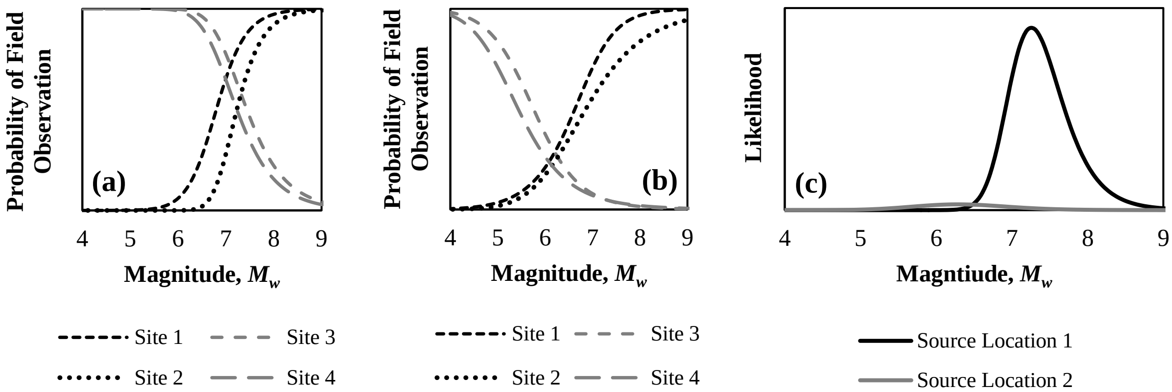


Fig. 4. Conceptual illustration of the proposed framework for locating the source of the earthquake that produced the field observations depicted in Fig. 3: (a) probabilities of field observations, given an earthquake at location one of varying M_w; (b) probabilities of field observations, given an earthquake at location two of varying M_w; (c) likelihood vs. M_w for source locations one and two.

2.2 Implementation Details

The implementation of Eqs. 1 and 2 is next discussed in detail. This includes general concepts, as well as the specific methods and models adopted herein for application to four modern earthquakes. While further details will be provided later, the framework will be applied to four events in California, USA, from which landslide inventories are available in Schmitt et al. (2017): (i) 1971 M_w6.6 San Fernando; (ii) 1983 M_w6.7 Coalinga; (iii) 1989 M_w6.9 Loma Prieta; and (iv) 1994 M_w6.7 Northridge. Owing to the modular form of the proposed framework, in which ground motions and landslides are separately predicted in series, it is easily modified for alternative and future prediction methods. In this regard, the Nowicki et al. (2018) landslide model, which uses geospatial variables, will be adopted in the current study for its ease of application. Using this model, the probability of a landslide, $P(x)$, is:

$$P(x) = \frac{1}{1+e^{-x}} \tag{3}$$

where $x = a + b \times \ln PGV + c \times Slope + d \times Lithology + e \times Land\ Cover + f \times CTI + \mathrm{g} \times \ln PGV \times Slope$

In Eq. (3), coefficients a through f are logistic regression coefficients trained on 23 landslide inventories. PGV is peak ground velocity (cm/sec); $Slope$ is the ground slope (degrees) computed from the 7.5 arc-second resolution Global Multi-resolution Terrain Elevation Data (GMTED) (Danielson and Gesch, 2011); $Lithology$ is obtained from the global lithology (GLiM) database of Hartmann and Moosdorf (2012); $Land\ Cover$ is obtained from the global land cover (GlobCover) maps of Arino et al. (2012); and CTI is the compound topographic index (a proxy for soil wetness) and is obtained from the Hydrologic Derivatives for Modeling and Analysis (HDMA) database (Verdin 2017). Per the recommendation of the U.S. Geological Survey (USGS 2021) who use the Nowicki et al. (2018) model in post-earthquake data products, we: (i) exclude areas with slopes less than 5°; and (ii) revise the lithologic 'd' coefficient for "unconsolidated sediments" from -3.22 to -1.36, which is the coefficient for "mixed sedimentary rocks" to better reflect that this unit is expected to be weak. Mapped in Fig. 5 for the Los Angeles metropolitan region, which will be studied herein, are the four "capacity" variables used by the Nowicki et al. (2018) model.

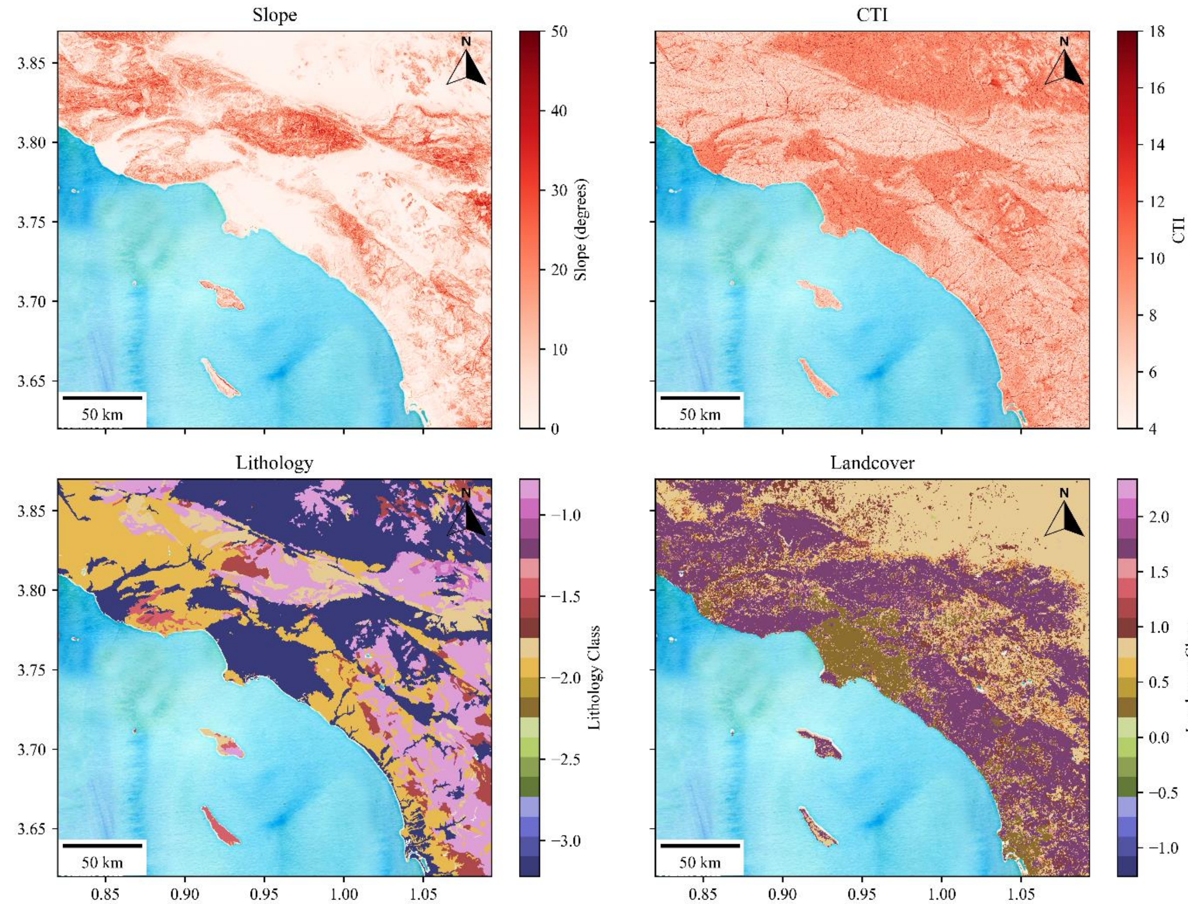


Fig. 5. Geospatial capacity variables used by the Nowicki et al. (2018) landslide model.

While empirical hazard models like Nowicki et al. (2018) have outperformed mechanistic models in some settings (Geyin et al. 2020), we generally assume that the adopted model is not the most accurate available. However, this model allows for the proposed framework to be readily implemented on multiple events, each with numerous study sites. In a paleolandslide study, wherein site- or region-specific profiling and material shear-strengths may be available, an improved empirical or mechanistic landslide model could be employed (e.g., limit equilibrium, sliding block, or nonlinear dynamic analysis). In cases where the mode of landsliding may be discerned, mode-dependent models, which potentially offer improved prediction efficiency, could be used. The basic approach, however, would not otherwise be different. To implement the framework, an analyst must:

(1) Create an array of possible source locations. In this paper, a 250 km by 250 km grid of surficial points, which will be treated as possible earthquake epicenters, was adopted and centered on the landslide field, such that the grid's borders far exceed the extent of landslides. Within the grid area (62,500 km^2 in this study), a finer point spacing gives greater spatial resolution while a coarser spacing allows for faster runtime. Balancing these ideals, a l km spacing was adopted for a 10,000 km^2 area at the center of the grid, while a 10 km spacing was adopted for the remainder of the 62,500 km^2 study area. In the event that points in the coarsely-spaced region are identified as likely sources, the grid can be recast at higher spatial resolution (i.e., to further constrain the likely source location, if possible).

(2) Select *N* study sites where the presence or absence of landslides was observed. Ideally, and as implemented in this work, an equal number of sites with and without landslides is selected to limit sampling bias.

(3) Select an appropriate GMM for the parameters required by the landslide model. In this paper, the Chiou and Youngs (2014) GMM for crustal earthquakes in active tectonic settings was adopted to predict *PGV* in all events.

(4) For each provisional seismic-source locale created in (1):

(5) For each provisional seismic-source M_w being considered (a range of M_w5 to M_w9 was used in this paper, in increments of 0.1M_w, with the M_w5 lower limit chosen following Keefer (1984)):

(6) For each of *N* study sites selected in (2), cycling from $i = 1$ to *N*:

(7) Compute the site-to-source distance(s) required by the GMM chosen in (3), as measured from study site *i* to the provisional seismic sources. In this paper, the required metrics include Joyner-Boore distance (R_{JB}) and the closest distance to fault rupture (R_{RUP}). Since the seismic sources generated in (1) are epicenters, rather than 3-dimensional faults, the correlations of Scherbaum et al. (2004) were used to estimate R_{JB} from epicentral distance (R_{EPI}), assuming a strike-slip fault with dip (δ) of 90°. The Scherbaum et al. (2004) correlations, which are magnitude dependent, effectively convert the point sources from (1) into a statistical realization of a 3-dimensional fault. R_{RUP} was calculated per the Pythagorean theorem using estimates of R_{JB} (computed as in the above) and the depth to top-of-rupture (Z_{TOR}). Z_{TOR} was estimated using the correlation of Kaklamanos et al. (2011), which requires as inputs the: down-dip rupture width (*W*); hypocentral depth (Z_{HYP}); and δ. Following the recommendation of Kaklamanos et al. (2011), *W* and Z_{HYP} were respectively estimated per the methods of Wells and Coppersmith (1994) and Scherbaum et al. (2004). It should be noted that the Scherbaum et al. (2004) R_{JB} correlation was used beyond the $5.0 < M_w < 7.5$ parameter space of the data used to train it. As a result, some adjustments to physically indefensible values were needed at very small and large M_w: (i) if $R_{JB} > R_{EPI}$, $R_{JB} = R_{EPI}$; (ii) if $R_{JB} < 0$, $R_{JB} = 0$.

(8) Using the GMM selected in (3) and site-to-source distances from (7), compute the PDF of expected *PGV* at study site *i*, wherein the time-averaged shear wave velocity over the top 30 m (V_{S30}) is consistent with site *i*. In this paper, the fault source was assumed to have strike-slip parameters ($\delta = 90°$ and rake, $\lambda = 0°$) and V_{S30} was estimated by the method of Heath et al. (2020). Topographic amplification was not considered but could be if incorporated into future GMMs. In general, the probability distribution of *PGV* is assumed to be a lognormal random variable defined by the median and lognormal standard deviation provided by the GMM. In this paper, *PGV* predictions beyond +/- 3 standard deviations of the median are truncated, as is typical in ground-motion modeling (since such values may be physically impossible) and the PDF is scaled upwards such that the area under the truncated PDF is one.

(9) For each possible *PGV* value at study site *i*, as computed in (8) for a given M_w and site-to-source distance pair (a *PGV* increment of 5 cm/s was adopted in this paper):

(10) Compute the probability of a field observation using Eq. 2a or 2b, depending on whether landslides were or were not observed at study site *i*. Completing these equations (i.e., by summing over all *PGV* values) gives the probability of field observation at site *i* for a given provisional source location and M_w. Repeating steps 6-10 for each provisional M_w being considered results in a probability of field observation versus M_w curve for each study site, examples of which are shown in Figs. 4a and 4b.

(11) Compute the likelihood of a provisional source locale, as a function of M_w, by multiplying the probabilities of field observations at all study sites (i.e., multiply the curves shown in Fig. 4a or 4b at each value of M_w). The result, an example of which is shown in Fig. 4c, is akin to a PDF of the causative earthquake magnitude, conditioned on a single seismic-source location.

(12) Repeating steps 5-11 for all provisional seismic source locations created in (1) results in a likelihood distribution for each, as described in (11).

(13) Following from (12), normalize the peak likelihood at each source location by the peak likelihood among all locations. This allows for the relative likelihoods of source locations to be assessed, with the most likely location having a normalized value of one. The M_w corresponding to the peak likelihood (see Fig. 4c) is the median M_w likely to produce the field observations, given that an earthquake at the source location occurs.

(14) From the arrays of normalized likelihood and median M_w, create geospatial surfaces or contours to identify the most-likely source location and median M_w, conditioned on that location.

(15) Finally, by aggregating PDFs from all potential source locations across the study area, an overall PDF of earthquake magnitude, considering all possible source locations, is produced. While a single source location will always be deemed "most likely," earthquakes at multiple other source locations have potential to produce the field of evidence,

albeit with lesser likelihood. In general, this overall PDF will be wider (i.e., have greater uncertainty) than the PDF conditioned on the single most likely source, especially when the source zone is not well constrained (e.g., due to limited field evidence).

3. Simulated Application

The proposed framework, as outlined above, provides probabilistic constraint on both the location and M_w of a seismic event considering landslide evidence. To investigate the degree of successful constraint and glean lessons for future use, simulated paleolandslide studies are next performed on four modern events with known seismic sources. As in any paleolandslide analysis (and paleoseismic research in general), the need exists, where judicious, to classify sites as "positive" and "negative" (i.e., landslides are, or are not, present). Inherent to the subsequent applications, it is thus assumed that the presence or absence of landslides would be accurately interpreted in a paleoseismic investigation, and moreover, that landslides would be properly attributed to a causative earthquake. These assumptions are not intended to diminish the importance of interpreting and dating earthquake-induced landslides in the field, which is both difficult and uncertain. The objective herein is to advance seismic-source inversion techniques, rather than to discuss or simulate the field investigations prerequisite for all such analyses.

In each event, the inverted, probabilistic source location and magnitude will be compared to the actual, known parameters. However, no prior knowledge of the seismic source is incorporated into the analyses. In addition, because the four earthquakes are modern, the number of study sites available from reconnaissance (i.e., slopes with and without observed landslides) is very large (e.g., 10,000+ in some events). Given that this far exceeds the number of field sites conceivable of a regional paleolandslide study, analyses will be performed on randomly selected samples of 100 study sites from each event (50 slopes with observed landslides, 50 slopes without). Because modern landslide inventories are comprised almost exclusively of positive observations, we assume slopes within a given survey area did not experience slides if none are mapped. While this inevitably introduces uncertainty, it facilitates rapid implementation of the proposed framework and is not dissimilar from a paleoseismic study, wherein an absence of evidence may not be evidence of absence. The preservation and sampling of evidence is further discussed later in the paper.

3.1 San Fernando, California

The 1971 M_w6.6 San Fernando, California earthquake occurred on the Sierra Madre Fault (Carena and Suppe 2002). An inventory of observed landslides was obtained from Morton (1971), as compiled by Schmitt et al. (2017), and is mapped in Fig. 6a. A representative example of the framework's output is in Fig. 6 for an analysis using 100 randomly selected study sites; the surface projection of the 1971 fault rupture is also shown. Mapped in Fig. 6a are contours of relative likelihood, wherein values near one and zero respectively denote source locations most and least likely to produce the observed evidence. Mapped in Fig. 6b are contours of M_w, indicating the rupture magnitude most likely to produce the observed evidence, given an earthquake at that location (i.e., the median M_w, conditioned on location).

As seen in Fig. 6a, the analysis identified a relatively well-constrained area as the likely source. This aligns with the actual rupture, such that the identified, most likely source falls within the surface projection of the fault that ruptured. In addition to this fault, other known faults in the region are mapped, such that their computed potential to produce the evidence field may be assessed. It can be seen that other potential seismic sources in the region are all judged unlikely to produce the observed evidence. As shown in Fig. 6b, the median magnitude corresponding to the most-likely source location was M_w6.6, matching the actual magnitude. Because M_w contours tend to map as concentric circles centered on the evidence field, as in Fig. 6b, these contours are provided in Appendix A for the remainder of events. In addition to the median M_w of M_w6.6, other M_w values at the most-likely source location could inevitably produce the evidence field (e.g., if ground motions were higher or lower than the median expectation). Shown in Fig. 7a is the full PDF of M_w, conditioned on the most likely source location, which is analogous to the conceptual PDFs shown in Fig. 4c. The 95% confidence interval (CI) is M_w6.25- M_w6.80. Moreover, as shown in Fig. 6a, earthquakes elsewhere have potential to produce the evidence field, albeit with lesser likelihood. By aggregating individual PDFs from all locations across the 62,500 km^2 study grid, an overall PDF that includes source-location uncertainty is produced. This PDF, which is also shown in Fig. 7a, has a median of M_w6.6 and 95% CI of M_w6.25-M_w6.90. Unless the causative fault can be constrained via on-fault (or some other) evidence, this latter, more uncertain PDF would be the more appropriate product of a paleolandslide analysis.

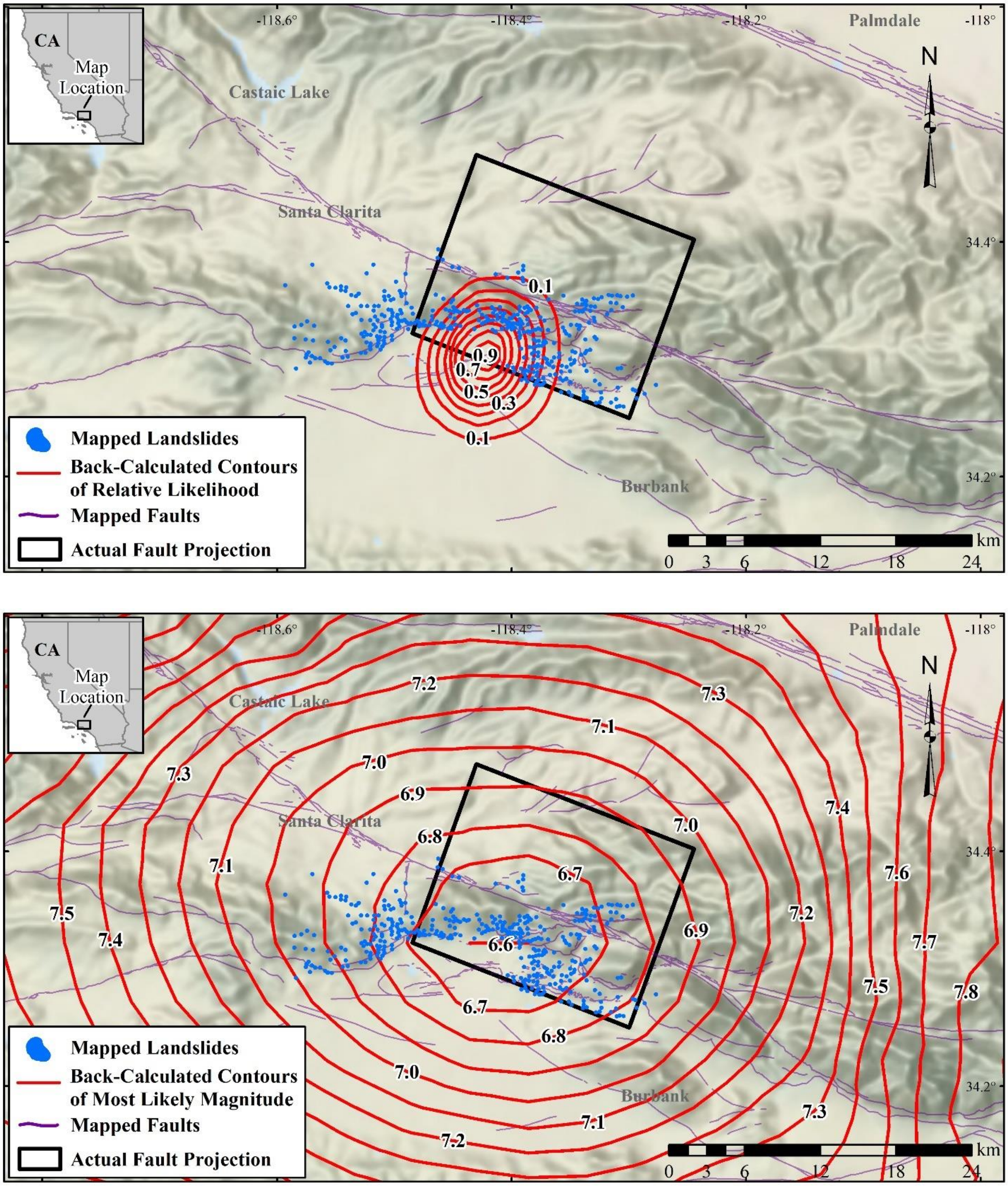


Fig. 6. Representative analysis of 100 study sites (50% positive, 50% negative) randomly selected from the 1971 M_w6.6 San Fernando, California earthquake inventory: **(a)** contours showing the relative likelihood of earthquake

source locations; and **(b)** corresponding M_w most likely to produce the field evidence, if the earthquake occurred at a given location. Black rectangles = actual fault projection (Carena and Suppe 2002); faint purple lines = other mapped faults in the region (California Geological Survey 2014); blue polygons = complete inventory of mapped landslides.

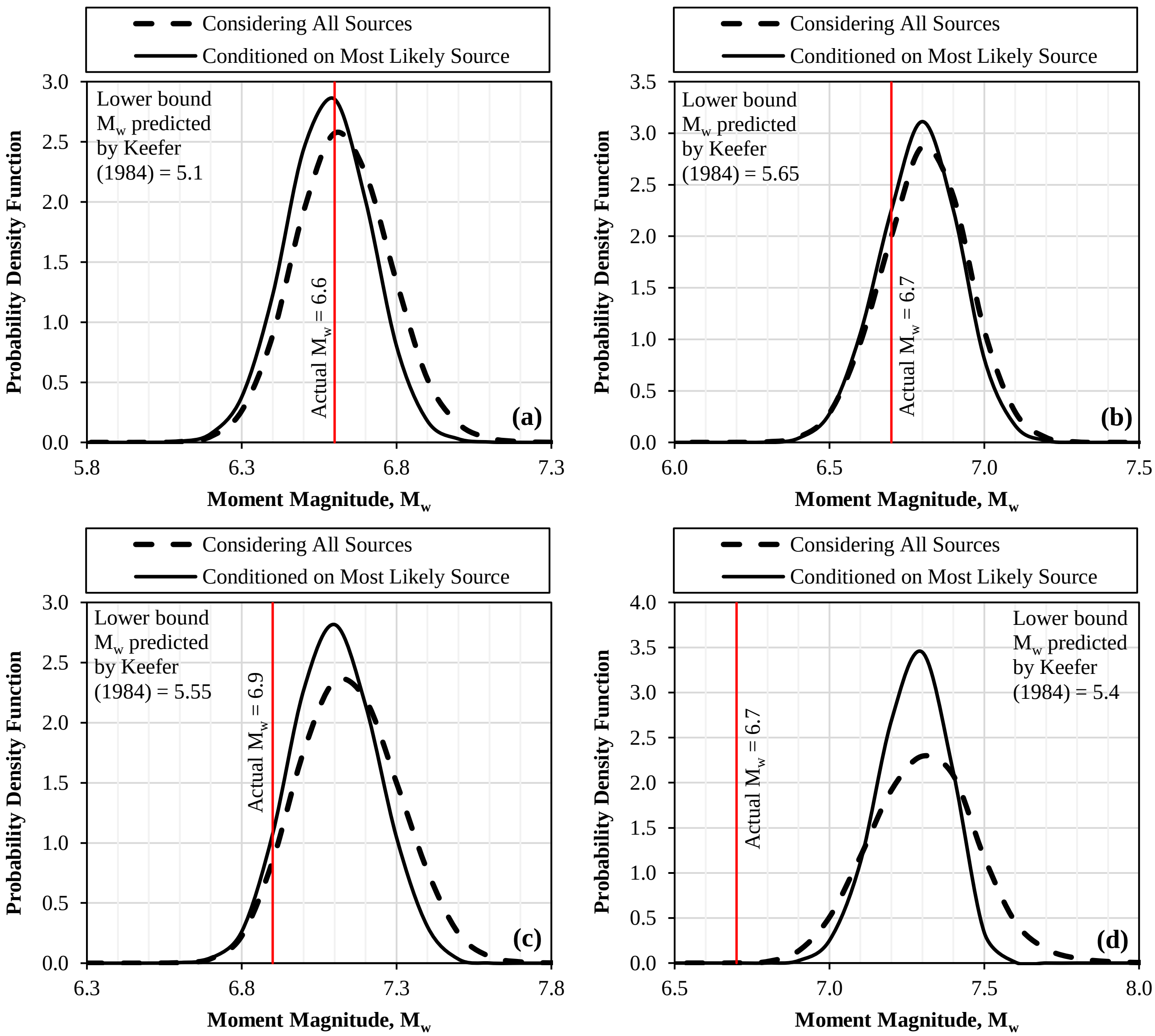


Fig. 7. PDFs of earthquake magnitude inverted from regional landslide evidence in the: (a) 1971 M_w6.6 San Fernando; (b) 1994 M_w6.7 Northridge; (c) 1989 M_w6.9 Loma Prieta; and (d) 1983 M_w6.7 Coalinga earthquakes. Shown are the PDFs conditioned on the most-likely source location (solid lines) as well as the aggregate PDFs considering source-location uncertainty (dashed lines).

For comparison, the magnitude-bound curve of Keefer (1984) (later corroborated by others) was used to predict the magnitude of the event, as has been employed in paleoseismic studies. To do so, the centroid of all mapped landslides was assumed to be the source location, from which the site-to-source distance of the most distal landslide was measured to obtain a minimum magnitude of M_w5.1 using Fig. 1. Applications of the "site-specific" approach, wherein a minimum ground-motion for sliding is established, would similarly result in a lower-bound magnitude. By contrast, the proposed framework allows for all uncertainties to be considered, resulting in a more meaningful, probabilistic understanding of source location and magnitude. In general, as the number of field sites increases, the uncertainty bounds on location and magnitude decrease, and vice-a-versa.

3.2 Northridge

The 1994 M_w6.7 Northridge, California earthquake occurred on a deeply buried blind thrust fault (Updike et al. 1996) in proximity to the 1971 San Fernando event. Landslide inventories were obtained from Harp and Jibson (1995) and Townsend et al. (2020), as compiled by Schmitt et al. (2017), and are mapped in Fig. 8. A representative output is shown in Fig. 8 for an analysis using 100 random study sites; the surface projection of the 1994 fault rupture is also shown. Mapped in Fig. 8 are contours denoting source locations most and least likely to produce the observed evidence. The analysis again identified a relatively well-constrained area as the likely source, ~12 km NW of that predicted for the 1971 San Fernando event. It can also be seen that the source with greatest likelihood is ~5 km from the actual fault projection, but that locations within the projection do have up to 40% relative likelihood (i.e., they are not unlikely). The median M_w (See Appendix A) at the most-likely source was M_w6.8 with a 95% CI of M_w6.5- M_w7.0, versus the actual M_w6.7. As with the San Fernando earthquake, this PDF - conditioned on the most-likely source location - and the overall PDF reflecting source-location uncertainty, are computed and shown in Fig. 7b. In this analysis, the two PDFs provide essentially the same M_w distribution. For comparison, the Keefer (1984) magnitude-bound curve was used to predict the magnitude as previously described. This resulted in a lower-bound estimate of M_w5.65. Thus, the proposed framework provides results that are informative, even if imperfect, especially when considering that an empirical landslide model using few site-specific parameters is being employed.

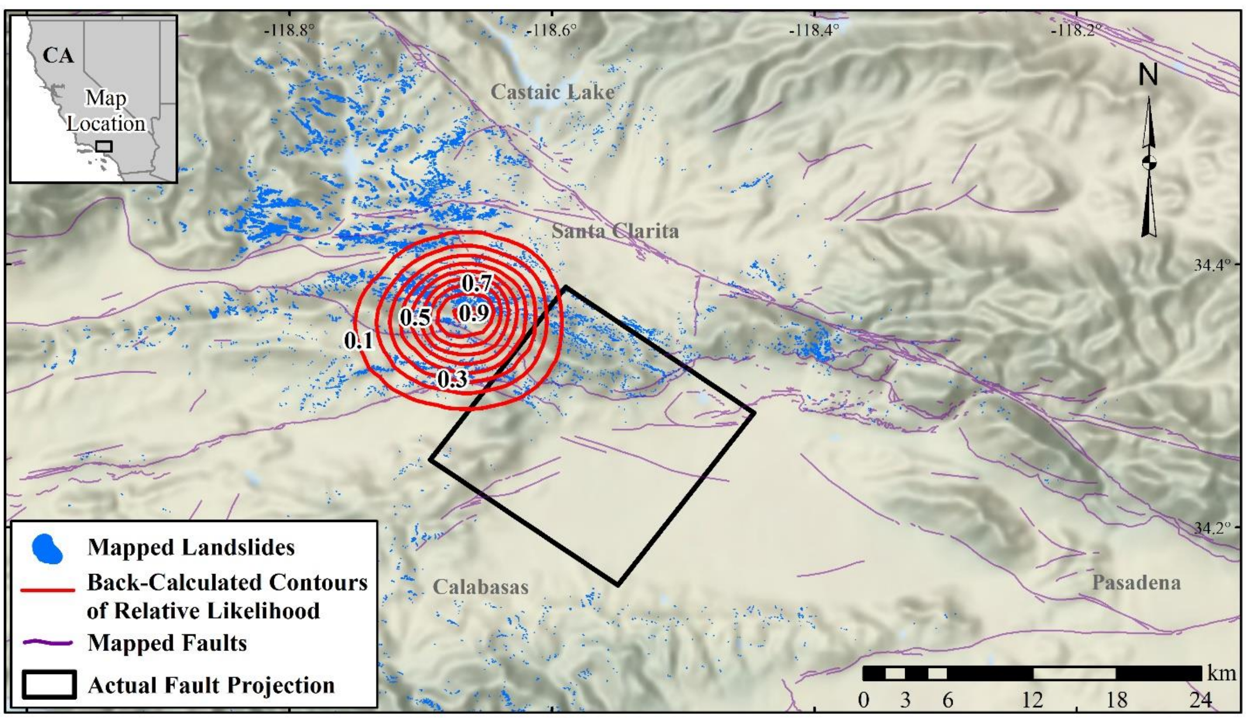


Fig. 8. Representative analysis of 100 study sites (50% positive, 50% negative) randomly selected from the 1994 M_w6.7 Northridge, California earthquake inventory: contours showing the relative likelihood of earthquake source locations. Black rectangles = actual fault projection (Updike et al. 1996); faint purple lines = other mapped faults in the region (California Geological Survey 2014); blue polygons = complete inventory of mapped landslides.

3.3 Loma Prieta

The 1989 M_w6.9 Loma Prieta earthquake occurred on the San Andreas Fault (Wald et al. 1991). An inventory of observed landslides was obtained from Keefer and Manson (1998), as compiled by Schmitt et al. (2017), and is mapped in Fig. 9. The results of a representative analysis of 100 study sites are also shown. The inversion identified a most-likely source approximately ~2 km from the actual fault projection, with some locations in the projection having up to 90% relative likelihood. The corresponding median magnitude was M_w7.1 (Appendix A), with a 95% CI of M_w6.8-M_w7.3, as shown in Fig. 7c. Aggregating conditional PDFs from across the study area resulted in an overall PDF with median M_w7.1 and 95% CI of M_w6.8-M_w7.4, versus the actual M_w6.9. By contrast, the Keefer (1984) magnitude-bound curve gives a lower-bound estimate of M_w 5.5. Thus, similar to the 1994 Northridge earthquake, the actual source parameters (i.e., location and magnitude) fall within the inverted distributions, even if not precisely at the medians.

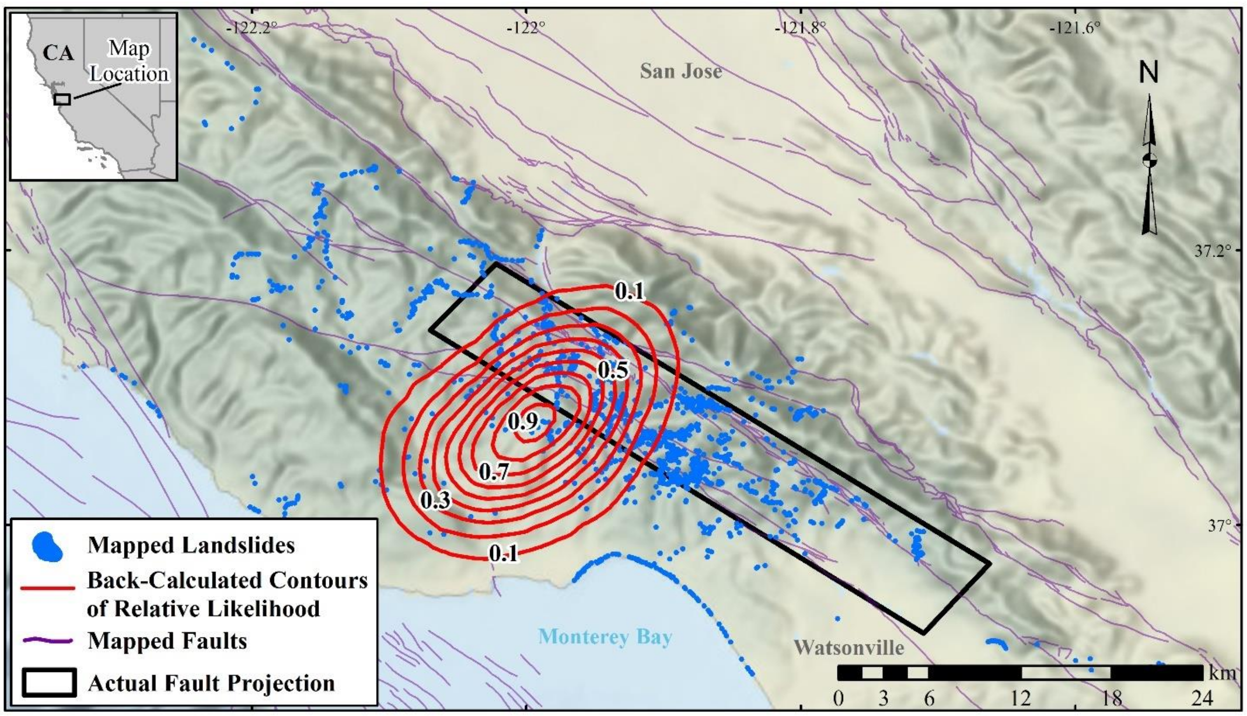


Fig. 9. Representative analysis of 100 study sites randomly selected from the 1989 M_w6.9 Loma Prieta earthquake inventory: contours showing the relative likelihood of source locations. Black rectangles = actual fault projection (Wald et al. 1991); faint purple lines = other mapped faults in the region (California Geological Survey 2014); blue polygons = complete inventory of mapped landslides.

3.4 Coalinga, California

The 1983 M_w6.7 Coalinga earthquake occurred on a previously unknown fault along the structural boundary of the San Joaquin Valley (Wentworth and Zoback 1990). An inventory of observed landslides was obtained from Harp and Keefer (1990), as compiled by Schmitt et al. (2017). The result of a representative analysis of 100 study sites is shown in Fig. 10. The inversion accurately located the source, with the identified, most-likely grid point falling within the actual fault projection. However, the median inverted M_w here (See Appendix A) was M_w7.3 with 95% CI of M_w7.0-M_w7.4, as shown in Fig. 7d. In contrast, the actual magnitude, M_w6.7, was deemed to have near-zero probability of producing the evidence field. Aggregating conditional PDFs resulted in a similar overall PDF with median M_w7.3 and 95% CI of M_w6.9- M_w7.6, versus the actual M_w6.7. Thus, the inversion framework significantly overpredicted the rupture magnitude, despite accurately predicting its location.

This discrepancy could have one or more causes. First, while epicentral grid points were converted to fault realizations using site-to-source distance correlations, these realizations may differ from the true source (e.g., with respect to fault orientation or depth). This will be further discussed later in the paper. Second, it can be seen in Fig. 10 that the fault ruptured beneath flat ground (i.e., on the edge of the San Joaquin Valley) and that this terrain extends a great distance east of the rupture location. As a result, very few landslides were observed at relatively short site-to-source distances, and none were observed east of the fault rupture. Analogous to the geolocation of earthquakes using seismic-wave arrivals, which relies on distributed seismic instruments, the proposed framework relies on distributed study sites (i.e., slopes with and without landslides) to "record" the ground motion. Thus, in events such as this, where the evidence field is constrained by topography and/or far from the source, it may be more difficult to detect the direction and rate of ground-motion attenuation, and therefore more difficult to constrain the source parameters. Third, the inversion framework relies on accurate predictions of ground-motions and landslides. Discrepancies between true and inverted seismic parameters are likely whenever these phenomena differ from expectations. An overpredicted magnitude could result, for example, if regional ground motions are greater than those predicted by the GMM, given the event magnitude (i.e., due to source, path, or site effects like topographic amplification). This could also result from landslides occurring more readily than expected, given the predicted ground motions. This is certainly plausible, given that the Nowicki et al. (2018) model does not consider local or regional variations in slope stratigraphy and material shear strength. Nonetheless, had the analysis been based on the Keefer (1984) magnitude-bound curve, a lower-bound estimate of M_w5.4 would result, providing an equally poor constraint.

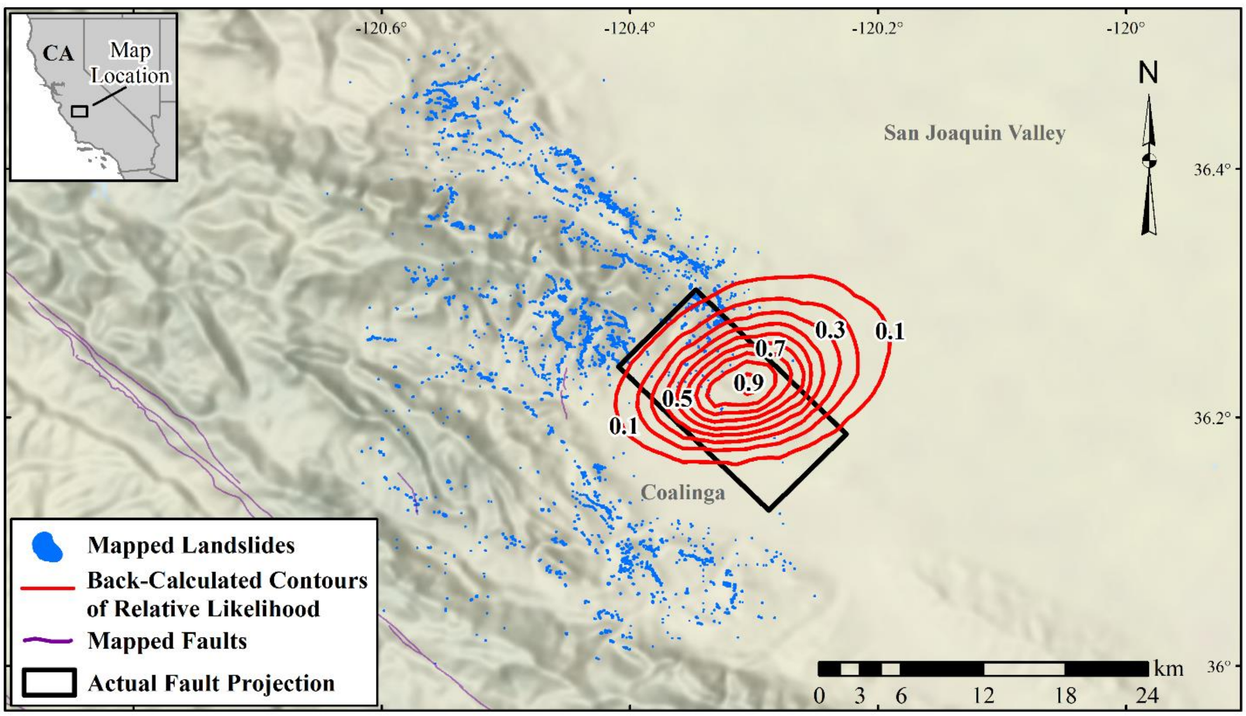


Fig. 10. Representative analysis of 100 study sites randomly selected from the 1983 M_w6.7 Coalinga earthquake inventory: contours showing the relative likelihood of source locations. Black rectangles = actual fault projection (Wentworth and Zoback 1990); faint purple lines = other mapped faults in the region (California Geological Survey 2014); blue polygons = complete inventory of mapped landslides.

4. Discussion and Conclusions

In regions where the modern earthquake catalog is poorly populated, analyses of paleolandslides may be used to infer information about the seismic hazard. However, existing analysis techniques: (i) require assumptions about the causative earthquake's location, which can lead to inaccurate results; (ii) generally use only "positive" observations, leading to lower-bound estimates of M_w; and (iii) have to-date been deterministic in nature, whereas paleolandslides involve multiple uncertainties. Accordingly, this paper proposed an approach by which seismic-source parameters are probabilistically constrained from regional landslide evidence, resulting in PDFs of rupture location and magnitude. Simulated paleoseismic studies were performed on four modern earthquakes in California, USA, wherein the analyses were blind to the actual source model (i.e., magnitude, location, geometry). While source parameters were accurately constrained in the 1971 San Fernando, 1989 Loma Prieta, and 1994 Northridge earthquakes, the analysis of the 1983 Coalinga earthquake highlighted possible limitations and lessons for use.

In these studies, provisional earthquake sources were modeled as epicenters and then converted to statistical fault realizations using empirical site-to-source distance correlations. For simplicity, and assuming the fault mechanism would be unknown in a paleoseismic study, an archetypical strike-slip fault was assumed, irrespective of the actual mechanism. However, because regional ground-motion patterns may be mechanism-dependent, the distribution of landslides can also vary, all else being equal (e.g., Tatard and Grasso 2013; Marc et al. 2017). The influence of fault mechanism might therefore be considered by repeating with other assumptions (i.e., normal, reverse) and coalescing the results via a logic tree. The branch weightings could be equal or informed by knowledge of the seismological setting (e.g., if one fault mechanism is believed to be more common). Alternatively, and while computationally demanding, the plausibility of 3-D source models could be directly evaluated with minimal modification. That is, the framework could quantify the likelihoods of faults with different strike, depth, mechanism, etc. producing an evidence field, so long as these faults result in different ground motions at landslide study sites. In this regard, because the framework relies on ground motions differing at different sites, it is generally best practice to compile spatially distributed study sites, given that multiple sites in proximity may provide little or no additional information. Given that the framework relies on predicting multiple phenomena in series, inaccuracies are likely whenever one or more of these phenomena differs from expectations. In the present study, the four application earthquakes were in California, USA, and were of relatively similar magnitude. Accordingly, these tests do not span the parameter space of geologic, seismologic, and climatic settings found on earth, nor do they include especially large- or small-magnitude events. However, because the framework is modular in nature, it can readily accommodate different ground-motion and landslide prediction models (e.g., region-specific models or models of differing complexity, depending on the availability of site-specific information). Above all, the framework remains conceptually valid for all earthquakes in all locations, so long as ground motions and landslides are predictable.

Given its modularity, the framework will also likely improve over time, since better component models are continuously proposed. In this regard, the Nowicki et al. (2018) landslide model was adopted herein for its ease of implementation to multiple earthquakes. In a paleolandslide study, wherein site- or region-specific profiling and material shear-strengths may be available, a more advanced model would potentially improve performance. Finally, the inversion framework, as demonstrated in this paper, did not simulate the field interpretation, dating, and inventorying prerequisite for inverse-analysis. As an example, and as in any paleolandslide study, the need exists to classify field observations as "positive" or "negative" (i.e., landslides did, or did not, occur). In practice, this presents a challenge, given the uncertain spatiotemporal potential for landslide evidence to be preserved in the geologic record. Inevitably, nearly all evidence from a given event is eventually erased. In such cases, the erroneous classification of a "positive" site as "negative" could result in the inverted, causative earthquake magnitude being less than actual. Similarly, a field of landslides interpreted to have resulted from one event could be augmented by additional spatially distributed earthquakes, or by weather events spaced closely in time. In such cases, the erroneous classification of a "negative" site as "positive" could result in the inverted, causative earthquake magnitude being greater than actual. As a result, the difficulty of accurate interpretation and analysis is likely to increase for increasingly older events. Additionally, the demonstrations presented herein used 100 randomly selected study sites, given the premise that a larger quantity would be unlikely in a paleoseismic study. It stands to reason, however, that: (i) analyses using a larger quantity of data are likely to be more accurate than those using a smaller quantity of data; and (ii) larger earthquakes impacting a larger geographic area are likely to require more data for accurate inversion, as compared to smaller earthquakes impacting a smaller area. To evaluate the finite-sample uncertainty of results (i.e., the uncertainty owed to incomplete data), the inversion can be repeated multiple times using bootstrap sampling, resulting in a distribution of PDFs. Inherently, this distribution would be wide when a small number of sites are studied (i.e., a different outcome could result from studying a different small number of sites) and narrow as the number of sites increases (i.e., the outcome would be relatively stable, regardless of which sites are studied). In this regard, uncertainties relating to field interpretation, evidence preservation, data selection, and temporal changes over long timescales were not present in the simulated studies of four modern earthquakes.

However, these applications did demonstrate how significant, quantifiable uncertainties can be accounted for via the total probability theorem, resulting in a probabilistic understanding of rupture location and magnitude, as inverted from landslide evidence. While paleolandslides have been studied and reported widely, less effort has arguably focused on advancing the methods for inverting seismic sources from field evidence. Towards that end, the approach proposed herein has the potential to provide new insights in regions of persistent seismic-hazard uncertainty.

5. Data Availability

All data analyzed herein is available in the public domain.

6. Supplemental Material

Appendix A, containing contour maps of most-likely M_w for each event, is available online.

7. Funding

The presented study is based on work supported by the National Science Foundation (NSF) under Grant No. CMMI-1751216, by the NSF Graduate Research Fellowship Program under Grant No. DGE-1762114, and by the University of Washington Royalty Research Fund (RRF). However, any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF or RRF.

8. References

Arino O, Ramos P, Jose J, Kalogirou V, Bontemps S, Defourny P, Van Bogaert E (2012) Global Land Cover Map for 2009 (GlobCover 2009). © European Space Agency (ESA) & Université catholique de Louvain (UCL), PANGAEA, https://doi.org/10.1594/PANGAEA.787668

Atwater BF, Stuiver M, and Yamaguchi DK (1991) Radiocarbon test of earthquake magnitude at the Cascadia subduction zone. Nature 353: 156-158.

Boulanger RW (2019) Nonlinear dynamic analyses of Austrian dam in the 1989 Loma Prieta earthquake. Journal of Geotechnical and Geoenvironmental Engineering 145(11): 05019011

Bull WB, King J, Kong F, Moutoux T, Phillips WM (1994). Lichen dating of coseismic landslide hazards in Alpine mountains. Geomorphology 10:253–264. https://doi.org/10.1016/0169-555X(94)90020-5

California Geological Survey (2014) Fault Activity Map of California. Scale not given. Accessed April 9, 2021, at: https://maps.conservation.ca.gov/cgs/fam/

Carena S, Suppe J (2002). Three-dimensional imaging of active structures using earthquake aftershocks: the Northridge thrust, California. Journal of Structural Geology 24:887-904

Chen XL, Zhou Q, Ran H, and Dong R (2012). Earthquake-triggered landslides in southwest China, Nat. Hazards Earth Syst. Sci., 12: 351-363, https://doi.org/10.5194/nhess-12-351-2012.

Chiou BS-J and Youngs RR (2014) Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. Earthquake Spectra 30(3): 1117-1153.

Clague J (2014) Paleolandslides. In: Davies T and Shroder F (Eds.) Landslide Hazards, Risks, and Disasters. Elsevier, pp 321-344

Comerci V, Vittori E, Blumetti AM, Brustia E, Di Manna P, Gerrieri L, Lucarini M, and Serva, L. (2015) Environmental effects of the December 28, 1908, Southern Calabria-Messina (Southern Italy) earthquake. Nat Hazards 76: 1849-1891. https://doi.org/10.1007/s11069-014-1573-x.

Crozier MJ (1992) Determination of paleoseismicity from landslides. In: Bell DH (ed) Landslides (Glissements de terrain), International Symposium, 6th, Christchurch, New Zealand, 1992, Proceedings: Rotterdam, A.A. Balkema, 2:1173-1180
Danielson JJ, Gesch DB (2011) Global multi-resolution terrain elevation data 2010 (GMTED2010) U.S. Geological Survey Open-File Report 2011–1073, 26 p
Delgado J, Garrido J, Lopez-Casado C, Martino S, Pelaez JA. (2011) On far field occurrence of seismically induced landslides. Engineering Geology 123:204-213. https://doi.org/10.1016/j.enggeo.2011.08.002
Engelhart SE, Horton BP, Nelson AR, Hawkes AD, Witter RC, Wang K, Wang PL, and Vane CH (2013) Testing the use of microfossils to reconstruct great earthquakes at Cascadia. Geology 41 (10): 1067–1070.
Geyin M, Baird AJ, Maurer BW (2020) Field assessment of liquefaction prediction models based on geotechnical versus geospatial data, with lessons for each. Earthquake Spectra 36(3):1386-1411.
Goldfinger C, Nelson CH, Morey AE, Johnson JE, Patton JR, Karabanov E, Gutiérrez-Pastor J, Eriksson AT, Gràcia E, Dunhill G, Enkin RJ, Dallimore A, and Vallier T (2012) Turbidite event history - methods and implications for Holocene paleoseismicity of the Cascadia subduction zone. U.S. Geological Survey Paper 1661–F, 170 pg.
Harp EL, Jibson RW (1995) Inventory of landslides triggered by the 1994 Northridge, California earthquake. U.S. Geological Survey Open-File Report 95-213, 17 p., 2 pl. scale 1:50,000 and 1:100,000, available at: https://pubs.usgs.gov/of/1995/ofr-95-0213/
Harp EL, Keefer DK (1990) Landslides triggered by the earthquake, in Rymer MJ, Ellsworth WL (eds.) The Coalinga, California, Earthquake of May 2, 1983, U.S. Geological Survey Professional Paper 1487, p. 335–347, 1 pl., scale 1:48,000, available at: https://pubs.usgs.gov/pp/1487/report.pdf
Hartmann J, Moosdorf N (2012) The new global lithological map database GLiM: A representation of rock properties at the Earth surface. Geochemistry, Geophysics, Geosystems, 13, Q12004, https://doi.org/10.1029/2012GC004370
Heath DC, Wald DJ, Worden CB, Thompson EM, Smoczyk GM (2020) A global hybrid VS 30 map with a topographic slope–based default and regional map insets. Earthquake Spectra 36 (3): 1570-1584.
Hupp CR, Osterkamp WR and Thornton JL (1987) Dendrogeomorphic evidence and dating of recent debris flows on Mount Shasta, northern California. US Geol. Surv. Professional Paper 1396-B, 39 pp
Jibson RW (1993) Predicting earthquake-induced landslide displacements using Newmark's sliding block analysis. Transportation research record 1411:9-17
Jibson RW, Keefer DK (1993) Analysis of the seismic origin of landslides: Examples from the New Madrid seismic zone. Geol. Soc. Am. Bull. 105:421-436
Jibson RW (1996) Use of landslides for paleoseismic analysis. Engineering Geology 43:291-323
Johnson RH (1987) Dating of ancient, deep-seated landslides in temperate regions. In: MG Anderson, KS Richards (Eds.) Slope Stability. John Wiley and Sons, New York, pp 561-600
Junquera-Torrado S, Moreiras SM, Rodriguez-Peces MJ, Sepulveda SA (2021) Linking earthquake-triggered paleolandslides to their seismic source and to the possible seismic event that originated them in a portion of the Argentine Precordillera (31°–33°S). Natural Hazards 106:43-78. https://doi.org/10.1007/s11069-020-04447-1
Kaklamanos J, Baise LG, Boore DM (2011) Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice. Earthquake Spectra 27(4):1219-1235
Katz O, Amit R, Yagoda-Biran G, Hatzor YH, Porat N, Medvedev B (2011) Quaternary earthquakes and landslides in the Sea of Galilee area, the Dead Sea Transform: paleoseismic analysis and implication to the current hazard. Isr. J. Earth Sci. 58:275–294. doi:10.1560/IJES.58.3-4.275
Keefer, DK (1984) Landslides caused by earthquakes. Geol. Soc. Am. Bull. 95:406-421
Keefer DK, Manson MW (1998) Regional distribution and characteristics of landslides generated by the earthquake, in Keefer DK (ed) The Loma Prieta, California, Earthquake of October 17, 1989—Landslides, U.S. Geological Survey Professional Paper, 1551-C, p. C7-C32, 4 pl., available at: https://pubs.usgs.gov/pp/pp1551/pp1551c/.
Le Roux O, Schwartz S, Gamond JF, Jongmans D, Bourles D, Braucher R, Mahaney W, Carcaillet J, Leanni L (2009) CRE dating on the head scarp of a major landslide (Se´chilienne, French Alps), age constraints on Holocene kinematics. Earth and Planetary Science Letters 280:239–245. https://doi.org/10.1016/j.epsl.2009.01.034
Mackey BH, Quigley MC (2014) Strong proximal earthquakes revealed by cosmogenic 3He dating of prehistoric rockfalls, Christchurch, New Zealand. Geology 42(11):975-978
Martino S, Prestininzi A, and Romeo RW (2014) Earthquake-induced ground failures in Italy from a reviewed database. Nat. Hazards Earth Syst. Sci., 14: 799-814.
Maurer BW, Green RA, Quigley MC, and Bastin S (2015) Development of magnitude-bound relations for paleoliquefaction analyses: New Zealand case study. Engineering Geology 197: 253-266.
Meunier P, Uchida T, and Hovius N (2013) Landslide patterns reveal the sources of large earthquakes. Earth and Planetary Science Letters 36: 27-33.

Morgenstern NU, Price VE (1965) The analysis of the stability of general slip surfaces. Geotechnique, 15(1):79-93
Morton DM (1971) Seismically triggered landslides in the area above the San Fernando Valley, in The San Fernando, California, Earthquake of February 9, 1971, U.S. Geological Survey Professional Paper 733, p. 99-109
Nowicki Jessee MA, Hamburger MW, Allstadt K, Wald DJ, Robeson SM, Tanyas H, Hearne M, Thompson EM (2018) A Global Empirical Model for Near-Real-Time Assessment of Seismically Induced Landslides. Journal of Geophysical Research: Earth Surface 123(8):1835-1859. https://doi.org/10.1029/2017JF004494
O'Donnell III RJ, Hawkes AD, Lane CS, Engelhard SE, Horton BP, Bobrowsky P, Sawai Y, Witter RC, Nelson AR, and Tanigawa K (2017) Assessing the utility of δ13C and bulk geochemistry in estuaries along the Cascadia subduction zone for coastal paleoseismology. Geological Society of America Abstracts with Programs 49(6): 58-2.
Ojala AEK, Markovaara-Koivisto M, Middleton M, Ruskeeniemi T, Mattila J, Sutinen R (2018) Dating of paleolandslides in western Finnish Lapland. Earth Surf. Process. Landforms 43: 2449–2462.
Panek T (2015) Recent progress in landslide dating: A global overview. Progress in Physical Geography: Earth and Environment 39(2):168-198. https://doi.org/10.1177/0309133314550671
Papadopoulos GA, Plessa A (2000) Magnitude-distance relations for earthquake–induced landslides in Greece. Engineering Geology 58:377–386
Peters R, Jaffe B, and Gelfenbaum G (2007) Distribution and sedimentary characteristics of tsunami deposits along the Cascadia margin of western North America. Sed Geo 200: 372-386.
Petersen MD, Frankel AD, Harmsen SC, Mueller CS, Haller KM, Wheeler RL, Wesson RL, Zeng Y, Boyd OS, Perkins DM, Luco N, Field EH, Wills CJ, Rukstales KS (2008) Documentation for the 2008 update of the United States national seismic hazard maps. USGS Open-File Report 2008-1128, 61 p
Petersen MD, Moschetti MP, Powers PM, Mueller CS, Haller KM, Frankel AD, Zeng Y, Rezaeian S, Harmsen SC, Boyd OS, Field N, Chen R, Rukstales KS, Luco N, Wheeler RL, Williams RA and Olsen AH (2014) Documentation for the 2014 update of the United States national seismic hazard maps. USGS Open-File Report 2014–1091, 243 p
Rasanen RA, Marafi NA, and Maurer BW (2021) Compilation and forecasting of paleoliquefaction evidence for the strength of ground motions in the US Pacific Northwest. Engineering Geology 292: 106253.
Rodríguez CE, Bommer JJ, Chandler RJ (1999) Earthquake-induced landslides 1980–1997. Soil Dynamics and Earthquake Engineering 18:325–346
Scherbaum F, Schmedes J, Cotton F (2004) On the Conversion of Source-to-Site Distance Measures for Extended Earthquake Source Models. Bulletin of the Seismological Society of America 94(3):1053-1069
Schmitt RG, Tanyas H, Nowicki Jessee MA, Zhu J, Biegel KM, Allstadt KE, Jibson RW, Thompson EM, van Westen CJ, Sato HP, Wald DJ, Godt JW, Gorum T, Xu C, Rathje EM, Knudsen KL (2017) An Open Repository of Earthquake-triggered Ground Failure Inventories. U.S. Geological Survey data release collection, accessed March 1, 2021, at https://doi.org/10.5066/F7H70DB4
Stout ML (1977) Radiocarbon dating of landslides in southern California. California Division of Mines and Geology, California Geology, May 1977, pp 99-105
Strasser M, Anselmetti FS, Fah D, Giardini D, Schnellmann M (2006) Magnitudes and source areas of large prehistoric northern Alpine earthquakes revealed by slope failures in lakes. Geology 34(12):1005-1008. https://doi.org/10.1130/G22784A.1
Struble WT, Roering JJ, Black BA, Burns WJ, Calhoun N, Wetherell L (2020) Dendrochronological dating of landslides in western Oregon: Searching for signals of the Cascadia A.D. 1700 earthquake. GSA Bulletin 132(7-8): 1775–1791. https://doi.org/10.1130/B35269.1
Sutinen R, Hyvonen E, Kukkonen, I (2014) LiDAR detection of paleolandslides in the vicinity of the Suasselka postglacial fault, Finnish Lapland. International Journal of Applied Earth Observation and Geoinformation 27:91–99. https://doi.org/10.1016/j.jag.2013.05.004
Tanyaş, H, Allstadt KE, and van Westen CJ (2018) An updated method for estimating landslide-event magnitude. Earth Surf. Process. and forms 43: 1836- 1847. https://doi.org/10.1002/esp.4359.
Townsend KF, Gallen SF, Clark MK (2020) Quantifying near-surface rock strength on a regional scale from hillslope stability models: JGR Earth Surface 125(7). https://doi.org/10.1029/2020JF005665
Updike RG, Brown WM, Johnson ML, Omdahl EM, Powers PS, Rhea S, Tarr AC (1996) USGS Response to an Urban Earthquake, Northridge '94. U.S. Geological Survey Open-File Report 96-263, 78 p
USGS (2021) Ground Failure Scientific Background https://earthquake.usgs.gov/data/ground-failure/background.php. Accessed 01 March 2021
Verdin KL (2017) Hydrologic Derivatives for Modeling and Applications (HDMA) database: U.S. Geological Survey data release, https://doi.org/10.5066/F7S180ZP

Vidale J, Atkinson G, Green R, Hetland E, Grant-Ludwig L, Mazzotti S, Nishenko S and Sykes L (2011) Report of the independent expert panel on New Madrid Seismic Zone earthquake hazards as approved by NEPEC on April 16, 2011. U.S. Geological Survey, 26 p

Wald DJ, Helmberger DV, Heaton TH (1991) Rupture Model of the 1989 Loma Prieta Earthquake from the Inversion of Strong-Motion and Broadband Teleseismic Data. Bulletin of the Seismological Society of America 81(5):1540-1572

Wang B (2020) Geotechnical investigations of an earthquake that triggered disastrous landslides in eastern Canada about 1020 cal BP. Geoenvironmental Disasters (2020) 7:21. https://doi.org/10.1186/s40677-020-00157-9

Watt JT, Brothers DA, Bennett SEK, Kluesner JW, Roland E, Conrad JE, Sliter RW, Dartnell P, Goldfinger C, Patton J, Michalak MJ, Sherrod BL, Gomberg J, and Wells RE (2017). "Toward a systematic characterization of Cascadia upper plate morphology, structure, and Quaternary deformation history: an integrated onshore-offshore approach." GSA Abstracts 49(6): 58.

Wells DL, Coppersmith KJ (1994) New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, And Surface Displacement. Bulletin of the Seismological Society of America 84(4):974-1002

Wentworth CM, Zoback MD (1990) Structure of the Coalinga Area and Thrust Origin of the Earthquake. In: Rymer MJ, Ellsworth WL (eds.) The Coalinga, California, Earthquake of May 2, 1983, U.S. Geological Survey Professional Paper 1487, p 417

Yagoda-Biran G, Hatzor YH, Amit R, Katz O (2010) Constraining regional paleo peak ground acceleration from back analysis of prehistoric landslides: Example from Sea of Galilee, Dead Sea transform. Tectonophysics 490:81-92. https://doi.org/10.1016/j.tecto.2010.04.029

APPENDIX D

Publication 4 (Journal Paper 3), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.

[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating

Rasanen, R.A.[a] and Maurer, B.W.[a,*]

ABSTRACT

The source of the 1886 Charleston, South Carolina earthquake influences the computed seismic hazard of the Southeastern U.S. and thus impacts public policy and engineering practice. However, because the 1886 rupture predated seismic instruments, its source is highly uncertain. This study presents probabilistic seismic-source inversions of the Charleston earthquake from liquefaction evidence and historical intensity reports. Using the latest predictive models and a novel inversion approach, we seek to constrain the magnitude, location, and orientation of the 1886 rupture. Probability distributions of rupture magnitude are conditioned on both the "Woodstock Fault" – a commonly inferred source of the 1886 event – and on an unknown source, wherein the uncertainties of fault location and orientation are considered. These distributions are compared to the M_w6.7-M_w7.5 distribution adopted by the U.S. National Seismic Hazard Model Project (NSHMP). Collectively, the results do not provide strong support for the hypothesized Woodstock Fault. This is not to say the Woodstock Fault does not exist or could not have caused the 1886 rupture, but rather, that the position of the 1886 source model cannot be constrained by the data and models studied herein, given the large uncertainties inherent to each. While this is at odds with the underlying assumption of prior studies, the results nonetheless generally uphold the magnitude distribution adopted by the NSHMP. The largest uncertainties inherent to this distribution are identified and could be diminished in the future. Finally, we note that the inversion methodology used here is not specific to any region, or to certain types of evidence, but can be applied to any seismic zone and to any co-seismic response. This methodology allows for uncertainty to be accounted for in a more complete and transparent manner when inverting seismic source parameters from macroseismic data. Of course, any limitations, biases, or unmodeled uncertainties inherent to these data must be understood, and their implications acknowledged, as further discussed herein.

[a] Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington 98195
*Corresponding author: bwmaurer@uw.edu

Keywords: seismic hazard, macroseismic intensity, soil liquefaction, inverse analysis

1. Introduction

Computed seismic hazards are especially uncertain in regions of infrequent seismicity, where the return periods of moderate-to-large earthquakes may exceed the historic observational period. To reduce this uncertainty, engineering geologists routinely perform forensic analyses of the macroseismic evidence (e.g., liquefaction, landslides, intensity reports) produced by prehistoric and pre-instrumental earthquakes (e.g., among many, Obermeier, 1998; Schneider et al., 2001; Green et al., 2005; Kuhn, 2005; Obermeier et al., 2005; Olson et al., 2005; Rodriguez-Marek and Ciani, 2008; Maurer et al., 2015a; Gheibi and Gassman, 2016; Yousuf et al., 2021; Chung et al., 2021; Rasanen et al., 2021; Bwambale et al., 2022). The goal of these analyses, in effect, is to constrain the seismic-source parameters of paleoearthquakes, such that these parameters may be input to probabilistic seismic hazard analyses. It follows that computed seismic hazards are, in some regions, heavily influenced by analyses of macroseismic evidence. The South Carolina Coastal Plain is one such region and is strongly influenced by interpretations of the 1886 Charleston earthquake.

The 1886 event induced widespread soil liquefaction across the Coastal Plain (Amick et al., 1990), damaged structures in multiple U.S. states, including most structures in Charleston (Dutton, 1889; Wong et al., 2005), produced perceptible shaking over 1500 km away in Canada (Bakun et al., 2002), and was larger in magnitude than any earthquake to since occur in the Southeastern U.S. The source of the Charleston event is thus a major seismic hazard for the region. A 2005 study, for example, predicted that a repeat of the 1886 event would cause 900 deaths, 44,000 injuries, and economic losses of $20 billion in South Carolina alone (Wong et al., 2005). In turn, the Charleston source controls the computed seismic hazard for much of the Southeastern U.S., particularly for long-period structures (Petersen et al., 2020), and thus impacts building codes, governing policies, and engineering practice. However, because the 1886 rupture predated seismic instruments and did not manifest at the surface, its exact location and magnitude remain uncertain, as do the regional amplitudes of resultant ground motions. To constrain these unknowns, and thus prepare for a similar event, numerous researchers have studied macroseismic evidence, as summarized in Table 1. Published confidence intervals (CIs) of the 1886 magnitude range from M_w6.4 to M_w7.8, as interpreted from intensity reports (e.g., Bakun and Hopper, 2004; Cramer and Boyd, 2014) and liquefaction evidence (e.g., Martin and Clough, 1994; Hayati and Andrus, 2008). Liquefaction features also suggest a history of recurrent earthquakes in the region extending back 6,000 years (Talwani and Schaeffer, 2001) with wide-ranging magnitude estimates of M_w5.1 to M_w7.8 (e.g., Hu et al., 2002; Gheibi et al., 2020). Considering the existing literature, Petersen et al. (2014, 2020) assigned to the Charleston seismic zone a magnitude probability distribution that ranged from M_w6.7 to M_w7.5 in the most recent U.S. National Seismic Hazard Model Project (NSHMP) maps.

While much has been learned about the 1886 Charleston earthquake, prior analyses of the macroseismic data (i.e., intensity reports, soil liquefaction) have several limitations. First, the analysis of this data has multiple uncertainties, yet existing studies tend either to be deterministic or to account for uncertainties informally. That is, they generally provide either a median estimate of the rupture magnitude or uncertainty bounds that are nominal in nature. It is often unclear what the bounds are, exactly, and which uncertainties are, and are not, accounted for. Second, most analyses assume that the 1886 event was caused by a particular fault (i.e., the "Woodstock Fault") with known characteristics, even though the fault(s) responsible for the event are debated and the characteristics of the Woodstock Fault are uncertain. The feasibility of the data to constrain the source model has arguably not been fully explored, given that nearly all studies provide a magnitude estimate conditioned on a single fault and do not investigate the uncertainty of this assumption. Third, the inverse analysis of intensity and liquefaction data requires a series of models for predicting these phenomena. Regionally distributed ground motions must be predicted, conditioned on a hypothetical source, to include site-response effects at the locations of study. The probability of field observations (i.e., the observed intensity or liquefaction response) must then be computed, conditioned on the expected ground motions. In this regard, major modeling advances have recently been made. The NGA-East project (Goulet et al., 2018) resulted in the most advanced understanding of Eastern North America (ENA) ground motions and site response (Harmon et al., 2019) to date. Models for correlating ground motions to macroseismic intensities, including ENA-specific relationships, have been updated (e.g., Cramer, 2020). And models for predicting the probability of liquefaction surface expression have been trained using all liquefaction case histories globally compiled to date (Geyin and Maurer, 2020).

Table 1. Prior estimates of the 1886 Charleston earthquake magnitude; estimates are in moment magnitude (M_w) and ranges are at the 95% confidence level, unless noted otherwise (m_b = body wave magnitude; M_s = surface wave magnitude).

Study	Study Type	Magnitude
Bollinger (1977)	MMI	6.8-7.1 (m_b)*
Nuttli et al. (1986)	MMI	6.7 (m_b), 7.7 (M_s)†
Martin and Clough (1994)	Liquefaction	7.0-7.5†
Johnston (1996)	MMI	6.8-7.8
Bakun and Hopper (2004)	MMI	6.4-7.2
Heidari and Andrus (2010)	Liquefaction	6.8-7.0†
Cramer and Boyd (2014)	MMI	6.7-7.3

* Upper bound magnitude estimate (range is not at the 95% confidence level)

† Magnitude range is either not given or is not at the 95% confidence level

Accordingly, this study presents probabilistic seismic-source inversions of the 1886 Charleston earthquake from historical intensity reports and liquefaction evidence. Each is studied using a novel approach wherein the above shortcomings are addressed directly. With this approach, the likelihood of a rupture with some location, geometry, and magnitude to produce a set of field observations (observed intensities or liquefaction responses) is computed. Repeating for enumerable hypothetical faults results in a regional scale understanding of the likely source parameters, to the degree the observational data permits. Probability distributions of earthquake magnitude, conditioned on both an unknown source and on the Woodstock Fault, are computed and compared to that used to develop the NSHMP maps (Petersen et al., 2014, 2020). In the following, prior studies of the 1886 macroseismic data (see Table 1) are discussed further. An overview of the analysis methodology is then presented, followed by implementation details. Lastly, the macroseismic data are analyzed and a variety of results are presented and discussed.

2. Prior analyses of 1886 macroseismic evidence

2.1 Analyses of MMI data

Following the 1886 earthquake, Dutton (1889) compiled intensity reports throughout ENA and developed isoseismal maps based on the Rossi-Forel intensity scale. Researchers have since reinterpreted these reports to the Modified Mercalli Intensity (MMI) scale and analyzed them to infer seismic parameters. Bakun et al. (2002), for example, compiled 1,034 MMI observations from these and other original reports. Prior analyses of the 1886 MMI data have typically used intensity-prediction equations (IPEs), which predict intensity as a function of rupture magnitude and site-to-source distance. Assuming some source location and adopting an IPE, researchers have constrained the causative M_w which best fits the MMI data (e.g., Bollinger, 1977; Nuttli et al., 1986; Johnston, 1996; Bakun and Hopper, 2004) as summarized in Table 1. Bakun and Hopper (2004), for example, developed an ENA-specific IPE, applied it to the Bakun et al. (2002) MMI data, and reported a magnitude of M_w6.9 (M_w6.4-7.2 at the 95% confidence level). Using different methods than prior researchers, Cramer and Boyd (2014) compared the mean MMI of the Bakun et al. (2002) dataset against those from two reference events in similar tectonic settings (M_w7.2 1929 Grand Banks, Canada and M_w7.6 2001 Bhuj, India) over a site-to-source distance of 600-1200 km. With this approach, Cramer and Boyd (2014) estimated a median magnitude of M_w7.0 with uncertainty of $\pm$ 0.3M_w. Collectively, existing studies of the MMI data have reported estimates of M_w6.4 to M_w7.8. In producing such estimates, these studies have generally assumed that the source was epicentrally located in the vicinity of what is typically called the Woodstock Fault, an inferred N striking, W dipping fault ~25 km NW of Charleston (e.g., Durá-Gómez and Talwani, 2009a,b; Chapman et al., 2016). The fault's surface projection, as hypothesized by Durá-Gómez and Talwani (2009a,b), is mapped in Figure 1. While most studies based on geophysical investigations or modern seismological data have supported this proposed alignment – at least in a general sense (e.g., Pratt et al., 2022) – dramatically different hypotheses for the 1886 earthquake have also been proposed. Marple and Hurd (2020), for example, recently suggested that the 40-km long "Deer Park lineament," which is oriented roughly E-W, may have been responsible.

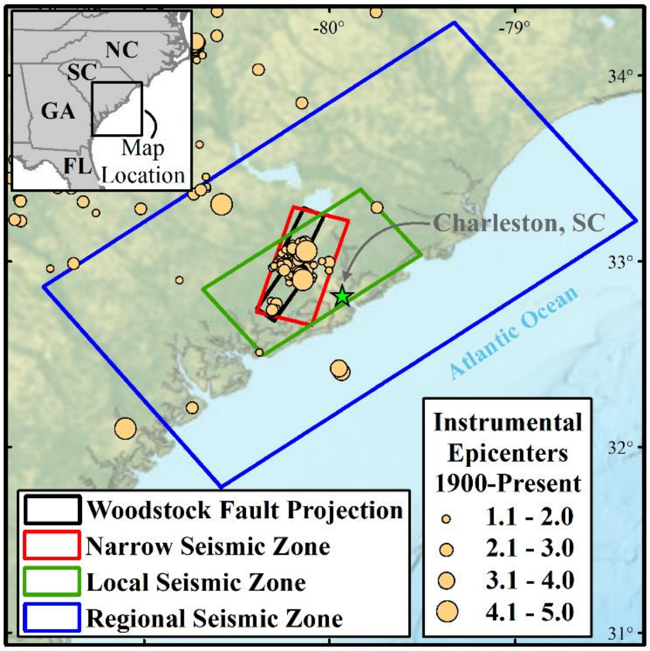


Figure 1. Woodstock Fault projection (Durá-Gómez and Talwani, 2009a,b) and the zonal weighting scheme assigned to the Charleston Seismic Zone in the most recent U.S. National Seismic Hazard Model Project (NSHMP) maps (Petersen et al., 2014, 2020).

Also shown in Figure 1 is the zonal weighting scheme assigned to the Charleston Seismic Zone by the NSHMP (Petersen et al., 2014, 2020). In this scheme, a "Narrow" zone with weight of 0.3 delineates the hypothesized Woodstock Fault while accounting for uncertainties in its position and branching structure. The "Local" and "Regional" zones, with respective weights of 0.5 and 0.2, collectively extend offshore and across the extents of the South Carolina Coastal Plain, thereby encompassing more distal faults and liquefaction features that have not been tied to the 1886 event. Each zone is assigned the same M_w6.7 to M_w7.5 probability distribution by Petersen et al. (2014, 2020). This weighting scheme was adopted from the Central and Eastern United States Seismic Source Characterization for Nuclear Facilities Project (Coppersmith et al., 2012), who concluded: "Neither the 1886 nor the prehistoric (i.e., pre-1886) earthquakes in the Charleston area can be definitively attributed to any specific fault or fault zone at the present time." In arriving at this conclusion, Coppersmith et al. (2012) noted: "the Charleston region is associated with a pattern of observed seismicity that is not particularly remarkable for drawing attention to the location of the 1886 earthquake." As summarized by Chapman et al. (2016), very different hypotheses have arisen from analyses of microseismicity data. While it should be emphasized that the NSHMP scheme does not describe the uncertainty of the 1886 rupture specifically, it does reflect the overall uncertainty of moderate-to-large "1886-like" events in the region. Thus, while prior studies have generally assumed that the 1886 source was in the "Narrow" zone, a scientific consensus has not been reached.

As a precursor to other analyses presented herein, the Bakun and Hopper (2004) IPE was first reimplemented on the Bakun et al. (2002) dataset and the same M_w6.9 estimate as Bakun and Hopper (2004) was obtained, indicating that the approach and dataset were correctly reproduced. Next, this approach was updated using the newest ENA IPE (Atkinson et al., 2014). Assuming the same source location and studying MMI data within 1000 km (the applicable distance of the Atkinson et al. (2014) IPE) and within all distances (to mirror Bakun and Hopper, 2004), median estimates of M_w8.0 and M_w8.2 were respectively obtained. The cause of the discrepant results obtained using Bakun and Hopper (2004) vs. Atkinson et al. (2014) can be seen in Figure 2, where both IPEs are plotted for three values of M_w. For a given M_w and epicentral distance, the Atkinson et al. (2014) model tends to predict a lesser MMI, indicating that a larger earthquake magnitude (i.e., M_w8.0 - M_w8.2) is needed to produce the same set of MMI observations.

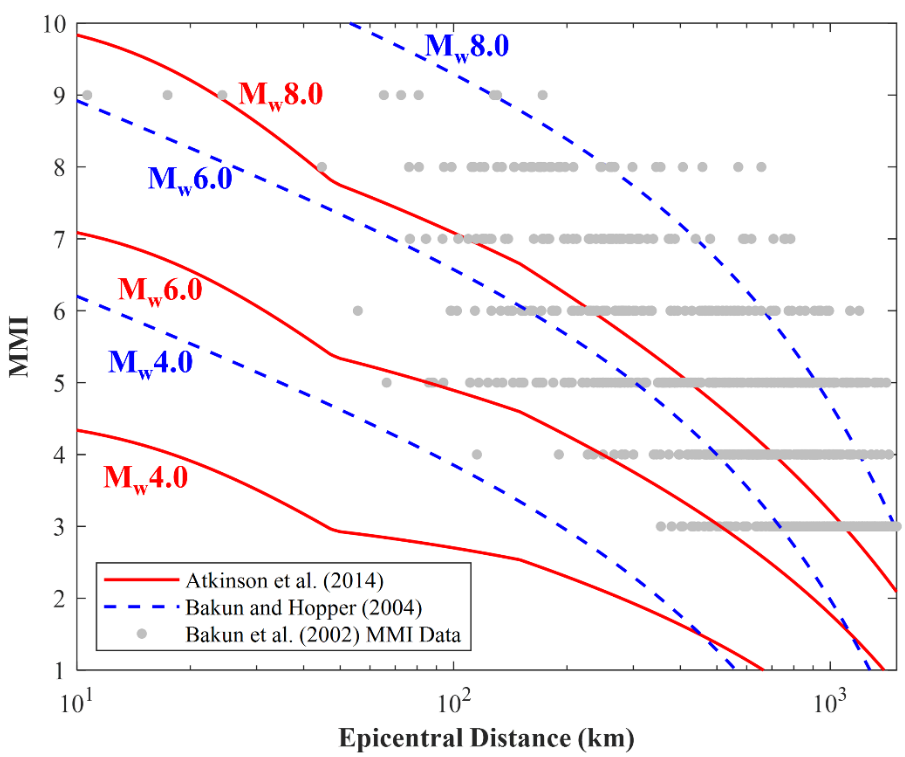


Figure 2. ENA-specific intensity prediction equations (IPEs) proposed by Bakun and Hopper (2004) and Atkinson et al. (2014) considering three values of M_w.

While it may appear, per the latest ENA IPE, that the 1886 rupture was much larger than previously thought, there are limitations that give rise to the work that follows. Most notably, perhaps, is that existing ENA IPEs do not allow for consideration of site effects when predicting MMI. The Atkinson et al. (2014) IPE, for example, is intended for site class C, or stiff soil and soft rock profiles (i.e., sites where V_{S30}, the time averaged shear-wave velocity over the upper 30 m, is 360 m/s < V_{S30} < 760 m/s). Accordingly, if some of the 1886 MMI observations were on softer sites (e.g., in river valleys or along waterways, which seems likely), then the estimate of M_w8.0-M_w8.2 obtained via the Atkinson et al. (2014) IPE could require significant reduction. Moreover, the MMI studies in Table 1: (i) predate the latest knowledge of ENA ground motions and site response (Goulet et al., 2018; Harmon et al., 2019); and (ii) do not rigorously account for uncertainty. In this study, probabilistic site-adjusted ground motion intensity measures (IMs) will be explicitly predicted using 17 ENA ground-motion models (GMMs) (whereas ground motions were only implicitly predicted in prior studies). In turn, MMI values will be probabilistically predicted at study sites using the latest IM-MMI models (e.g., Cramer, 2020). These predictions will be repeated for a multitude of fault locations, orientations, and magnitudes to compute the likelihood that each source would produce the 1886 MMI observations of Bakun et al. (2002). This approach, which will be subsequently presented in detail, incorporates ground-motion IM uncertainty, IM-MMI uncertainty, and source location uncertainty to produce a probability distribution of rupture magnitude. It must be noted that MMI data is also subject to measurement uncertainty and reporting bias (e.g., Hough et al., 2000; Cramer and Boyd, 2014). While site-specific measurement uncertainties, correction factors, or weighting schemes could be accommodated, a thorough reinterpretation of the more than 1000 original intensity reports would be required, to include possible reassignment of MMI values in the Bakun et al. (2002) dataset and development of observation-specific uncertainties. In the current effort, however, we study the existing data directly, treat measurement uncertainty in a simple manner, and assign all observations equal weight. It should be noted that all prior studies of the 1886 intensities have also used these data. Thus, while our methodology has important advantages over prior efforts (e.g., the capacity to probabilistically constrain the rupture location), a future study might further benefit from rigorous reinterpretation of the original intensity reports.

2.2 Analyses using liquefaction

Paleoliquefaction evidence suggests that at least seven moderate-to-large earthquakes have impacted the South Carolina Coastal Plain in the last 6,000 years (e.g., Gohn et al., 1984; Obermeier et al., 1985; Obermeier et al., 1987; Talwani and Schaeffer, 2001). Three of these events are interpreted to have a source in the vicinity of Charleston and a recurrence rate of ~500 years (Talwani and Schaeffer, 2001). The 1886 event was the most recent of these to generate liquefaction and the only instance in which liquefaction was well documented as it occurred. Surface manifestations (e.g., ejecta, ground cracks) were mapped by Earle Sloan, among others, and compiled by Dutton (1889). Additional liquefaction evidence was subsequently discovered during trenching investigations (e.g., Obermeier et al., 1985; Talwani and Cox, 1985). Collectively, this evidence has been analyzed to determine the magnitude of the earthquake that caused it. Liquefaction models (e.g., Green et al., 2019; Maurer et al., 2015b) conventionally predict the future triggering and surface manifestation of liquefaction, given in-situ geotechnical test data and some seismic loading. In an inverse analysis, these models are used in reverse to constrain the seismic loading that would, and would not, produce the observed response. By comparing this loading to that forward predicted by a GMM for an assumed source, the magnitude of that source may be constrained. Studying primarily standard penetration test (SPT) data from sites of interest, Martin and Clough (1994) carried out such an analysis with the Seed et al. (1984) SPT-based liquefaction triggering model and the Ishihara (1985) liquefaction manifestation model. Assuming a seismic source at the centroid of reported intensity (roughly consistent with the hypothesized Woodstock Fault), adopting GMMs then available (e.g., Chapman et al., 1989), and employing considerable judgement, Martin and Clough (1994) estimated that an M_w7.0-M_w7.5 event could produce liquefaction consistent with that observed.

Studying cone penetration test (CPT) data, Hayati and Andrus (2008) used the Robertson and Wride (1998) CPT-based triggering model and the Iwasaki et al. (1978) manifestation model to estimate a magnitude of M_w6.8-M_w7.3. Because triggering models are trained almost exclusively with case-history data from Holocene deposits, “aging correction” factors (or “deposit resistance” corrections) have been proposed when applying triggering models in older soils. Specifically, it has been argued that aging effects, or increases in the cyclic strength of soils over time, may be resolved into gains measurable by large strain penetration tests and gains influenced by soil fabric phenomena undetectable at larger strain (e.g., Maurer et al., 2014). Thus, penetration resistance may correlate to liquefaction resistance differently in Pleistocene soils than in Holocene soils. Accordingly, Hayati and Andrus (2008) employed aging correction factors in select geologic units. In producing their estimate of M_w6.8-M_w7.3, Hayati and Andrus (2008) assumed that the Woodstock fault was the source and that a peak ground acceleration (*PGA*) of 0.3 g occurred throughout Charleston, citing previous ground-motion predictions (e.g., Silva et al., 2003). Using this methodology, they constrained the magnitude range for which predictions from liquefaction models matched observations of response. Heidari and Andrus (2010) used a similar methodology as Hayati and Andrus (2008) but applied the updated aging correction factors of Hayati and Andrus (2009) to obtain an estimate of M_w6.8-M_w7.0. In addition to these studies, researchers have studied paleoliquefaction interpreted to be from older, possibly similar events. Gheibi et al. (2020), for example, studied evidence induced by a "Charleston Source" approximately ~550 and ~5,000 YBP. Assuming the source to be the Woodstock fault, they computed respective minimum magnitudes of M_w6.6-M_w7.2 and M_w6.2-M_w6.7 for these two events, where the uncertainty stems from which GMM is adopted to predict median ground motions.

Like prior studies of the MMI data, those of liquefaction evidence have greatly improved knowledge of the regional seismic hazard, but also have limitations that motivate the present study. In brief, existing studies: (i) predate both the NGA East Project (Goulet et al., 2018; Harmon et al., 2019) and the latest liquefaction models trained on all globally available data (Geyin and Maurer, 2020); and (ii) do not rigorously account for uncertainty. In this regard, prior studies do not account for source-model uncertainty (i.e., they assume a single seismic source), do not account for the uncertainty of ground motions conditioned on that source, and do consider the prediction of liquefaction in any probabilistic sense. In general, published uncertainties of the 1886 rupture magnitude, whether derived from MMI or liquefaction data, are arguably nominal in nature. In some studies, for example, only a single uncertainty is considered, such as the epistemic uncertainty of which deterministic model is used (say, to compute ground motions). In such cases, published uncertainty bounds (e.g., M_w6.8-M_w7.0) are ranges of the estimated *median* magnitude considering one source of uncertainty. This may be distinctly different from the total uncertainty of the 1886 magnitude, which could be much greater. Consider, for example, that *instrumental* magnitudes have 95% CIs exceeding some of the ranges in Table 1 (e.g., Werner and Sornette, 2008).

In this study, and analogous to the analysis of MMI data, probabilistic site-adjusted ground motion IMs will be predicted by 17 ENA GMMs. Conditioned on these IMs, the probability of liquefaction manifestation will be computed by the fragility functions of Geyin and Maurer (2020), wherein multiple models for soil aging effects will be ensembled. These predictions will be repeated for a multitude of hypothetical sources to compute the likelihood that each would produce the observed regional liquefaction response. Like the study of MMI data, this will result in a probability distribution of earthquake magnitude and, arguably, a more complete and transparent understanding of the 1886 source model, at least insofar as can be gained from the macroseismic evidence available for analysis.

3. Macrosesimic data

3.1 Analyses using liquefaction

Bakun et al. (2002) compiled 1,034 intensity reports from the 1886 event, including those of Dutton (1889) and Bollinger and Stover (1976), and assigned MMI values per the: (i) MMI definitions of Wood and Neumann (1931); and (ii) USGS National Earthquake Information Center practice for assigning intensity (Stover and Coffman, 1993). In compiling these data, Bakun et al. (2002) either excluded MMI = 1 and MMI = 2 reports (i.e., documented instances where shaking was not felt) or found no such reports to compile. Like all previous studies of these data, we assume that MMI values were accurately assigned by Bakun et al. (2002), with proper consideration of structural vulnerability to avoid biasing estimates. While matters of MMI uncertainty and bias could be important and will be further discussed, they are largely beyond the scope of the present analysis. It should be emphasized that our analysis is not an endorsement of the Bakun et al. (2002) dataset. Rather, we study this data because it is the most recent compilation of 1886 intensity reports and because many prior publications studied either the Bakun et al. (2002) compilation, or earlier collections of data that Bakun et al. (2002) built upon. As previously stated, a future study could further benefit from a complete reinterpretation of the more than 1000 original intensity reports. While the coordinates of one MMI report from Ottawa, Ontario were judged to be erroneous and corrected by judgement, we otherwise adopt the Bakun et al. (2002) MMI data as presented therein. The locations of these data are shown in Figure 3 and extend ~1500 km from Charleston, SC. A larger, alternative version of Fig. 3 that provides an MMI raster surface is also provided in the electronic supplement.

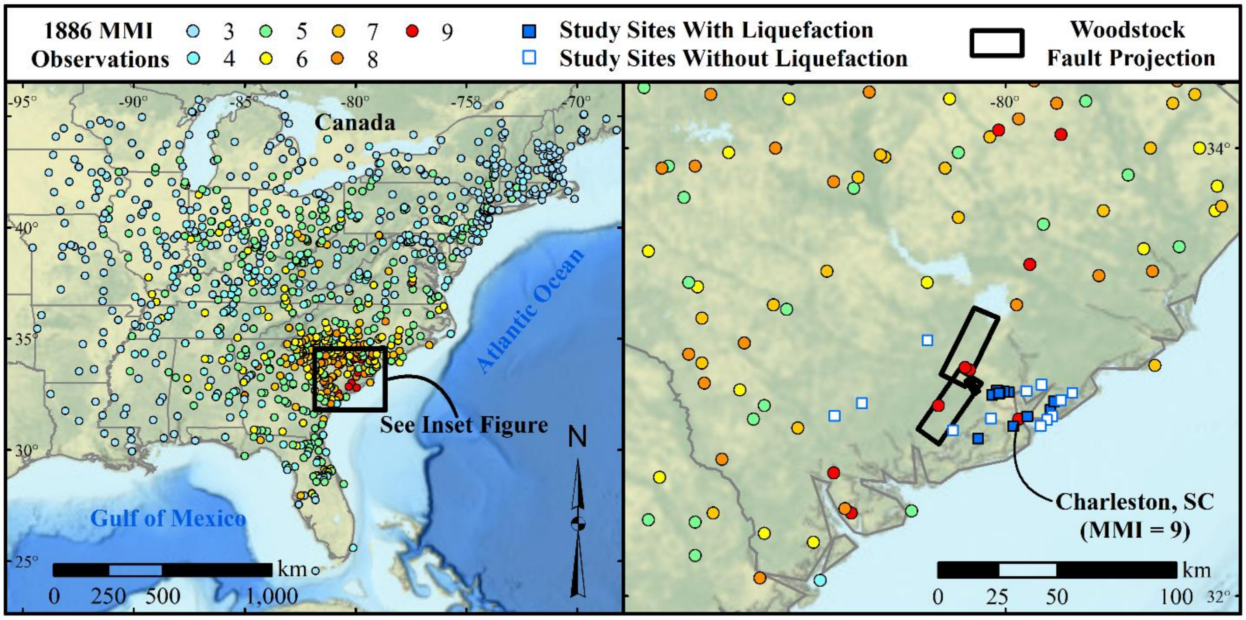


Figure 3. 1886 MMI observations (Bakun et al., 2002) and liquefaction evidence, as introduced subsequently. Also shown is the Woodstock Fault projection (Durá-Gómez and Talwani, 2009a,b).

3.2 Liquefaction data

Twenty-four sites where liquefaction manifestations were or were not observed in 1886, and where CPT testing was subsequently performed, will be studied. Liquefaction manifested (typically in the form of large sand boils) at twelve of these sites, whereas no evidence of liquefaction was observed at the remainder. These sites are summarized in Table 2, where citations are provided for the observed liquefaction response and for the geotechnical tests, which are all available in the public domain. These study sites are also mapped later in the paper and CPT data from each site is plotted in the electronic supplement.

Table 2. Summary of 1886 liquefaction data analyzed herein.

CPT ID	Longitude	Latitude	Manifestation	Geotechnical Reference	Liquefaction Reference
BKY07	-79.9061	32.9150	No	USGS (2021)	Dutton (1889)
BKY09	-79.8385	32.9443	No	USGS (2021)	Dutton (1889)
BKY23	-79.9855	32.9115	Yes	USGS (2021)	Dutton (1889)
BKY24	-80.0071	32.9118	Yes	USGS (2021)	Dutton (1889)
CHN01	-79.7900	32.8030	No	USGS (2021)	Dutton (1889)
CHN07	-79.8134	32.7874	No	USGS (2021)	Dutton (1889)
CHN12	-79.7989	32.8310	Yes	USGS (2021)	Amick et al. (1990)
CHN15	-79.6998	32.9073	No	USGS (2021)	Dutton (1889)
CHN28	-79.8428	32.7597	No	USGS (2021)	Dutton (1889)
CHN29	-79.7840	32.8682	Yes	USGS (2021)	Amick et al. (1990)
CHN31	-79.7520	32.8741	No	USGS (2021)	Dutton (1889)
CHN32	-80.0267	32.9043	Yes	USGS (2021)	Dutton (1889)
CHN33	-80.0323	32.9166	Yes	USGS (2021)	Dutton (1889)
CHN34	-80.0391	32.9170	Yes	USGS (2021)	Dutton (1889)
CHN50	-80.1235	32.7023	Yes	USGS (2021)	Dutton (1889)
CHN59	-79.9655	32.7575	Yes	USGS (2021)	Dutton (1889)
CHN64	-80.0604	32.8982	Yes	USGS (2021)	Dutton (1889)
CREC1	-80.0655	32.7921	No	Boller (2008)	Martin and Clough (1994); Boller (2008)
FHS3	-80.3507	33.1420	No	Hasek (2016)	Williamson and Gassman (2014); Hasek (2016)
HA74	-80.0300	32.9050	Yes	Heidari & Andrus (2012)	Dutton (1889)
HWD2	-80.2355	32.7394	No	Hasek (2016)	Talwani and Cox (1985); Hasek (2016)
S99634DS1	-79.9015	32.8017	Yes	Heidari & Andrus (2010)	Dutton (1889)
WLC2	-80.6444	32.8607	No	Geiger (2010)	Geiger (2010)
WRAP2	-80.7666	32.8040	No	Geiger (2010)	Geiger (2010)

In compiling these study sites, an interpreted, potential lateral spread at Fort Dorchester, SC (Talwani et al., 2011) was omitted because the liquefaction response was ambiguous and because lateral spreading – a distinct and complex manifestation of liquefaction – is not intended to be predicted by the liquefaction model adopted herein. Whereas the MMI data are spread throughout ENA, the compiled liquefaction data are within 100 km of Charleston. Because liquefaction is mechanistically implausible, irrespective of soil properties, at *PGA*s less than ~0.09 g (de Magistris, 2013), the maximum site-to-source distance of liquefaction observations is inherently limited. By corollary, an analysis of sites where liquefaction was not observed, but which are very far from Charleston, would not provide meaningful constraint of the source model, given that the computed probability of such an observation is 100% even for very large earthquakes. In addition, while the MMI data from 1886 are unlikely to grow significantly, additional liquefaction data could be compiled. That is, the liquefaction response was documented in 1886 at more than twenty-four sites, but costly geotechnical testing must also be performed at each site. In this regard, it is known that CPTs have been performed near additional sites of observation, yet these data are privately held and could not be obtained for analysis. Nonetheless, a larger dataset could be studied in the future.

4. Methodology

The methodology that will be used to analyze macroseismic data was first introduced by Rasanen and Maurer (2021, 2022), who collectively demonstrated and validated its use on landslide and liquefaction evidence produced by eleven modern earthquakes with known source models. This is the first application of the method to a prehistoric or pre-instrumental earthquake. The methodology will be covered in two sections. The first provides a succinct conceptual overview. The second describes in detail the application of the methodology to the 1886 Charleston earthquake.

4.1 Conceptual Overview

Our goal is to probabilistically constrain the 1886 source model. This is accomplished by computing the likelihood that a rupture with some magnitude, location, and geometry, would produce a set of field observations (reported MMI or liquefaction observations), wherein uncertainties inherent to these observed outcomes are considered. In general, the likelihood of a parameter having some value, given a set of observations, is the product of the probabilities of those observations, conditioned on the parameter value. In other words, the likelihood of a rupture having some magnitude (M_w), location (L), and geometry (G), given a set (*x*) of field observations at *N* different sites, can be computed as:

$$Likelihood\ (L, G, M_w | x) = P(X = x | L, G, M_w) = \prod_{i=1}^{N} P(X_i = x_i | L, G, M_w) \quad (1)$$

where $P(X_i = x_i | L, G, M_w)$ is the probability of what was observed at site *i* (i.e., the observed MMI or liquefaction response) given an earthquake with parameters L, G, and M_w. By repeating for enumerable possibilities, the actual rupture parameters may be probabilistically constrained by the likelihood function (product of the probabilities of N observations), such that different source models will be found more and less likely to produce the observed evidence.

If the evidence is MMI data, then the probability of any one field observation (obs) is:

$$P(\text{obs}|\text{EQK}: L, G, M_w) = \int_{IM} \int_{MMI} P(\text{obs}|MMI)\, f(MMI|IM)\, f(IM|L, G, M_w) \cdot \text{d}MMI \cdot \text{d}IM \quad (2)$$

where $f(IM|L, G, M_w)$ is the probability of an IM value conditioned on fault parameters L, G and M_w, and site parameter V_{S30}, as computed by a GMM that considers site response; $f(MMI|IM)$ is the probability of an MMI value conditioned on the IM value, and possibly on other parameters, as computed by an IM-MMI model; and $P(obs|MMI)$ is the binomial probability that the predicted MMI is equal to the observed MMI. In this work, we consider predictions and observations to agree if within ± 0.5 MMI (e.g., predicted MMIs of 5.51 and 6.49 agree with an observed MMI of 6). Thus, we assign an implicit, uniform measurement uncertainty of ± 0.5 MMI, but do not otherwise model the uncertainty of reported MMIs. In this regard, a thorough reinterpretation of the original intensity data, to include assignment of site-specific uncertainty distributions and corrections for bias, would be a valuable endeavor. In the current effort, however, we adopt the Bakun et al. (2002) MMI dataset, which has been studied by other modern investigators of the 1886 earthquake.

If the evidence is observed liquefaction response rather than MMI, then the probability of field observation is computed per Eq (3) if liquefaction manifestation was observed and per Eq (4) otherwise:

$$P(\text{Manifestation}|\text{EQK}: L, G, M_w) = \int_{IM} P(\text{Manifestation}|IM, M_w)\, f(IM|L, G, M_w) \cdot \text{d}IM \quad (3)$$

$$P(\text{No Manifestation}|\text{EQK}: L, G, M_w) = 1 - \int_{IM} P(\text{Manifestation}|IM, M_w)\, f(IM|L, G, M_w) \cdot \text{d}IM \quad (4)$$

In these equations, uncertainty is not assigned to the observed liquefaction response; $f(IM|L, G, M_w)$ has the same meaning as above; and $P(\text{Manifestation}|IM, M_w)$ is computed by a probabilistic model that predicts the triggering of liquefaction at depth and its subsequent manifestation at the ground surface using subsurface geotechnical data. Ultimately, the uncertainties that are, and are not, accounted for in this work will be explicitly discussed.

Application of this method is demonstrated conceptually in Figure 4 considering four MMI observations and two hypothetical sources for the earthquake that produced the observations. In actual analyses, a very large number of sources is considered. Figure 5 illustrates how the relative likelihoods of these two sources are assessed. Shown in Figure 5a are the computed probabilities of individual observations, given a rupture of source one, as computed by Eq. 2 for varying M_w. In Figure 5b, this is repeated considering a rupture of source two. In Figure 5c, the likelihood of each source is computed as a function of M_w by Eq. 1 (i.e., the product of the four probability distributions in Figure 5a or 5b). In this simple example, source one has a far greater peak likelihood of producing the set of field observations, whereas source two is very unlikely to do so, regardless of its M_w. By repeating this process for an array of hypothetical sources and generating a likelihood distribution for each, the characteristics of the causative rupture (e.g., location, orientation, magnitude) are probabilistically constrained, to the degree that evidence permits. In this work, we compute M_w probability-distributions conditioned on both the hypothesized Woodstock fault and on an unknown source. The latter is accomplished by aggregating probability distributions from all hypothetical sources and thus includes the uncertainties of fault location and geometry, whereas the former assumes the fault location and geometry are known.

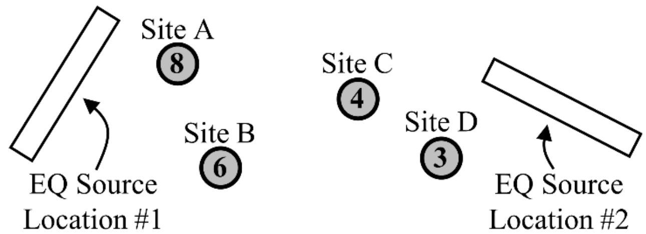


Figure 4. Hypothetical MMI inverse analysis consisting of four field sites, where MMI values were reported for each of the site locations (the value in the circle). In addition, two hypothetical sources for the earthquake that produced these observations are shown.

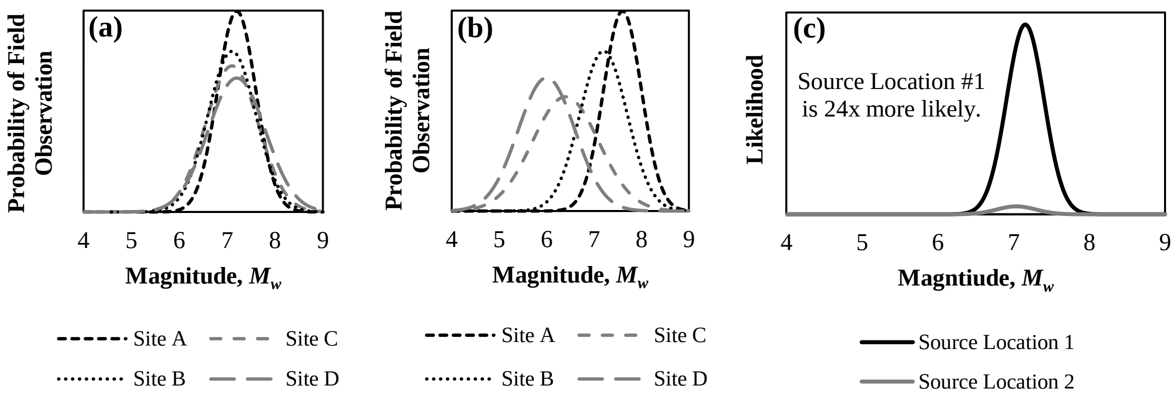


Figure 5. Approach for computing the likelihood of the sources depicted in Figure 4: (a) probabilities of individual observations, given an earthquake at location one of variable M_w; (b) probabilities of individual observations, given an earthquake at location two of variable M_w; (c) M_w-likelihood distributions for source locations one and two.

4.2 Implementation Details

The implementation of Eqs. 1-4 is next described in detail. This includes both general methods transferrable to other regions and to other forms of evidence, as well as the specific models adopted for the 1886 earthquake. Because the analyses of MMI and liquefaction data are procedurally similar, we first fully describe the former and then succinctly discuss differences specific to the latter. For completeness, however, the liquefaction analysis procedure is fully replicated in the electronic supplement.

As detailed in the following, two approaches will be used to model hypothetical seismic sources. These sources are first treated as earthquake epicenters on a grid pattern, which we term the "epicenter search". To use modern GMMs, which are applicable only to faults, and because modeling large earthquakes as points is inappropriate, each epicenter is converted to an amorphous fault realization using site-to-source distance correlations, which implicitly assume some fault geometry as a function of magnitude. With this approach, the source location may be investigated (i.e., probabilistically constrained) without consideration of rupture geometry. While this reduces computational expense, it considers only a "median" geometry and thus omits one source of uncertainty. Accordingly, and following constraint of the epicentral region, seismic sources are next treated as faults having an array of locations, lengths, and orientations, which we term the "fault search." Of these enumerable faults, one is the hypothesized Woodstock Fault, which we highlight in the results for obvious reasons.

4.2.1 MMI Analysis Procedure

The analysis of MMI data is completed via the following 15 steps:

(1) Create an array of seismic sources. In this paper, a 62,500 km^2 grid of surficial points was centered on Charleston. Within this grid, a finer point spacing increases spatial resolution while a coarser spacing decreases computational demand (which can be significant, and tractable only with high performance computing). Having found that 1 km and 10 km spacing produce nearly identical outcomes, we adopted the latter spacing when studying MMI observations. These surficial points will be treated first as earthquake epicenters (i.e., the "epicenter search") and second as surface projections of the centroids of faults (i.e., the "fault search").

(2) Select N study sites (i.e., the *MMI* observations compiled by Bakun et al. (2002)).

(3) Select an appropriate GMM. In this paper, the 17 GMMs developed by the NGA East Project (Goulet et al., 2018) were coalesced in a logic tree using the weights proposed by Goulet et al. (2018). Using this scheme, the least and greatest model weights were ~2% and 10%, respectively.

(4) For each seismic source created in (1):

(5) For each seismic-source M_w considered (a range of M_w4 to M_w8.2 – the applicable range of the adopted GMMs – was used in this paper):

(6) For each of N study sites selected in (2), cycling from $i = 1$ to N:

(7) Compute the site-to-source distance(s) required by the GMM chosen in (3), as measured from study site i to the seismic source selected in (4). For the GMMs adopted herein, the only such metric required is the closest distance to fault rupture (R_{RUP}). When the seismic sources in (1) are treated as epicenters, rather than faults, the correlations of Scherbaum et al. (2004) were used to estimate a median R_{RUP} from epicentral distance (R_{EPI}). In effect, these correlations, which are magnitude dependent, convert each epicenter from (1) into a median realization of a

multidimensional fault. Alternative approaches to estimating R_{RUP} from point sources are provided by Bommer et al. (2016) and Thompson and Worden (2018). Ultimately, the sources in (1) are explicitly modeled as faults via the "fault search" to determine whether the field evidence can constrain the 1886 source more precisely. In doing so, R_{RUP} is directly measured from fault planes and the uncertainty of fault orientation is considered.

(8) Using the GMM in (3), M_w from (5), and site-to-source distances from (7), compute the probability density function (PDF) of expected *PGA* at site *i*, modified for site effects. In general, this PDF is a lognormal random variable described by a median and lognormal standard deviation, which are given by a GMM. *PGA* predictions beyond ± 3 standard deviations of the median were truncated, as is typical, and the PDF was scaled such that the area beneath it was one. In this study, the ENA weighted GMM predicts *PGA* for reference rock conditions. Accordingly, the V_{S30}-dependent model of Harmon et al. (2019), which is ENA-specific and developed as part of the NGA-East Project, was used to adjust *PGA*s for local site effects. V_{S30} was estimated at the site of each field observation using the Heath et al. (2020) maps.

(9) For each possible *PGA* value at study site *i*, as computed in (8) for a given M_w and R_{RUP} pair:

(10) Select an IM-MMI model. In this paper, the Atkinson and Kaka (2007), Worden et al. (2012), Caprio et al. (2015), and Cramer (2020) models were adopted. The inputs to these models vary but generally include *PGA*, M_w, and R_{RUP}. The median MMI predicted by each model is plotted in Figure 6 as a function of *PGA* considering an M_w6 event at R_{RUP} = 100 km.

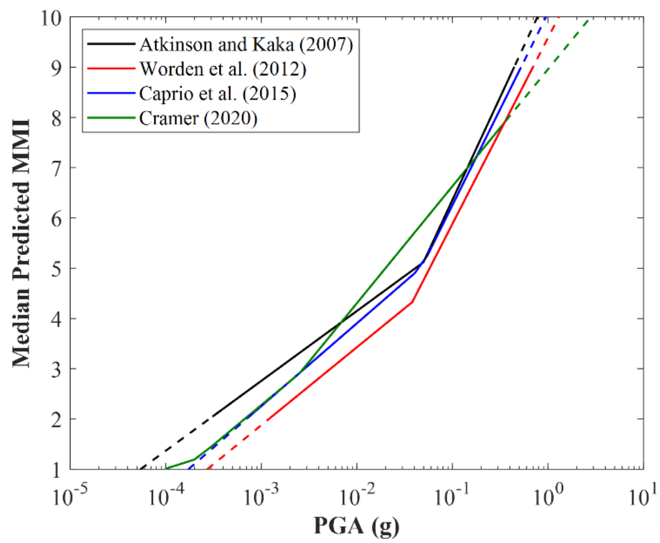


Figure 6. Predicted median MMI versus PGA according to four recent models (where applicable, M_w = 6.0; R_{RUP} = 100 km). Solid lines indicate the range of each model's proposed applicability (typically the range of training data); dashed lines indicate extrapolation beyond these respective bounds.

(11) Using the IM-MMI model selected in (10), the R_{RUP} from (7), and the *PGA* from (9), compute the PDF of expected *MMI* at study site *i*. In this paper, *MMI* predictions beyond $1 \leq MMI \leq 10$ were truncated because the definition of *MMI* gives a lower bound of 1 and because intensities greater than 10 are rarely assigned in practice (Stover and Coffman, 1993). Following truncation (when applicable) the PDF was scaled such that the area underneath it was one.

(12) For each possible *MMI* value at study site *i*, as computed in (11):

(13) Compute the probability of field observation as described in Eq. 2. Completing this equation (i.e., by multiplying the probability of field observation by the probabilities of *MMI* and *PGA*, then summing over all *MMI* and *PGA* values) gives the probability of the field observation at site *i* for a given seismic source location, geometry, and M_w. Repeating steps 6-13 for all M_w results in a probability of field observation curve for each source, examples of which are in Figures 5a and 5b.

(14) Compute the likelihood of a seismic source (as evidenced by MMI) as a function of M_w by multiplying the probabilities of all field observations (i.e., multiply the curves in Figures 5a or 5b at each value of M_w). The result, an example of which is shown in Figure 5c, is a likelihood distribution of M_w conditioned on a single seismic source (i.e., assuming that a fault with given location and geometry, but unknown M_w, is responsible for producing the field observations).

(15) Repeating steps 5-14 for all seismic sources created in (1) results in a likelihood distribution of M_w for each. Collectively, this field of distributions describes the locations and magnitudes of earthquakes that are, and are not, likely to produce the field evidence. To allow for relative likelihoods to be compared visually, we normalize the likelihood of each source by the peak likelihood among all sources, such that the most likely source has a normalized peak value of one. We then map contours of likelihood to identify this location. The M_w distribution at this location (see Figure 5c) is the PDF of the inverted M_w conditioned on the most likely source. Finally, by aggregating PDFs from all potential sources in (1), an overall PDF of M_w, considering all possible sources, is produced. While a single source will always be "most likely," earthquakes at other locations typically also have potential to produce the evidence. This latter PDF, conditioned on all possible sources, includes that uncertainty.

4.2.2 Liquefaction Analysis Procedure

The analysis of liquefaction is akin to that of MMI, differing only in steps (10) through (13) of the preceding. Once site-adjusted *PGA*s are predicted at sites of observation (in this case, sites with liquefaction observations), the factor of safety against liquefaction is computed throughout the soil profile using the Boulanger and Idriss (2014) CPT-based triggering model, which is a function of *PGA*, M_w, and subsurface geotechnical data. As part of this procedure, the effects of soil aging are separately accounted for with three proposed models, in addition to a control analysis without any such accounting. Two of these models, henceforth termed "A" and "B", are regional-scale corrections based on the Weems et al. (2014) geology map of the Charleston region, whereas Model "C" uses site-specific measurements. Model A uses the Hayati and Andrus (2009) aging-correction model based on the measured to estimated shear-wave velocity ratio (MEVR). As MEVR increases, the shear stiffness of soil measured at small strain exceeds that inferred from large strain measurements, which may be interpreted as an indicator of cementation. Median MEVR values for each geologic unit were adopted from the regional sampling of Heidari and Andrus (2012). Model B uses the Hayati and Andrus (2009) correction model based on depositional age, which we estimated for each geologic unit from the Weems et al. (2014) map. As a deposit's age increases, an increasingly larger correction is applied to the computed liquefaction resistance. Lastly, Model "C" again adopts the MEVR-based approach of Hayati and Andrus (2009), but in this case MEVR is directly computed at each study site using data from seismic CPTs (i.e., CPT resistances and shear-wave travel times) per the method of Andrus et al.

(2009). This approach produced corrections that tended to exceed those of Models A and B. While corrections from Model C – being based on site-specific subsurface data – should be most efficient, all such corrections are likely very uncertain and the best approach to account for soil aging is actively debated. We thus argue that all three models, which scale upward the computed liquefaction triggering resistance at depth, warrant consideration. However, given that the field observations are of surface manifestations (i.e., liquefaction vents and dikes) rather than observations of liquefaction at discrete depths, surface manifestations must be predicted for proper comparison to field observations. Accordingly, the liquefaction potential index *(LPI)* of Iwasaki et al. (1978) was adopted, given its longstanding use:

$$LPI = \int_0^{20\,m} F\left(FS_{liq}\right) \times w(z)dz \tag{5}$$

where $F\left(FS_{liq}\right)$ and $w(z)$ weight the respective influences of FS_{liq} and depth, z, on surface manifestation. In brief, *LPI* assumes that manifestation depends on the FS_{liq} in each soil stratum, the thickness of all liquefied strata, and the proximity of those strata to the surface. Given this definition, *LPI* can range from 0 to 100. Next, the probability of liquefaction manifestation was computed at each study site, considering all values of M_w from (5) and all values of *PGA* from (9), using the fragility function of Geyin and Maurer (2020), which relates the probability of liquefaction manifestation to *LPI*, and which was trained on all globally available CPT-based liquefaction case histories. Lastly, and analogous to (13), the probability of what was observed in the field was computed by Eq. 3 or 4, depending on whether manifestations were or were not observed. Thus, the fact that the liquefaction hazard differs at different study sites is directly measured and accounted for via site-specific CPT measurements and the latest CPT-based liquefaction triggering and manifestation models. Using these data and tools, the probability of surficial liquefaction manifestation is computed for a range of M_w and *PGA* at each study site. The liquefaction analysis procedure is fully replicated in the electronic supplement.

5. Results and discussion

Using the above data and methodologies, the source of the 1886 Charleston rupture was rigorously investigated. Results are separately presented for MMI and liquefaction evidence. For each, three types of analyses were performed, namely the: (i) "epicenter search;" (ii) "fault search;" and (iii) "Woodstock Fault," wherein the hypothesized Woodstock Fault is directly considered. The first provides preliminary probability distributions of the source location and magnitude. The second provides final distributions that include additional source-model uncertainties and conveys whether the orientation of the causative fault can be constrained. The third provides results conditioned on the singular Woodstock Fault and should be adopted if all other seismic sources, known and unknown, are dismissed as sources of the 1886 earthquake.

5.1 Analysis of MMI Data: Epicenter Search

Using the "epicenter search," the inversion methodology was applied with the Atkinson and Kaka (2007), Worden et al. (2012), Caprio et al. (2015), and Cramer (2020) IM-MMI models. For brevity, these models are henceforth titled AK07, WEA12, CEA15, and C20. For these analyses, MMI observations were studied if within 600 km of Charleston. While the sensitivity of this decision will be analyzed further, it was made based on a trait of the Bakun et al. (2002) MMI dataset that we view as problematic, but which prior researchers have not discussed. Namely, Bakun et al. (2002) did not include "negative" observations (i.e., MMI values of 1 or 2) where ground motions were not perceived. This is notable, given that a *distribution* of MMI values is naturally experienced at any given distance. Problematically, the omission of low MMI values asymmetrically truncates this distribution at large distances (i.e., those where some observers report MMI 1 or 2). The analysis of this dataset at large distances thus potentially has the effect of biasing inverted magnitudes toward larger values, since those magnitudes need not adhere to small MMI values that were experienced, but which were systematically undocumented. The analysis of data only at small distances is equally undesirable, of course, given the inherent benefits of analyzing more data over a wider spectrum of attenuation. Based on these competing interests and given that we begin to observe evidence of MMI truncation at distances exceeding 600 km, this threshold was provisionally selected.

The resulting geospatial contours of seismic-source likelihood are shown in Figure 7 for the WEA12 model. Also shown is the surface projection of the hypothesized Woodstock fault (Durá-Gómez and Talwani, 2009a,b), as well as instrumental epicenters (M_w > 1.0; 1900-present). These contours show the source locations most and least likely to produce the MMI evidence. Using the WEA12 model, a zone of high likelihood is computed just northwest of the Woodstock Fault projection, with the most likely source located 2 km from the mapped projection. By contrast, source locations to the south and east are comparatively very unlikely to produce the MMI observations. For brevity, results using the AK07, CEA15, and C20 models are summarized in Figure 7 via stars, which denote the epicenters deemed most likely by each model. Complete contour maps, analogous to those in Figure 7 for WEA12, are provided in the electronic supplement. While CEA15 produces a result similar to WEA12, with the most likely source inside the Woodstock Fault projection, AK07 and C20 suggest most likely sources ~10 km and ~35 km east of the projection. Thus, all analyses point to a source north of Charleston. Some IM-MMI models strongly support an earthquake source in the immediate vicinity of the hypothesized Woodstock Fault, while others do not.

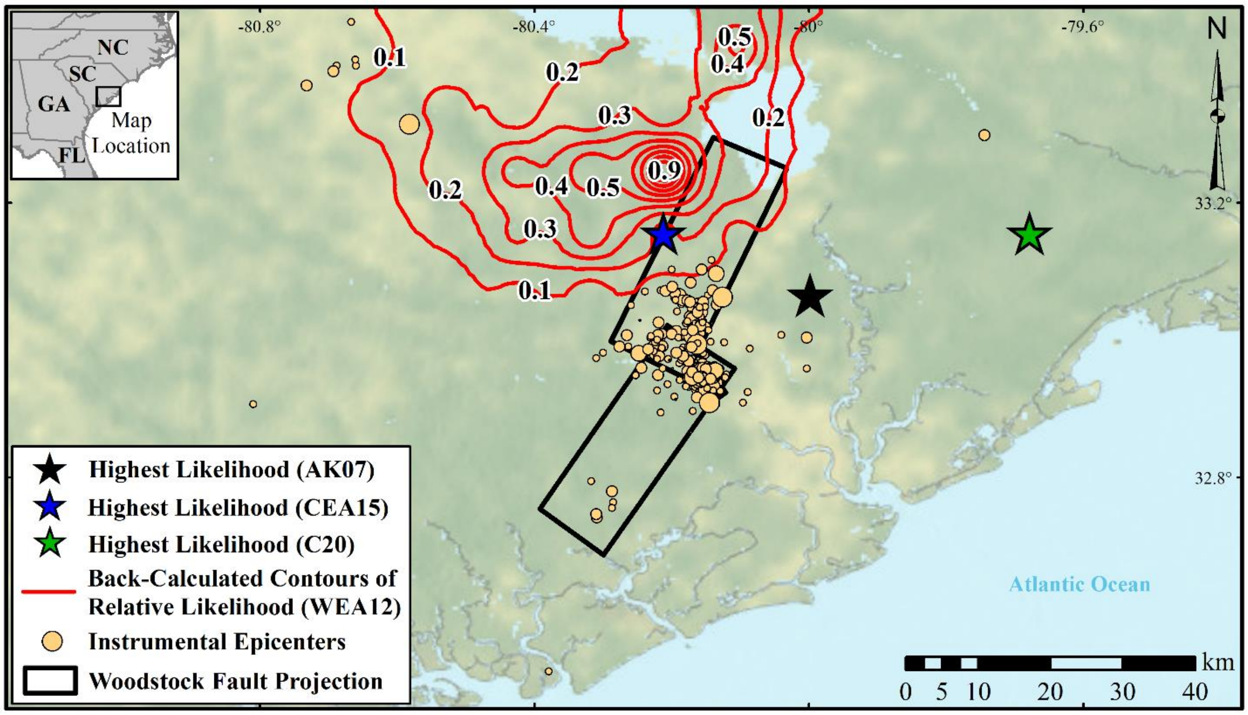


Figure 7. Likelihood contours produced using the WEA12 model which indicate source locations that are, and are not, likely of producing the MMI evidence. Stars = the most likely source locations per the AK07, CEA15, and C20 models; AK07 = Atkinson and Kaka (2007), WEA12 = Worden et al. (2012), CEA15 = Caprio et al. (2015), C20 = Cramer (2020); black rectangles = the Woodstock fault projection (Durá-Gómez and Talwani 2009a,b); tan circles = instrumental epicenters (M_w > 1.0; 1900-present) (USGS, 2022).

Shown in Figure 8 are the cumulative distribution functions (CDFs) of M_w inverted from MMI data, as computed with each of the four IM-MMI models. These results include the uncertainty of an unknown source (epicenter) location, the uncertainties of expected shaking intensities conditioned on a given source, considering also the epistemic uncertainty of GMM selection, and the uncertainties of the IM-MMI models. For comparison, the CDF assigned to the Charleston source in the NSHMP (Petersen et al., 2014, 2020) is also shown and ranges from M_w6.7 to M_w7.5

with a median of M_w7.1. AK07 and C20 produce nearly identical CDFs with a median of M_w7.1. WEA12 and CEA15 produce similarly identical CDFs, but with a median of M_w7.3. The grouping of these outcomes can be traced to Figure 6. AK07 and C20 tend to predict larger MMIs for a given PGA. As a result, smaller magnitude ruptures result from the use of AK07 and C20 within the inversion methodology. Each of the individual CDFs has relatively low uncertainty. This can be attributed to: (i) the large quantity of MMI observations; (ii) the fact that not all uncertainties are considered (e.g., those of site conditions or rupture geometry, among others); and (iii) the relatively large variance of MMI observations over any given site-to-source distance. Regarding the last, extreme outliers from the mean MMI trend have the effect of yielding very low likelihoods for small and large magnitude events. That is, only a narrow range of magnitudes are likely to simultaneously produce MMI observations of, say, 3, 5, and 7 at the same site-to-source distance and on similar site profiles. Conversely, a wider range of magnitudes could simultaneously produce MMI observations of 4, 5, and 6. While geostatistical analyses could conceivably identify and delete observations that are extreme outliers, the justification would be largely speculative without an intensive reinvestigation. In other words, we are unaware of any objective basis for deleting some MMI reports but not others in the absence of a complete reinterpretation of the more than 1000 original intensity reports. Lastly, to include the epistemic uncertainty of IM-MMI model selection, each CDF was weighted to produce an ensemble CDF, as shown in Figure 8. This selection of weights is heuristic, as no quantitative justification could be identified (e.g., based on model residuals in ENA). The weights and justifications are as follows. C20 (0.4) is the latest ENA-specific model and was trained using the largest quantity of ENA data. AK07 (0.3) is also ENA-specific, but also found no need for region-specific models within North America. For that reason, and because differences in methodology (e.g., data selection, intensity scales, regression techniques) can result in very different models with apparently similar performance (i.e., on the respective training sets), we hesitate to omit other reputable models even if nonspecific to ENA. Accordingly, WEA12 (0.15), which was trained only on California data, and CEA15 (0.15), which was trained on global data including some from ENA, are also weighted in the ensemble. While CEA15 might otherwise warrant greater weighting, its training set was truncated at a site-to-source distance of 200 km, which calls into question its suitability in ENA, where ground motions are commonly felt at greater distances (as in the 1886 event).

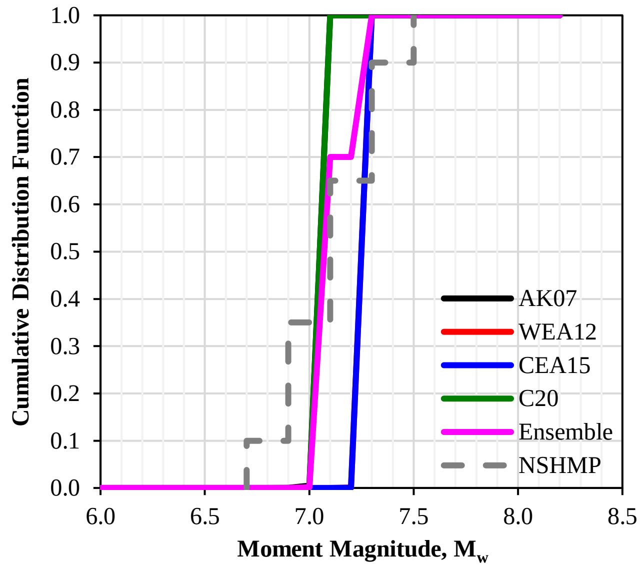


Figure 8. CDFs of rupture magnitude, as inverted from MMI observations using four IM-MMI models, wherein seismic sources are modelled as epicenters. AK07 = Atkinson and Kaka (2007), WEA12 = Worden et al. (2012), CEA15 = Caprio et al. (2015), C20 = Cramer (2020), and NSHMP = National Seismic Hazard Model Project (Petersen et al., 2014, 2020).

5.2 Analysis of MMI Data: Fault Search (Seismic Sources Modelled as Faults)

To determine whether MMI observations can constrain the seismic source beyond a general area, the gridded sources were next treated as surface projections of the centroids of faults. For simplicity, the faults were initially assumed to be strike-slip with a dip of 90 degrees. The orientations of the faults were discretized into strike azimuths of 30˚ increments (i.e., 0˚, 30˚, 60˚…). The lengths of the faults were computed via the magnitude-dependent correlation of Wells and Coppersmith (1994). The depth to the top of rupture was estimated using the correlation of Kaklamanos et al. (2011), which produces estimates consistent with the inferred depths of modern ruptures in the region (Chapman et al., 2016). Thus, while not every aspect of rupture geometry was assigned uncertainty, the analyses can nonetheless determine whether some fault alignments (e.g., the hypothesized Woodstock Fault) are more likely than others to produce the observed evidence. With this approach, the inversion methodology was again applied to the Bakun et al. (2002) dataset using each of the four IM-MMI models.

Results are shown in Figure 9 for the WEA12 model in the style of a heat map. Faults more likely to produce the MMI observations have thicker lineweight and are darker in color. Faults very unlikely to produce the evidence have thin lineweight and light color, and thus blend with the map's background. Also shown is the surface projection of the hypothesized Woodstock Fault

(Durá-Gómez and Talwani 2009a,b) as well as instrumental epicenters. While most modelled faults are relatively unlikely to produce the evidence, the faults deemed most likely align with the hypothesized Woodstock Fault, albeit their magnitudes are greater than most prior estimates. The singular fault most likely to produce the MMI evidence, for example, has a median M_w of 7.60. Faults striking E-W and N-S have lesser, but nontrivial, likelihoods of producing the MMI evidence.

For brevity, results using the AK07, CEA15, and C20 models are each summarized in Figure 9 by single lines, which denote the singular faults deemed most likely by each model. Complete heat maps, analogous to those in Figure 9 for WEA12, are provided in the electronic supplement. These three models produce results that are similar to one another and different from WEA12, with the most likely faults oriented perpendicular to the Woodstock Fault and located partly offshore. These most-likely faults range in median magnitude from M_w7.20 to M_w7.90. A close inspection of these results (electronic supplement) reveals that each of these models finds faults consistent with the Woodstock Fault to have low relative likelihoods of producing the MMI evidence, as compared to all other hypothetical faults. It is worth noting that the two ENA-specific IM-MMI models (AK07 and C20) point to faults inconsistent with the Woodstock Fault. While these results do not necessarily indicate that the Woodstock Fault was not responsible for the 1886 earthquake, they do indicate that the causative fault cannot currently be constrained by the available MMI evidence. That is, faults striking SE and partly offshore are just as likely to produce this evidence as faults striking NE and onshore. While different assumptions about fault depth, length, and dip invariably change the results, a parametric analysis indicates that these changes are minor, upholding the overall conclusion that the MMI data cannot constrain the 1886 fault rupture. If anything, the data point to a fault striking SE and partly offshore. This conclusion might be different if: (i) more MMI observations were available in the near field; and (ii) the affected area was not on a coastline. Analogous to the inversion of an epicenter from wave-arrival times, which relies on distributed instruments, this inversion relies on distributed observers to “record” the amplitudes of ground motions. Thus, in events such as this, where the near-field is not more densely populated with observations than the far field, and where observations are geographically asymmetric, it may be more difficult to detect directions and rates of ground-motion attenuation, and thus more difficult to constrain the causative fault from which motions propagated.

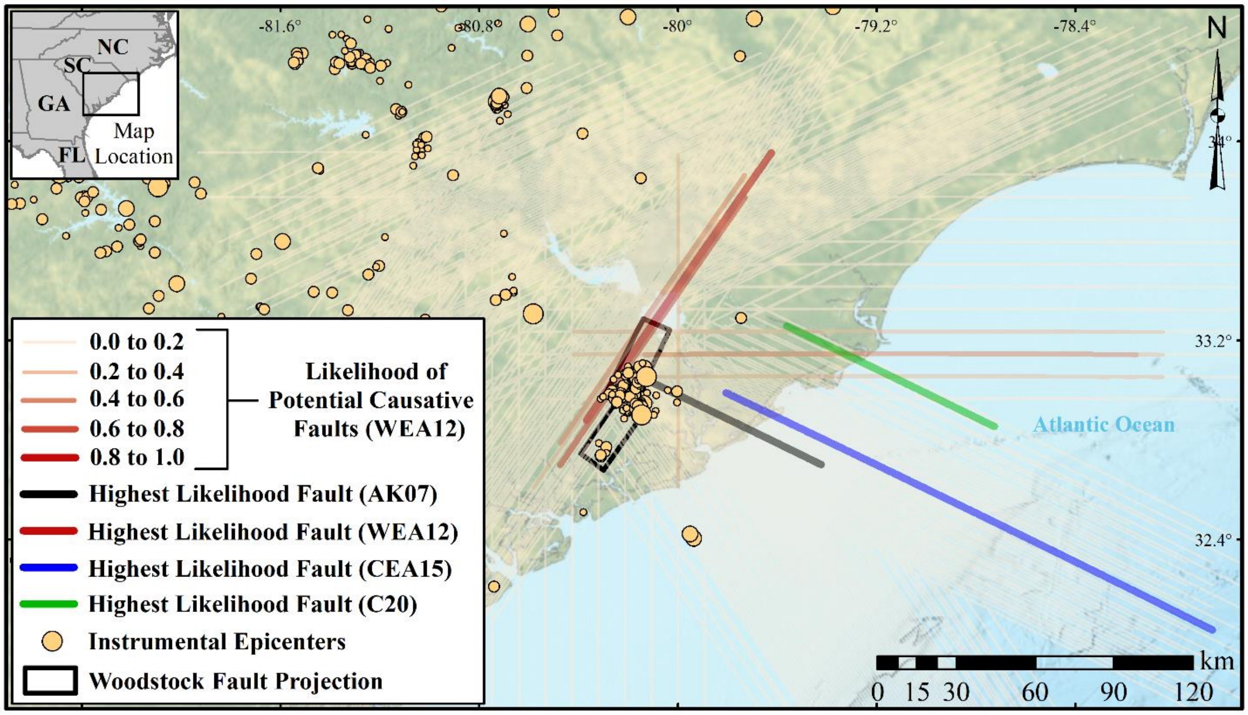


Figure 9. Likelihoods of hypothetical faults with differing orientations and lengths to produce the MMI evidence, as obtained using the WEA12 model. Also shown are the singular faults deemed most likely using the AK07, CEA15, and C20 models; AK07 = Atkinson and Kaka (2007), WEA12 = Worden et al. (2012), CEA15 = Caprio et al. (2015), C20 = Cramer (2020); black rectangles = the Woodstock projection (Durá-Gómez and Talwani 2009a,b); tan circles = instrumental epicenters (M_w>1.0; 1900-present) (USGS, 2022).

Shown in Figure 10 are the CDFs of M_w inverted from MMI data via the “fault search”, as computed with each of the four IM-MMI models. These results include the uncertainty of an unknown source location, the uncertainty of unknown fault orientation, the uncertainties of expected shaking intensities conditioned on a given source and considering the epistemic uncertainty of GMM selection, and the uncertainties of the IM-MMI models. It should be emphasized that these results do not include every source of uncertainty. Omitted, for example, are the uncertainties of: (i) fault depth and length; (ii) fault dip; and (iii) site V_{S30}. For each of these inputs only a median prediction was used. For comparison, the CDF assigned to the Charleston source in the NSHMP (Petersen et al., 2014, 2020) is also shown. Like the “epicenter search” results in Figure 8, the AK07 and C20 models produce similar CDFs with a lesser median of ~M_w7.2, while WEA12 and CEA15 produce similar CDFs with a greater median of ~M_w7.7. It is readily apparent that the epistemic uncertainty of IM-MMI model selection is considerable. To account for this uncertainty, each CDF was weighted per the prior scheme to produce an ensemble CDF, as shown in Figure 10, which has a median of ~M_w7.25 and 95% CI of M_w7.05 to M_w7.85. In this regard, efforts to better quantify the suitability of various IM-MMI models to ENA, and in turn to refine the weighting scheme used herein, could have significant influence on the overall conclusions of this study (and presumably also on studies of other seismic sources). Despite this ambiguity, the ensemble CDF in Figure 10 suggests that ruptures larger than M_w7.5 – the largest value considered in the NSHMP – could produce the observed evidence.

Of course, it should also be emphasized that the numerous hypothetical faults considered in our methodology and aggregated to form the results in Figure 10 may not exist. That is, some of these gridded faults are more likely to produce the 1886 MMI data than the hypothesized Woodstock Fault, but these various faults are not necessarily present. Conversely, there is evidence (e.g., seismological data, geophysical surveys) supporting the Woodstock Fault's existence (Durá-Gómez and Talwani 2009a,b; Chapman et al., 2016). Accordingly, the preceding analyses are next repeated, conditioned solely on the Woodstock Fault, thereby removing all source uncertainties aside from magnitude.

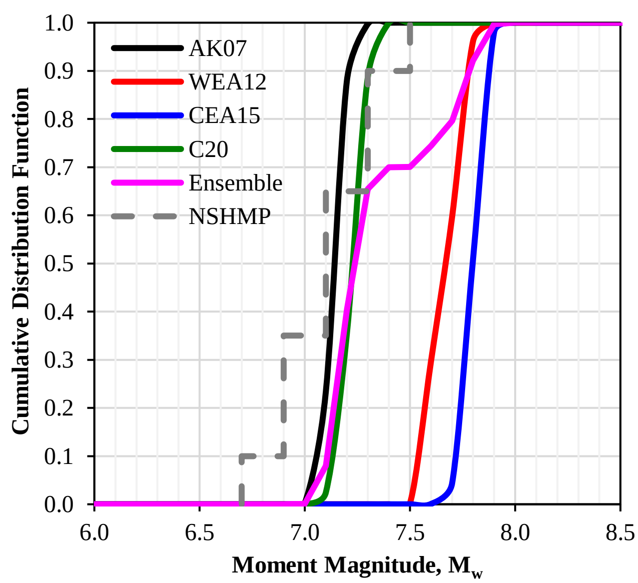


Figure 10. CDFs of rupture magnitude, as inverted from MMI observations using four IM-MMI models, wherein seismic sources are modelled as faults. AK07 = Atkinson and Kaka (2007), WEA12 = Worden et al. (2012), CEA15 = Caprio et al. (2015), C20 = Cramer (2020), and NSHMP = National Seismic Hazard Model Project (Petersen et al., 2014, 2020).

5.3 Analysis of MMI Data: Woodstock Fault

While the parameters of the Woodstock Fault are uncertain, we adopt the source model of Durá-Gómez and Talwani (2009a,b) exactly as proposed therein, the surface projection of which appears in Figures 1, 7, and 9. While Chapman et al. (2016) do not identify the Woodstock Fault by name, they infer and describe from more recent seismological records a source that is similar: “We interpret …that the modern seismicity is the lingering aftershock sequence of the 1886 shock and that the mainshock occurred on a south-striking, west-dipping fault plane ...” To evaluate the prior site-to-source distance threshold of 600 km (i.e., that which was used to exclude MMI observations), these analyses were also repeated using MMI data at different site-to-source distances. Shown in Figure 11, for example, are the CDFs of M_w conditioned on the Woodstock Fault and inverted from MMI data in two distance bins: (a) < 600 km; and (b) < 400 km. Several observations are made from Figure 11 as follows.

First, it is observed that the ensemble CDF in Figure 11a is only marginally less uncertain than the ensemble CDF in Figure 10, meaning that conditioning the analysis on the Woodstock Fault does not significantly alter the results. In other words, the uncertainties of fault location and geometry are relatively minor, given the data available for analysis. As previously discussed in Figure 9, these data are unable to constrain the causative fault. That is, faults with diverse locations and orientations have similar likelihoods of producing the MMI evidence, which may be due to a paucity of near-field observations. However, because similar rupture magnitudes are inferred for these various faults irrespective of their positions, the CDFs conditioned on the Woodstock Fault and the CDFs conditioned on an unknown fault are similar.

Second, a site-to-source distance bias is observed from Figures 11a and 11b. While the CDFs become more vertical as more observations are included (reflecting a decrease in uncertainty) the CDFs also tend to increase in magnitude. Observations at distances up to 600 km suggest larger rupture magnitudes than those at distances up to 400 km. The median magnitude, for example, increases by ~$0.25M_w$ according to CEA15, by ~$0.1M_w$ according AK07 and C20, and by ~$0.01M_w$ according to WEA12. It could be shown that the prior “epicenter search” and “fault search” results have a similar degree of sensitivity. This distance bias could be present in either: (i) one or more of the component models utilized (e.g., the GMMs, site-response model, or IM-MMI models); or (ii) in the MMI data itself. With respect to the adopted models, all were shown to be unbiased during their respective trainings and cannot be further tested in the absence of additional data. It is worth noting again, however, that the CEA15 model, which exhibits the greatest site-to-source distance sensitivity in Figure 11, is herein applied to data much more distant than it was trained on. It was for this reason that CEA15 was given low weighting despite having a large global training set. With respect to the MMI data, we reiterate that analyzing more data over a wider spectrum of attenuation is beneficial but remain adamant that the Bakun et al. (2002) dataset is apt to introduce bias at large distances, given that small MMI values were systematically undocumented. To further probe the issue of IM-MMI model applicability, the analyses were performed using only the data to which each model is applicable (as stated by the original authors, or otherwise interpreted by the present authors). Specifically, the WEA12, AK07, CEA15, and C20 models were respectively applied to observations within site-to-source distances of 400 km, 800 km, 200 km, and 1500 km. The results are shown in Figure 11c. The most salient changes are: (i) the CEA15 CDF is more uncertain (due to fewer field observations) and has a reduced median of ~M_w7.2; and (ii) the C20 CDF is less uncertain (due to more field observations) and has an increased median of ~M_w7.35. In summary, we are unsure why the results show evidence of site-to-source distance bias at distances less than 600 km. However, we argue that the bias observed at larger distances (e.g., the C20 result in Figure 11c) is at least partly due to the aforementioned bias of the Bakun et al. (2002) dataset. As such, we prefer not to glean new conclusions from Figure 11c or champion its results over those in Figure 11a. Our preferred CDF conditioned on the Woodstock Fault thus has a median of M_w7.10 and 95% CI of M_w7.0 to M_w7.75. Given an unknown source, our preferred CDF (Figure 10) has median of ~M_w7.25 and 95% CI of M_w7.05 to M_w7.85. For either result, the epistemic uncertainty of IM-MMI model selection has significant influence.

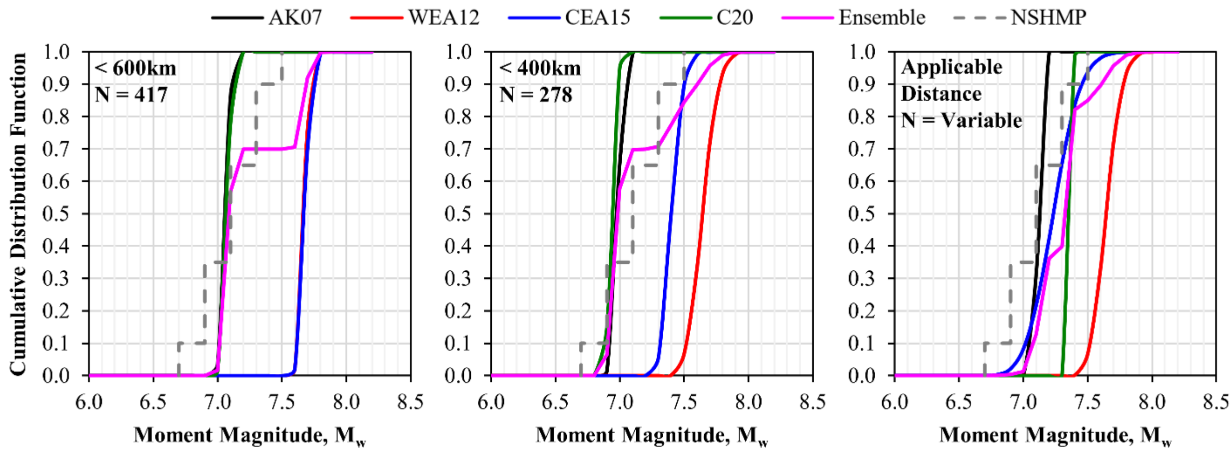


Figure 11. CDFs of rupture magnitude, as inverted from MMI observations using four IM-MMI models and assuming the Woodstock Fault (Durá-Gómez and Talwani, 2009a,b) to be the earthquake source. AK07 = Atkinson and Kaka (2007), WEA12 = Worden et al. (2012), CEA15 = Caprio et al. (2015), C20 = Cramer (2020); NSHMP = National Seismic Hazard Model Project (Petersen et al., 2014, 2020).

5.4 Analysis of Liquefaction Data: Epicenter Search

The inversion methodology was next applied to the 24 study sites summarized in Table 2 where the 1886 liquefaction response was observed, or has since been investigated, and where CPT testing has been performed. Analogous to the analysis of MMI data, the seismic source was first treated as an epicenter with unknown location. The resulting contours of seismic-source likelihood, which depict the epicenters most and least likely to produce the observed liquefaction response, are mapped in Figure 12, as are the 24 study sites. Also shown are instrumental epicenters and the hypothesized Woodstock Fault (Durá-Gómez and Talwani, 2009a,b). These initial results include no correction to account for the effects of soil aging on liquefaction. As shown in Figure 12, the analysis identified a source area east of the Woodstock fault. While the single most likely epicenter is 9 km from the proposed fault projection, epicenters within the projection have up to 70% relative likelihood. Notably, these geolocation results are largely insensitive to the treatment of soil aging. Because aging correction models tend to adjust the computed liquefaction resistances by similar amounts, the inverted magnitude tends to be significantly affected, whereas the inverted, most likely source location does not. For this reason, geolocation results are not shown for each aging-correction model. All give similar results to those in Figure 12, indicating that an epicenter more than ~20 km in any direction from the northern Charleston Peninsula (e.g., the Charleston Airport) is relatively unlikely to produce the observed evidence.

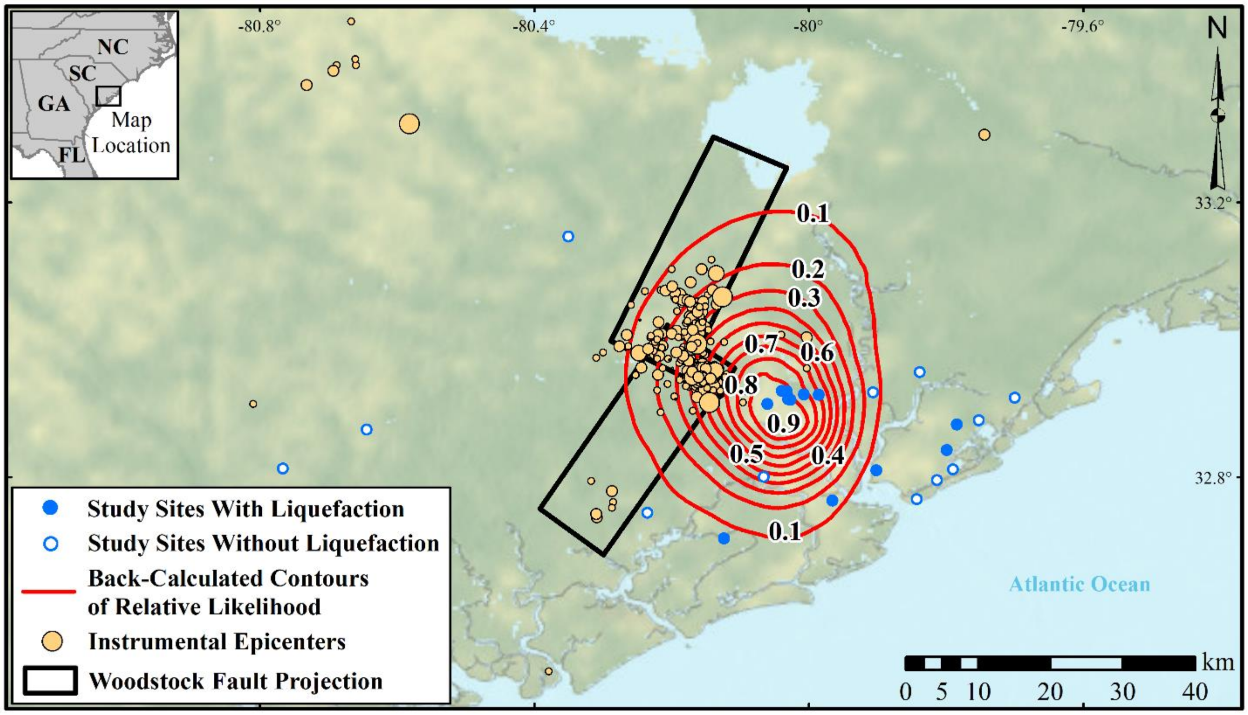


Figure 12. Likelihood contours indicating source locations that are, and are not, likely of producing the observed liquefaction response. Black rectangles = the Woodstock fault projection (Durá-Gómez and Talwani 2009a,b); tan circles = instrumental epicenters (M_w>1.0; 1900-present) (USGS, 2022).

Shown in Figure 13 are the CDFs of M_w inverted from liquefaction evidence, as computed using: (i) no aging correction; and (ii) each of the three correction methods discussed in the *Liquefaction Analysis Procedure* (i.e., Models "A, B, and C"). It is apparent from Figure 13 that the uncertainties of *whether* and *how* to correct for the effects of soil aging have significant influence, with the inverted median magnitudes ranging from ~M_w6.3 (no aging correction) to ~M_w7.0 (Model C). Irrespective of soil aging, it is also apparent that a magnitude inverted from the available liquefaction evidence is: (i) more uncertain than a magnitude inverted from the available MMI evidence; and (ii) more uncertain than past studies have reported. The former is attributable to there being far more MMI data than liquefaction data. The latter, as previously discussed, is attributable to published uncertainty bounds (e.g., "M_w6.8-M_w7.0") being ranges of the *median* magnitude considering one source of uncertainty, which is distinctly different from a CDF of magnitude. Consequently, these results suggest more uncertainty than is adopted in the NSHMP. Magnitudes below M_w6.7 and above M_w7.5, for example (i.e., the limits of the NSHMP weighting), have non-trivial probabilities of producing the field evidence. It should also be noted that these results do not include uncertainty within the aging-correction models (which are certainly uncertain). That is, the models provide a median correction factor. To include the epistemic uncertainty of model selection, each CDF was weighted to produce an ensemble. As with the MMI results, our selection of weights is based more on judgement than on quantitative evidence. While we agree with the developers of aging-correction models that corrections are likely warranted, these corrections are actively debated and likely have large uncertainty. The weights selected and their justifications are as follows. Model C (0.5) is the only site-specific method. It uses detailed subsurface geotechnical measurements from each study site to compute site-specific corrections and thus arguably warrants the greatest weighting. In contrast, Models A (0.25) and B (0.25) both rely on a sampling of the median characteristics of the geologic unit in which each study site resides (e.g., the age of the unit). These models, which do not use site-specific information, are therefore better suited for regional scale analyses (i.e., where subsurface data is unavailable). Nonetheless, we argue these models warrant weighting given the overall uncertainty and debate surrounding aging corrections. The ensemble CDF has a median of ~M_w6.90 and 95% CI of M_w6.05 to M_w7.85.

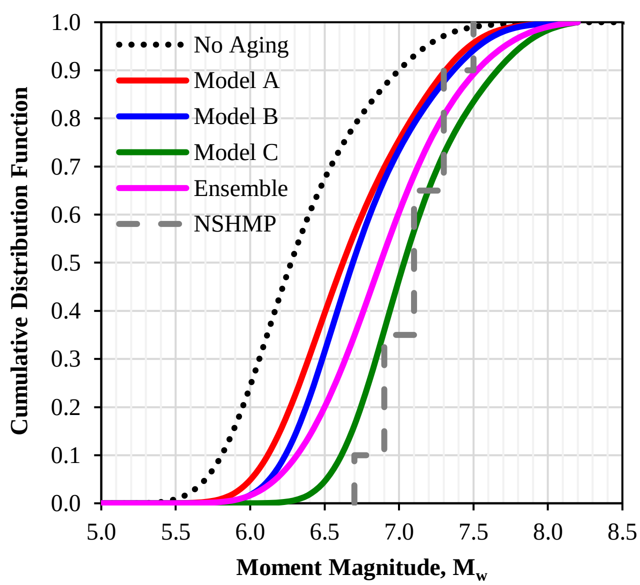


Figure 13. CDFs of rupture magnitude, as inverted from liquefaction observations using three age correction models, wherein seismic sources are modelled as epicenters. NSHMP = National Seismic Hazard Model Project (Petersen et al., 2014, 2020).

5.5 Analysis of Liquefaction Data: Fault Search (Seismic Sources Modelled as Faults)

To determine whether the liquefaction data can constrain the seismic source beyond a point location, the epicenters were next converted to faults. The discretization and treatment of hypothetical fault sources was identical to the MMI analyses. The results without correction for soil aging are shown in Figure 14. It can be seen that the data are incapable of constraining the seismic source to the Woodstock Fault, or to any other hypothetical alignment. The faults deemed most likely to produce the field evidence strike SE (like the predominant result from MMI analysis), but faults with nearly any orientation also have high relative likelihoods. These faults all strike through the general area previously identified via the epicenter search (i.e., through the northern Charleston Peninsula). While faults striking NE do have some likelihood of producing the evidence, these faults deviate from the proposed position of the Woodstock Fault. Results using each of the aging correction models are very similar to those in Figure 14 and are therefore not presented. The use of these models increases the inverted magnitudes (and thus increases the fault lengths in Figure 14), but otherwise has little effect on the inferred fault location or orientation. As with the MMI analyses, this should not be interpreted to mean that the Woodstock Fault or any similar alignment was not the source of the 1886 earthquake. Rather, this should be interpreted to mean that many aspects of the causative fault cannot be constrained with the available macroseismic evidence. This evidence is only marginally supportive of the Woodstock Fault's existence, which is not to say it doesn't exist. This conclusion might be different if more liquefaction study sites were available to the analysis, or if the component prediction models were less uncertain.

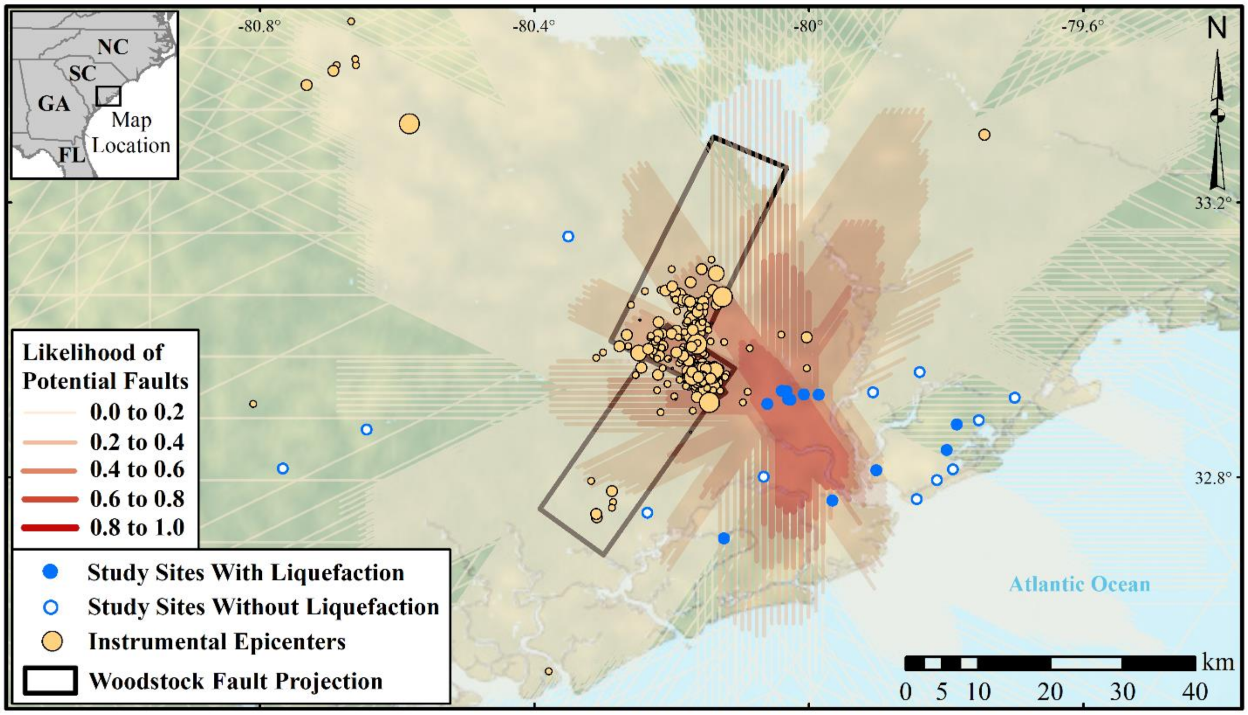


Figure 14. Likelihoods of hypothetical faults with differing orientations and lengths to produce the liquefaction evidence (no aging correction). Black rectangles = Woodstock fault projection (Durá-Gómez and Talwani 2009a,b); tan circles = instrumental epicenters (M_w>1.0; 1900-present) (USGS, 2022).

Shown in Figure 15 are the CDFs of M_w inverted from liquefaction data via the "fault search." The large uncertainty of a magnitude inferred from the liquefaction data is again apparent, as is the influence of soil aging and its correction. It is again emphasized that these results do not include every source of uncertainty. Omitted, for example, are the uncertainties of: (i) fault depth and length; (ii) fault dip; (iii) site V_{S30}; and (iv) the aging-correction factor. For each of these inputs only a median prediction was used. To account for the uncertainty of selecting aging-correction models, each CDF was weighted per the prior scheme to produce an ensemble CDF, as shown in Figure 15. This ensemble has a median of M_w6.95 and 95% CI of M_w6.15 to M_w7.85, suggesting a similar median magnitude as the NSHMP (Petersen et al., 2014, 2020) but much greater uncertainty.

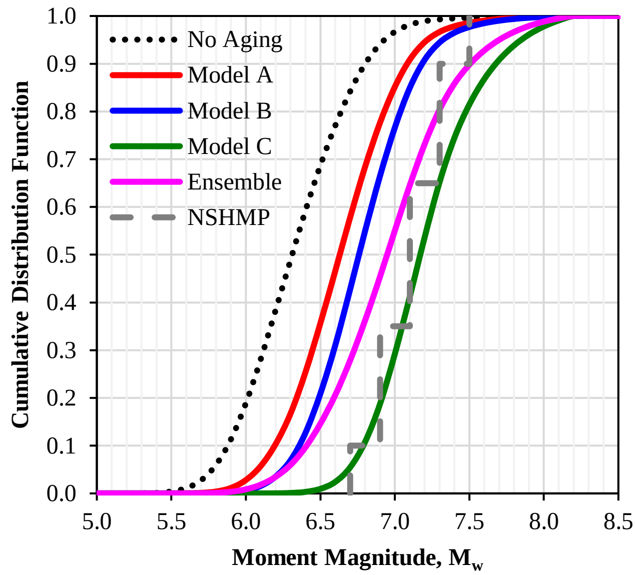


Figure 15. CDFs of rupture magnitude, as inverted from liquefaction observations using three age correction models, wherein seismic sources are modelled as faults. NSHMP = National Seismic Hazard Model Project (Petersen et al., 2014, 2020).

5.6 Analysis of Liquefaction Data: Woodstock Fault

Finally, the preceding analyses were repeated, but conditioned solely on the Woodstock Fault as hypothesized by Durá-Gómez and Talwani (2009a,b). The computed CDFs of M_w, including the weighted ensemble, are shown in Figure 16. Our preferred CDF (i.e., the ensemble) has a median of M_w7.20 and 95% CI of M_w6.30 to M_w8.10. Relative to the CDF conditioned on an unknown source in Figure 15, conditioning on the Woodstock Fault increases the inverted magnitude by ~0.2M_w and increases the M_w uncertainty. This counterintuitive increase in uncertainty is attributable to the epistemic uncertainty of selecting aging-correction models, which is observed to increase with increasing rupture magnitude. Because conditioning on the Woodstock Fault increases the inverted rupture magnitude, differences between the aging-correction models become more apparent. Nonetheless, and similar to the analysis of MMI data, the uncertainties of fault location and geometry appear relatively minor compared to other uncertainties, given the data available for analysis. As shown in Figure 14, faults with diverse orientations were found to have similar likelihoods of producing the liquefaction evidence. In other words, the data are sufficient to constrain the source to a general area but are insufficient to constrain the source to a specific fault alignment. However, because similar magnitudes are inferred for these faults irrespective of their positions, the CDFs conditioned on the Woodstock Fault and the CDFs conditioned on an unknown source are similar.

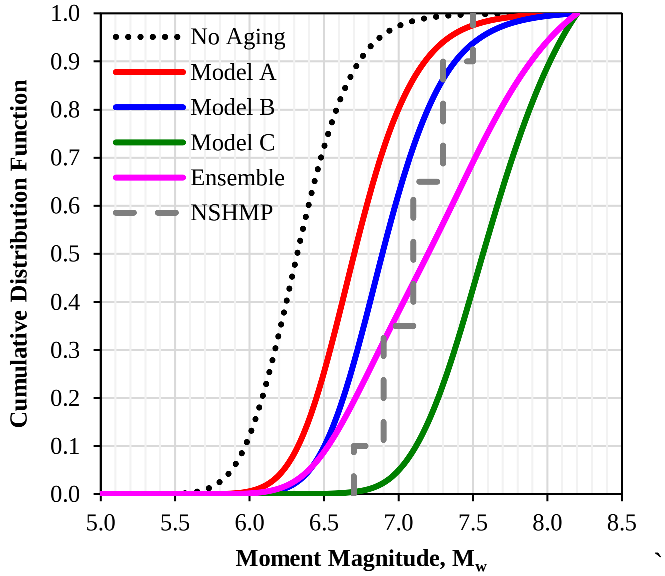


Figure 16. CDFs of rupture magnitude, as inverted from liquefaction observations using three age correction models and assuming the Woodstock Fault (Durá-Gómez and Talwani, 2009a,b) to be the earthquake source. NSHMP = National Seismic Hazard Model Project (Petersen et al., 2014, 2020).

5.7 Forward Predictions of an 1886-Like Rupture

Based on the preceding, which indicated that the intensity and liquefaction data alone cannot well constrain the source of the 1886 event beyond a general region, we forward predict the median PGAs expected in an "1886-like" rupture. As part of this effort, two fault locations and three magnitudes are considered. The purpose of these predictions is to: (i) further illustrate why constraint of the source model is difficult, given data limitations; and (ii) provide predictions for this scenario event using the latest predictive models, which may be useful for regional-scale consequence modelling or science communication. Shown in Figure 17, for example, are predictions considering an M_w7.1 rupture and two strike-slip fault scenarios: (i) fault 1, which is similar to the Woodstock fault; and (ii) fault 2, which is further east, perpendicular to fault 1, and similar to hypothetical faults that were shown to have high likelihoods of producing the MMI and liquefaction evidence (see Figures 9 and 14). Among previously hypothesized sources of the 1886 earthquake, fault 2 is most similar to the Deer Park lineament proposed by Marple and Hurd (2020), although this fault is not as well supported by modern seismicity and is very different from most hypothesized sources. Although the predicted PGAs do differ in the nearfield (e.g., within 50 km of the source), the overall expectation at regional scale is obviously similar for two perpendicular faults. Importantly, and as previously discussed, the Bakun et al. (2002) dataset does not contain an abundance of 1886 MMI observations in the vicinity of Charleston, or in the near field more generally. Only 16 observations, for example, are within 100 km of Charleston. As a result, the predicted PGAs, and by corollary MMIs, are relatively similar at most observation sites.

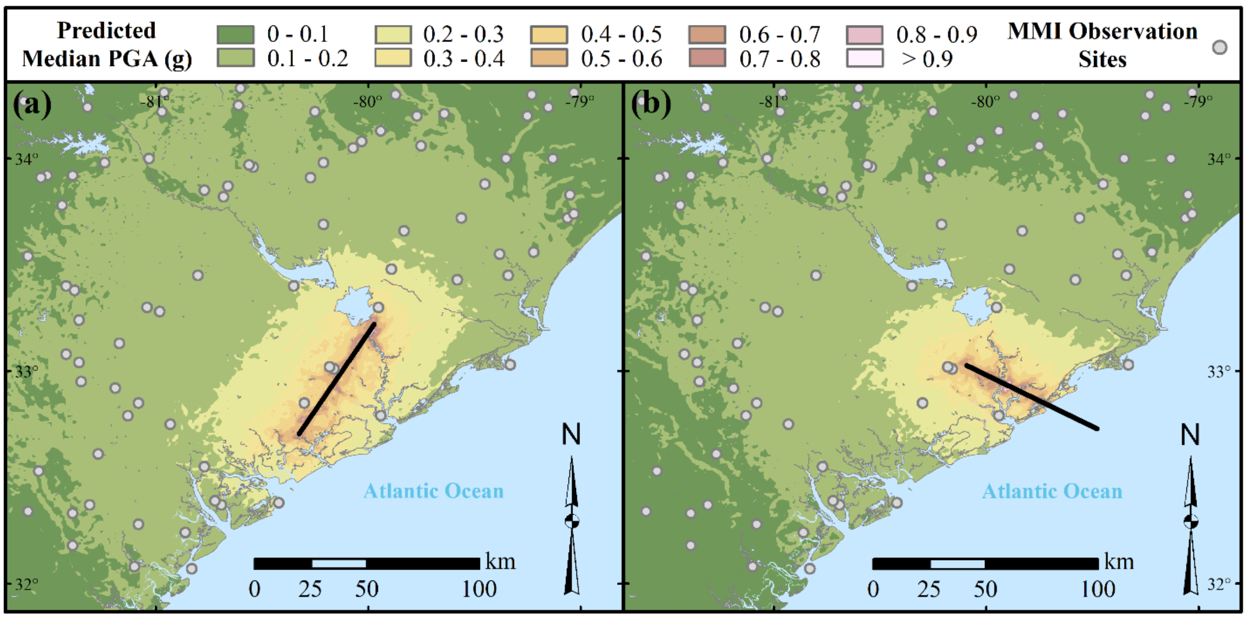


Figure 17. Predictions of median PGA in the Charleston region, as computed using 17 NGA-East GMMs (Goulet et al., 2018) and the Harmon et al. (2019) site-response model, and considering an M_w7.1 rupture on: (a) fault scenario 1; and (b) fault scenario 2, as described in the text.

Plotted in Figure 18a, considering these two fault scenarios, are the predicted MMIs at each of the intensity report locations in the Bakun et al. (2002) dataset. For this example, the Atkinson and Kaka (2007) IM-MMI model was adopted. It can be seen that large differences between the predicted MMIs (say, > 0.5 MMI) are rare. In turn, the difficulty of source-model constraint is apparent, given that the two scenarios lead to perceptible differences at relatively few observation sites. This ambiguity is analogous to that which arises when inverting faults from seismic displacements, wherein the classic "double couple" model requires near-field or supplemental information to solve. Plotted in Figure 18b are the MMI values predicted for fault scenario 1 versus the MMI values observed in 1886; an analogous plot is shown in Figure 18c for fault scenario 2. While scenario 2 fits the data marginally better, as measured by mean absolute error, the fits are essentially the same, as expected. Building off prior results, it cannot be concluded that scenario 1, which approximates the Woodstock Fault, fits the MMI data better than alternative scenarios.

While a version of Figure 18a at sites of liquefaction evidence would show larger differences between expected MMIs (given that all such sites are in the near field), the large finite-sample uncertainty of this smaller dataset leads to a similar outcome, with a variety of faults having relatively high likelihoods of producing the evidence (see Figure 14). It follows that additional near-field evidence (whether MMI or liquefaction) could be especially influential to future studies of the 1886 earthquake. Following the same approach, predictions were made for M_w6.6 and M_w7.6 ruptures. These results are shown in the electronic supplement, where each of the forward predictions is also provided as a GIS map package.

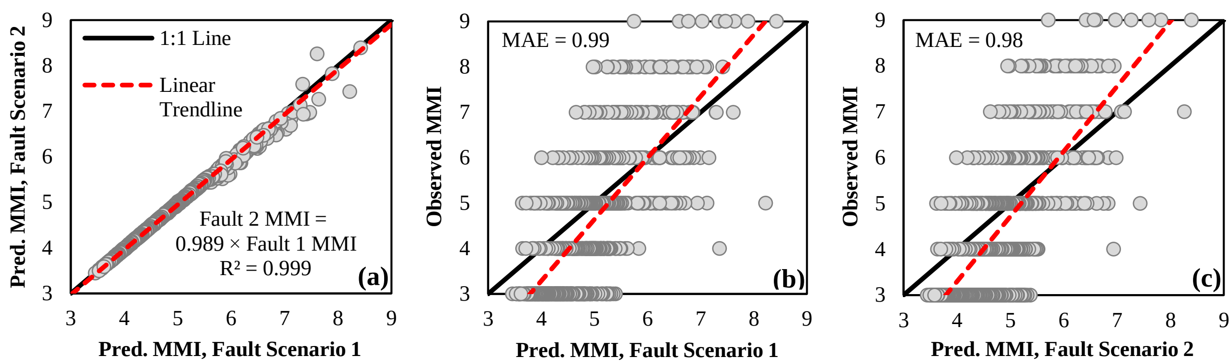


Figure 18. MMI values predicted by the Atkinson and Kaka (2007) model at sites of MMI observations: (a) fault scenario 1 predictions vs. fault scenario 2 predictions; (b) predictions vs. observations for fault scenario 1; (c) predictions vs. observations for fault scenario 2. MAE = mean absolute error.

5.8 Limitations and Uncertainties

The findings presented herein are inherently tied to the field evidence currently available and adopted for analysis. Reinterpretation or augmentation of these observations, which we assume to be independent events, would potentially change these findings, as would the adoption of new component models (e.g., to predict ground motions, site response, MMI, liquefaction triggering, or liquefaction manifestation). Undoubtedly, the modeling of these phenomena will continually advance, warranting future analyses of the 1886 Charleston earthquake. ENA ground-motion modeling has advanced, for example, yet GMMs remain especially uncertain and untested at large magnitudes and will continue to evolve. And, as a supplement to empirical GMMs, physics-based ground motion simulations could provide new insights into the 1886 rupture via more explicit modeling of influential site, path, and source effects (e.g., stress drop), as has been shown for other historical earthquakes (e.g., Lozos, 2016). While empirical GMMs are ubiquitous in earthquake science/engineering and implicitly account for many salient effects, a more explicit accounting could help to reduce uncertainty. Moreover, the goal of this study was to assess the degree to which macroseismic evidence can constrain the 1886 source model (i.e., rupture magnitude and position), rather than to perform a broad investigation of all geophysical and modern seismological evidence, which would fall under the purview of other investigators with different expertise. As one example, our inversions of rupture magnitude conditioned on an unknown source allowed for randomly oriented fault sources. However, with support of the interpreted regional stress regime, it may be reasonable to limit the range of orientations considered, or alternatively, to subjectively assign weights to each orientation, thereby increasing or decreasing the computed likelihoods, all else being equal. These possibilities could be investigated in the future. In addition, it should be noted that many, but not all, sources of uncertainty were accounted for in the analyses. Neglected, for example, were the uncertainties of site characterization (i.e., V_S30) and site observations (i.e., MMI and liquefaction responses). The inclusion of these and other uncertainties could potentially broaden the M_w CDFs computed herein. It must similarly be emphasized that a thorough reinterpretation of the more than 1000 original intensity reports, to include the assignment of site-specific uncertainties, bias corrections, and/or weighting schemes, could potentially change the results we present. And, as previously emphasized, our ensemble M_w CDFs utilize judgement-based weighting schemes. While justifications were provided for these weights, readers might argue for other weights, and thus draw other conclusions from our suites of M_w CDFs. Ultimately, future analyses will confirm or revise the conclusions reached in this study and summarized below.

6. Conclusions

Prior studies of MMI and liquefaction data resulting from the 1886 Charleston, SC, earthquake have several limitations. Namely, these studies tend to: (i) either be deterministic or account for uncertainties in an informal manner (e.g., it is often unclear what published uncertainty bounds represent and which uncertainties are, and are not, included); (ii) assume that the 1886 event was caused by a particular fault (i.e., the Woodstock Fault) without investigating the uncertainty of this assumption or the ability of the field data to constrain source traits beyond magnitude (i.e., fault location, geometry); and (iii) rely on models for predicting various phenomena (e.g., ground motions, site response, liquefaction response, MMI) that have since been superseded or augmented (e.g., by the NGA East project's 17 GMMs). Accordingly, this study presented probabilistic seismic-source inversions of the 1886 earthquake using the latest predictive models and a novel inversion methodology. With this approach, the likelihood of a rupture with some magnitude, location, and geometry to produce a set of field observations is computed. Repeating for enumerable hypothetical sources results in a regional scale constraint of the likely source traits, to the extent that observational data permits. With this approach, magnitude probability distributions conditioned on both an unknown source and on the hypothesized Woodstock Fault (Durá-Gómez and Talwani, 2009a,b) were computed and compared to that used in the U.S. NSHMP (Petersen et al., 2014, 2020). The most salient findings of this study, subject to the limitations and uncertainties summarized in section 5.8, are:

- Neither the location nor orientation of the 1886 fault rupture could be confidently constrained by the macroseismic evidence and models utilized herein, given the large uncertainties inherent to each. Hypothetical faults in a range of locations and with various alignments were deemed to have high relative likelihoods of producing this evidence. Considering both types of evidence and all analyses, faults striking SE and partly offshore were predominantly identified as having the greatest likelihood. Yet these faults ranged in location and other faults, with very different strikes, were also often found to be relatively likely.
- Collectively, the evidence does not provide strong support for the hypothesized Woodstock Fault. One analysis (that of MMI data using the WEA12 IM-MMI model) found a NE-striking "Woodstock like" fault to be the most likely source of the field evidence. Yet most analyses – in so far as supporting the Woodstock Fault – point to the likelihood of a seismic source somewhere north of Charleston but deem the Woodstock Fault to be relatively unlikely. This is not to say the fault does not exist or could not have been the source of the 1886 rupture (geophysical investigations and modern seismological data clearly suggest that active faults exist in the area), but rather, that many aspects of the 1886 source model cannot be well constrained with the available macroseismic evidence and models. This result might change if: (i) more MMI and liquefaction evidence were available – particularly in the near field; (ii) the MMI evidence was reinterpreted to remove or correct outliers, and to assign observation-specific uncertainties; or (iii) if the various required component models were less uncertain.

- In the absence of these developments, certain aspects of the 1886 fault rupture can only be constrained with other seismologic, geologic, and/or coseismic data and interpretation.
- When conditioned on the Woodstock Fault proposed by Durá-Gómez and Talwani (2009a,b), and generally supported by others, our preferred CDF of M_w inverted from MMI data has a median of M_w7.10 and 95% CI of M_w7.0 to M_w7.75. Of all uncertainties considered, the epistemic uncertainty of IM-MMI model selection was larger than any other, since different models may give significantly different MMI predictions for a given IM. In this regard, efforts to better quantify the suitability of IM-MMI models to the study region, and in turn, to refine the weighting scheme used herein, could have a significant influence on our overall conclusions.
- The results from MMI analysis show some site-to-source distance bias, with magnitudes inverted from more distal MMI observations tending to be larger. Possible reasons for this bias were discussed but could be further investigated in the future.
- When conditioned on the Woodstock Fault, our preferred CDF of M_w inverted from liquefaction data has a median of M_w7.20 and 95% CI of M_w6.30 to M_w8.10. The greater M_w uncertainty, relative to that from MMI analysis, is attributable to there being fewer liquefaction study sites, the compilation of which requires both an observation of liquefaction response *and* subsurface geotechnical testing. In addition, the uncertainties of whether and how to correct for soil-aging effects considerably augmented the uncertainty. As such, efforts to quantify the suitability and uncertainties of aging-correction models could have significant influence on our overall conclusions.
- When conditioned on an unknown seismic source, the CDFs of M_w inverted from MMI and liquefaction data did not differ greatly from the CDFs conditioned on the Woodstock Fault. Our interpretation is that while faults with a range of locations and alignments were found similarly likely to produce the field evidence, these faults were inferred to have mostly similar magnitude distributions. Thus, the results of this study pertaining to the magnitude of the 1886 rupture would not necessarily change if the rupture's exact position was known. This should not be interpreted to mean that the uncertainties of source location and geometry are inconsequential to inverse-analyses of macroseismic data. As demonstrated via simulated inversions of modern earthquakes (Rasanen and Maurer, 2021, 2022), these uncertainties are often considerable. This is especially the case, for example, when the effects of a distant,

large M_w rupture cannot be distinguished from the effects of a nearby, small M_w rupture. In such cases, constraint of the source location can dramatically reduce the overall uncertainty of the inverted magnitude.

- Collectively, the results largely support the M_w distribution adopted by the NSHMP, which ranges from M_w6.7 to M_w7.5 with a median of M_w7.1. While analyses indicate that $M_w < 6.7$ ruptures have potential to produce the observed liquefaction response (particularly when the uncertainty of selecting an aging-correction model is considered), the MMI evidence suggests a near-zero likelihood of such magnitudes. Conversely, both types of evidence suggest that $M_w > 7.5$ events have potential to produce the field evidence. While this conclusion hinges on which component models are adopted to predict soil-aging effects and MMI, we fail to find conclusive evidence for outright rejecting some models in favor of others. In the absence of such evidence, the possibility of $M_w > 7.5$ ruptures would merit consideration.
- Ultimately, the flexible inversion methodology employed herein is not specific to ENA, or to certain types of macroseismic evidence, but rather is applicable to any seismic zone and to any co-seismic response for which probabilistic prediction models exist. This methodology allows for uncertainty to be accounted for in a more complete and transparent manner when inverting seismic parameters from macroseismic evidence.

Declaration of competing interests

The authors declare no competing interests.

Acknowledgements

The presented study is based on work supported by the National Science Foundation (NSF) under Grant No. CMMI-1751216 and by the NSF Graduate Research Fellowship Program under Grant No. DGE-1762114. However, any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF. We also gratefully acknowledge Silvia Mazzoni of the John Garrick Institute for Risk Sciences at UCLA for developing NGA East ground motion characterization tools based on the results of Goulet et al. (2018), which were utilized in this study.

Appendix A. Supplementary data

All data and models utilized in this study are publicly available, as cited herein. The supplemental material includes 13 additional figures that largely pertain to the analysis of MMI data using the "epicenter search" and "fault search" methods. The results presented in these figures are summarized and discussed in the main text. Also included are maps (figures and GIS files) of the predicted median PGAs in an "1886-like" event, considering three magnitudes and two fault alignments, as well as measurements from 24 CPTs.

REFERENCES

Amick, D., G. Maurath, and R. Gelinas (1990). Characteristics of seismically induced liquefaction sites and features located in the vicinity of the 1886 Charleston, South Carolina earthquake, *Seismological Research Letters*, **61**, no. 2, 117-211.

Andrus R.D., H. Hayati, and N.P. Mohanan (2009). Correcting liquefaction resistance for aged sands using measured to estimated velocity ratio, *J. Geotech. Geoenviron. Eng.*, **135**, no. 6, 735-744, doi: 10.1061/(ASCE)GT.1943-5606.0000025

Atkinson G.M., and S.I. Kaka (2007). Relationships between felt intensity and instrumental ground motion in the central United States and California, *Bull. Seismol. Soc. Am*. **97**, no. 2, 497-510, doi: 10.1785/0120060154

Atkinson, G.M., C.B. Worden, and D.J. Wald (2014). Intensity prediction equations for North America, *Bulletin of the Seismological Society of America*, **104**, no. 6, 3084-3093, doi: 10.1785/0120140178

Bakun W.H., and M.G. Hopper (2004). Magnitudes and locations of the 1811-1812 New Madrid, Missouri, and the 1886 Charleston, South Carolina, earthquakes, *Bulletin of the Seismological Society of America*, **94**, no. 1, 64-75, doi: 10.1785/0120020122.

Bakun W.H., A.C. Johnston, and M.G. Hopper (2002). Modified Mercalli Intensities (MMI) for large earthquakes near New Madrid, Missouri, in 1811-1812 and near Charleston, South Carolina, in 1886, *U.S. Geol. Surv.* Open File Rep. 02-184.

Boller, R. (2008). Geotechnical investigations at three sites in the South Carolina coastal plain that did not liquefy during the 1886 Charleston earthquake (*Master's thesis*). Available from https://tigerprints.clemson.edu/all_theses/301/

Bollinger, G.A. (1977). Reinterpretation of the intensity data for the 1886 Charleston, South Carolina, earthquake, *U.S. Geol. Surv.* Profess. Pap. 1028, 17–32.

Bollinger, G.A. and C.W. Stover (1976). List of intensities for the 1886 Charleston, South Carolina earthquake, *U.S. Geol. Surv.* Open-File Rept. 76-66, 31pp.

Bommer, J.J., B. Dost, B. Edwards, P.J. Stafford, J. van Elk, D. Doornhof, and M. Ntinalexis (2016). Developing an application-specific ground motion model for induced seismicity, *Bull. Seismol. Soc. Am.* **106**, 158–173, doi: 10.1785/0120150184

Boulanger R.W., and I.M. Idriss (2014). CPT and SPT based liquefaction triggering procedures. Report no. UCD/CGM-14/01. Davis, CA: Center for Geotechnical Modeling, University of California, Davis.

Bwambale, B., Andrus, R. D., Heidari, T., Gathro, J., & Cramer, C. H. (2022). Influence of source-to-site distance and diagenesis on liquefaction triggering of 200,000-year-old beach sand. *Engineering Geology*, *298*, 106557, 10.1016/j.enggeo.2022.106557.

Caprio M., B. Tarigan, C.B. Worden, S. Wiemer, and D.J. Wald (2015). Ground motion to intensity conversion equations (GMICEs): A global relationship and evaluation of regional dependency, *Bull. Seismol. Soc. Am.* **105**, no. 3, 1476-1490, doi: 10.1785/0120140286

Chapman, M.C., M.S. Sibol, and G. A. Bollinger (1989). Investigation of anomalous earthquake intensity levels along the coastal-plain-Piedmont boundary in South Carolina and Georgia, *Virginia Tech Seismological Observatory Report* 1360.

Chapman, M.C., J.N. Beale, A.C. Hardy, and Q. Wu (2016). Modern seismicity and the fault responsible for the 1886 Charleston, South Carolina, earthquake, *Bulletin of the Seismological Society of America*, **106**, no. 2, 364-372, doi: 10.1785/0120150221.

Chung, J., Okok, A., and Rogers, J. D. (2021). Geologic impacts and calculated magnitudes of historic earthquakes in the central United States. *Engineering Geology*, *280*, 105923, doi: 10.1016/j.enggeo.2020.105923.

Coppersmith, K. J., Salomone, L. A., Fuller, C. W., Glaser, L. L., Hanson, K. L., Hartleb, R. D., ... & Tuttle, M. P. (2012). Central and eastern United States (CEUS) seismic source characterization (SSC) for nuclear facilities project (No. DOE/NE-0140). Electric Power Research Institute (EPRI).

Cramer, C.H. (2020). Updated GMICE for central and eastern North America extending to higher intensities, *Seimol. Res. Lett.* **91**, no. 6, 3518-3527, doi: 10.1785/0220200061

Cramer, C.H., and O.S. Boyd (2014). Why the New Madrid earthquakes are **M** 7-8 and the Charleston earthquake is ~**M** 7, *Bulletin of the Seismological Society of America*, **104**, no. 6, 2884-2903, doi: 10.1785/0120120257

de Magistris, F.S., G. Lanzano, G. Forte, and G. Fabbrocino (2013). A database for PGA threshold in liquefaction occurrence, *Soil Dynamics and Earthquake Engineering*, **54**, 17-19, doi: 10.1016/j.soildyn.2013.07.011

Durá-Gómez, I., and P. Talwani (2009a). Finding faults in the Charleston area, South Carolina: 1. seismological data. *Seismological Research Letters* **80**, no. 5, 883–900, doi: 10.1785/gssrl.80.5.883

Durá-Gómez, I., and P. Talwani (2009b). Finding faults in the Charleston area, South Carolina: 2. complementary data. *Seismological Research Letters* **80**, no. 5, 901–919, doi: 10.1785/gssrl.80.5.901

Dutton, C.E. (1889). The Charleston earthquake of August 31, 1886: *U.S. Geological Survey*, Ninth Annual Report 1887-88, 203-528.

Gheibi, E., & Gassman, S. L. (2016). Application of GMPEs to estimate the minimum magnitude and peak ground acceleration of prehistoric earthquakes at Hollywood, SC. *Engineering Geology*, *214*, 60-66, doi: 10.1016/j.enggeo.2016.09.016.

Geiger, A. (2010). Liquefaction analysis of three Pleistocene sand deposits that did not liquefy during the 1886 Charleston, South Carolina earthquake based on shear wave velocity and penetration resistance (*Master's thesis*). Available from https://tigerprints.clemson.edu/all_theses/815/

Geyin M. and B.W. Maurer (2020). Fragility functions for liquefaction-induced ground failure, *Journal of Geotechnical and Geoenvironmental Engineering* **146**, no. 12, doi: 10.1061/(ASCE)GT.1943-5606.0002416

Gheibi, E., S. Gassman, and P. Talwani (2020). Regional assessment of prehistoric earthquake magnitudes in the South Carolina Coastal Plain, *Bulletin of Engineering Geology and the Environment*, **79**, no. 3, 1413-1427, doi: 10.1007/s10064-019-01627-7

Gohn, G.S., R.E. Weems, S.F. Obermeier, and R.L. Gelinas (1984). Field studies of earthquake-induced liquefaction-flowage features in the Charleston, South Carolina area: preliminary report: *U.S. Geological Survey Open-File Report* 84-670, 26 p.

Goulet, C., Y. Bozorgnia, N. Abrahamson, N. Kuehn, L. Al Atik, R. Youngs, R.W. Graves, G.M. Atkinson (2018). Central and eastern North America ground-motion characterization (NGA-East Final Report, PEER Report Number 2018/08). Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California.

Green, R.A., Obermeier, S.F., and Olson, S.M. (2005). Engineering geologic and geotechnical analysis of paleoseismic shaking using liquefaction effects: field examples. *Engineering Geology*, **76**, no. 3-4, 263-293. doi: 10.1016/j.enggeo.2004.07.026

Green, R.A., J.J. Bommer, A. Rodriguez-Marek, B.W. Maurer, P.J. Stafford, B. Edwards, … J., Van Elk (2019). Addressing limitations in existing 'simplified' liquefaction triggering evaluation procedures: application to induced seismicity in the Groningen gas field. *Bulletin of Earthquake Engineering*, **17**, no. 8, 4539-4557. doi: 10.1007/s10518-018-0489-3

Harmon, J., Y.M.A., Hashash, J.P. Stewart, E.M. Rathje, K.W. Campbell, W.J. Silva, and O. Ilhan (2019). Site Amplification Functions for Central and Eastern North America – Part II: Modular Simulation-Based Models, *Earthquake Spectra*, **35**, no. 2, 815-847, doi: 10.1193/091117EQS179M

Hasek, M.J. (2016). Age-related liquefaction resistance of Pleistocene coastal plain sands in South Carolina (*Doctoral dissertation*). Retrieved from https://scholarcommons.sc.edu/etd/3600.

Hayati, H. and R.D. Andrus (2008). Liquefaction potential map of Charleston, South Carolina based on the 1886 earthquake, *Journal of Geotechnical and Geoenvironmental Engineering*, **134**, no. 6, 815-828, doi: 10.1061/(ASCE)1090-0241(2008)134:6(815)

Hayati, H., and R.D. Andrus (2009). Updated liquefaction resistance correction factors for aged soils, *Journal of Geotechnical and Geoenvironmental Engineering*, **135**, no. 11, 1683–1692, doi: 10.1061/(ASCE)GT.1943-5606.0000118

Heidari, T., and R.D. Andrus (2010). Mapping liquefaction potential of aged soil deposits in Mount Pleasant, South Carolina, *Eng. Geol.* **112**, 1–12, doi: 10.1016/j.enggeo.2010.02.001

Heidari, T., and R.D. Andrus (2012). Liquefaction potential assessment of Pleistocene beach sands near Charleston, South Carolina, *J. Geotech. Geoenviron. Eng.* **138**, no. 10, 1196-1208, doi: 10.1061/(ASCE)GT.1943-5606.0000686

Heath, D.C., D.J. Wald, C.B. Worden, E.M. Thompson and G.M. Smoczyk (2020) A global hybrid Vs30 map with a topographic slope-based default and regional map insets. *Earthquake Spectra* **36**, 1570–1584, doi: 10.1177/8755293020911137

Hough, S. E., Armbruster, J. G., Seeber, L., and Hough, J. F. (2000). On the modified Mercalli intensities and magnitudes of the 1811–1812 New Madrid earthquakes. *Journal of Geophysical Research: Solid Earth*, **105**(B10), 23839-23864.

Hu, K., S.L. Gassman, and P. Talwani (2002). Magnitudes of prehistoric earthquakes in the South Carolina coastal plain from geotechnical data, *Seismological Research Letters*, **73**, no. 6, 979-991, doi: 10.1061/9780784479087.112

Ishihara, K. (1985). Stability of natural soil deposits during earthquakes, *Proc. llth Conf. on Soil Mech. and Found. Engrg.*, International Society of Soil Mechanics and Foundation Engineers, San Francisco, Calif., 321-376.

Iwasaki T., F. Tatsuoka, K. Tokida and S. Yasud (1978) A practical method for assessing soil liquefaction potential based on case studies at various sites in Japan. In: Proceedings of the 2nd international conference on microzonation for safer construction research and application, (2): 885–896, 30 Oct– 3 Nov, Seattle, Washington.

Johnston, A.C. (1996). Seismic moment assessment of stable continental earthquakes - III. New Madrid 1811–1812, Charleston 1886 and Lisbon 1755, *Geophys. J. Int.*, **126**, no. 2, 314–344, doi: 10.1111/j.1365-246X.1996.tb05294.x

Kaklamanos J., L.G. Baise and D.M. Boore (2011). Estimating unknown input parameters when implementing the NGA ground motion prediction equations in engineering practice, *Earthquake Spectra* **27**, no. 4, 1219–1235, doi: 10.1193/1.3650372

Kuhn, G. G. (2005). Paleoseismic features as indicators of earthquake hazards in North Coastal, San Diego County, California, USA. *Engineering geology*, *80*(1-2), 115-150, doi: 10.1016/j.enggeo.2005.04.006.

Lozos, J.C. (2016). A case for historic joint rupture of the San Andreas and San Jacinto faults. *Science advances*, **2**(3), e1500621.

Marple, R. T., & Hurd, J. D. (2020). Interpretation of lineaments and faults near Summerville, South Carolina, USA, using LiDAR data: Implications for the cause of the 1886 Charleston, South Carolina, earthquake. *Atlantic Geology: Journal of the Atlantic Geoscience Society/Atlantic Geology: revue de la Société Géoscientifique de l'Atlantique*, *56*, 73-95, doi: 10.4138/atlgeol.2020.003.

Martin, J.R., II, and G.W. Clough (1994). Seismic parameters from liquefaction evidence, *J. Geotech. Engrg.*, **120**, no. 8, 1345-1361, doi: 10.1061/(ASCE)0733-9410(1994)120:8(1345)

Maurer, B.W., R.A. Green, M. Cubrinovski, and B.A. Bradley (2014). Assessment of aging correction factors for liquefaction resistance at sites of recurrent liquefaction. In *Proceedings of the 10th US National Conference on Earthquake Engineering*.

Maurer, B. W., Green, R. A., Quigley, M. C., & Bastin, S. (2015a). Development of magnitude-bound relations for paleoliquefaction analyses: New Zealand case study. *Engineering Geology*, *197*, 253-266, doi: 10.1016/j.enggeo.2015.08.023.

Maurer, B.W., R.A. Green, and O.D.S Taylor (2015b). Moving towards an improved index for assessing liquefaction hazard: lessons from historical data. *Soils and Foundations*, **55**, no. 4, 778-787. doi. 10.1016/j.sandf.2015.06.010

Nuttli, O.W., G.A. Bollinger, and R.B. Herrmann (1986). The 1886 Charleston, South Carolina, earthquake: A 1986 perspective, *U.S. Geological Survey*, **985**, doi: 10.3133/cir985

Obermeier, S. F. (1998). Liquefaction evidence for strong earthquakes of Holocene and latest Pleistocene ages in the states of Indiana and Illinois, USA. *Engineering Geology*, *50*(3-4), 227-254, doi: 10.1016/S0013-7952(98)00032-5.

Obermeier, S.F., G.S. Gohn, R.S. Weems, R.L. Gelinas, and M. Rubin (1985). Geologic Evidence for Recurrent Moderate to Large Earthquakes Near Charleston, South Carolina, *Science*, **277**, no. 4685, 408-411.

Obermeier, S. F., Olson, S. M., & Green, R. A. (2005). Field occurrences of liquefaction-induced features: a primer for engineering geologic analysis of paleoseismic shaking. *Engineering Geology*, *76*(3-4), 209-234, doi: 10.1016/j.enggeo.2004.07.009.

Obermeier, S. F., R.E. Weems, and R.B. Jacobson, (1987). Earthquake induced liquefaction features in the coastal South Carolina region, *U.S. Geol. Surv. Open File Report* 87-504.

Olson, S. M., Green, R. A., & Obermeier, S. F. (2005). Geotechnical analysis of paleoseismic shaking using liquefaction features: a major updating. *Engineering Geology*, *76*(3-4), 235-261, doi: 10.1016/j.enggeo.2004.07.008.

Petersen, M.D., M.P. Moschetti, P.M. Powers, C.S. Mueller, K.M. Haller, A.D. Frankel, … A.H. Olsen (2014). Documentation for the 2014 update of the United States national seismic hazard maps. *U.S. Geol. Surv. Open-File Report* 2014–1091.

Petersen, M.D., A.M. Shumway, P.M. Powers, C.S. Mueller, M.P. Moschetti, A.D. Frankel, … Y. Zeng (2020). The 2018 update of the US National Seismic Hazard Model: Overview of model and implications, *Earthquake Spectra*, **36**, no. 1, 5-41, doi: 10.1177/8755293019878199

Pratt, T. L., Shah, A. K., Counts, R. C., Horton, J. W., & Chapman, M. C. (2022). Shallow Faulting and Folding in the Epicentral Area of the 1886 Charleston, South Carolina, Earthquake. *Bulletin of the Seismological Society of America*, doi: 10.1785/0120210329.

Rasanen, R.A., and B.W. Maurer (2021). Probabilistic seismic source inversion from regional landslide evidence. *Landslides*, doi: 10.1007/s10346-021-01780-9

Rasanen, R.A., and B.W. Maurer (2022). Probabilistic seismic source location and magnitude via inverse analysis of paleoliquefaction evidence. *Earthquake Spectra*, **38**, no. 2, 1499-1528. doi: 10.1177/87552930211056355

Rasanen, R. A., Marafi, N. A., & Maurer, B. W. (2021). Compilation and forecasting of paleoliquefaction evidence for the strength of ground motions in the US Pacific Northwest. *Engineering Geology*, *292*, 106253, doi: 10.1016/j.enggeo.2021.106253.

Robertson, P.K., and C.E. Wride (1998). Evaluating cyclic liquefaction potential using the cone penetration test. *Canadian Geotechnical Journal*, ***35***, no. 3, 442-459. doi: 10.1139/t98-017

Rodriguez-Marek, A., & Ciani, D. (2008). Probabilistic methodology for the analysis of paleoliquefaction features. *Engineering geology*, *96*(3-4), 159-172, 10.1016/j.enggeo.2007.10.007.

Scherbaum F., J. Schmedes and F. Cotton (2004). On the conversion of source-to-site distance measures for extended earthquake source models. *Bulletin of the Seismological Society of America* **94**, no. 3, 1053–1069, doi: 10.1785/0120030055

Schneider, J. A., Mayne, P. W., & Rix, G. J. (2001). Geotechnical site characterization in the greater Memphis area using cone penetration tests. *Engineering Geology*, *62*(1-3), 169-184, doi: 10.1016/S0013-7952(01)00060-6.

Seed, H.B., K. Tokimatsu, L.F. Harder, and R.M. Chung (1984). The influence of SPT procedures in soil liquefaction resistance evaluations, *Report No. UBC/EERC-84/15,* Earthquake Engineering Research Center, University of California, Berkeley, Calif.

Silva, W., I. Wong, T. Siegel, N. Gregor, R. Darragh, and R. Lee (2003). Ground motion and liquefaction simulation of the 1886 Charleston, South Carolina, earthquake, *Bulletin of the Seismological Society of America*, ***93***, no. 6, 2717-2736, doi: 10.1785/0120030029

Stover, C.W. and J.L. Coffman (1993). Seismicity of the United States, 1568-1989 (Revised), *U.S. Geological Survey* Prof. Pap. 1527, 418pp.

Talwani, P., and J. Cox (1985). Paleoseismic evidence for recurrence of earthquakes near Charleston, South Carolina, *Science*, **228**, no. 4711, 379-381, doi: 10.1126/science.229.4711.379

Talwani, P., M. Hasek, S. Gassman, W.R. Doar III, and A. Chapman (2011). Discovery of a sand blow and associated fault in the epicentral area of the 1886 Charleston earthquake, *Seismological Research Letters*, **82**, no. 4, 589-598, doi: 10.1785/gssrl.82.4.589

Talwani, P., and W.T. Schaeffer (2001). Recurrence rates of large earthquakes in the South Carolina coastal plain based on paleoliquefaction data, *Journal of Geophysical Research*, **106**, no. B4, 6621-6642, doi: 10.1029/2000JB900398

Thompson, E. M., and C.B. Worden (2018). Estimating rupture distances without a rupture. *Bulletin of the Seismological Society of America*, **108**, no. 1, 371-379, doi: 10.1785/0120170174

USGS (2021). Cone Penetration Test (CPT) Map and Database. U.S. Geological Survey, Earthquake Hazards Program. https://earthquake.usgs.gov/research/cpt/data/

USGS (2022). Advanced National Seismic System (ANSS) Comprehensive Catalog of Earthquake Events and Products: Various. U.S. Geological Survey, Earthquake Hazards Program. doi: 10.5066/F7MS3QZH

Weems, R.E., W.C. Lewis, and E.M. Lemon, Jr. (2014). Surficial geologic map of the Charleston region, Berkeley, Charleston, Colleton, Dorchester, and Georgetown Counties, South Carolina, *U.S. Geological Survey* Open-File Report 2013–1030, 1 sheet, scale 1:100,000, doi: 10.3133/ofr20131030

Wells D.L., and K.J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull. Seismol. Soc. Am*. **84**, no. 4, 974–1002, doi: 10.1785/BSSA0840040974

Werner, M. J., and D. Sornette (2008). Magnitude uncertainties impact seismic rate estimates, forecasts, and predictability experiments, J. Geophys. Res.*,* ***113****, no. B8, doi:* 10.1029/2007JB005427

Williamson, J.R., and S.L. Gassman (2014). Identification of liquefiable coastal plain soils using DMT, SPT, and CPT profiles, *Geo-Congress 2014*, February 23-26, Atlanta, Georgia, doi: 10.1061/9780784413272.206

Wong, I., J. Bouabid, W. Graf, C. Huyck, A. Porush, W. Silva, T. Siegel, G. Bureau, R. Eguchi, and J. Knight (2005). Potential losses in a repeat of the 1886 Charleston, South Carolina, earthquake, *Earthquake Spectra*, **21**, no. 4, 1157-1184, doi: 10.1193/1.2083907.

Wood, H.O., and F. Neumann (1931). Modified Mercalli Intensity scale of 1931, *Bull. Seismol. Soc. Am*. **21**, 277-283.

Worden, C.B., M.C. Gerstenberger, D.A. Rhoades, and D.J. Wald (2012). Probabilistic relationships between ground-motion parameters and Modified Mercalli Intensity in California, *Bull. Seismol. Soc. Am*. **102**, no. 1, 204-221, doi: 10.1785/0120110156.

Yousuf, M., Bukhari, S. K., & Bhat, G. R. (2021). Using paleo-liquefaction features to determine the likely source, magnitude and ground accelerations of pre-historic earthquakes in the Kashmir Basin (Northwestern Himalaya), India. *Engineering Geology*, *293*, 106302, 10.1016/j.enggeo.2021.106302.

APPENDIX E

Publication 5 (Journal Paper 4), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Soil liquefaction in an M9 Cascadia Subduction Zone Earthquake: Looking Forwards and Backwards." *In Preparation.*	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.
[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest

Ryan A. Rasanen[f], Nasser A. Marafi[2], and Brett W. Maurer[1]

Abstract: In the U.S. Pacific Northwest (PNW), the historic earthquake record is often insufficient to provide inputs to seismic-hazard analyses or to inform ground-motion predictions for certain seismic sources (e.g., the Cascadia Subduction Zone, CSZ). As a result, paleoseismic studies are commonly used to infer information about the seismic hazard. However, among the many forms of coseismic evidence, soil liquefaction provides the best, if not only, evidence from which the intensities of previous ground motions may be constrained. Accordingly, the overarching goal of this research is to use paleoliquefaction to elucidate previous ground motions in the PNW – both for CSZ events and others – and to further constrain the locations, magnitudes, and recurrence rates of such ruptures. Towards that goal, this paper: (i) reviews current paleoliquefaction inverse-analysis methods and their limited, prior applications in the PNW; (ii) compiles all PNW paleoliquefaction evidence from the literature into a GIS database, resulting in data from 185 study sites (e.g., feature locations, types, sizes, and ages); and (iii) develops maps – specific to the CSZ – that forecast paleoliquefaction for 30 different simulations of a CSZ event. These maps can be used to guide field explorations for new evidence, such that they are conducted efficiently and strategically, considering the apparent utility of evidence toward constraint of CSZ ground-motion models. Of additional utility, this process provides regional ground-motion predictions for physics-based simulations of an M9 event, to include expected site effects. Collectively, the maps of expected shaking intensity and liquefaction may be useful in downstream hazard modelling, regional loss estimation, policy development, and science communication. Ultimately, as more paleoliquefaction evidence is identified and studied, better constraint of regional ground-motion hazards will result.

1. Introduction

1.1 Significance of Paleoliquefaction Evidence

In regions experiencing infrequent moderate-to-large earthquakes, the historic record may be insufficient to provide accurate inputs to seismic-hazard analyses (i.e., the locations, magnitudes, and recurrence-rates of fault ruptures) or to inform ground-motion predictions for certain seismic sources. As an example, the 1700 A.D. Cascadia Subduction Zone (CSZ) earthquake was likely far bigger than

---

[f] Department of Civil and Environmental Engineering, University of Washington, Seattle USA.
[2] Risk Management Solutions, Newark, California USA.

any subsequent rupture in the U.S. Pacific Northwest (PNW), yet there is little-to-no eyewitness account of it (Thrush and Ludwin 2007), let alone ground-motion records. As a result, paleoseismic evidence must be relied on to elucidate the seismic hazard. In the case of the CSZ, the 1700 earthquake is believed to be one of approximately 40 similar events during the past 10,000 years, all judged to be M8.0 to M9.0 or greater, as inferred from dendrochronology (e.g., Atwater et al., 1991), turbidites (e.g., Goldfinger et al., 2012; Atwater et al., 2014), tsunami deposits (e.g., Peters et al., 2007), soil liquefaction (e.g., Obermeier 1995), microfossils (e.g., Engelhart et al., 2013), geochemical markers (e.g., O'Donnell et al., 2017), and seafloor morphology (e.g., Watt et al., 2017), among other indicators of seismicity. Collectively, this evidence has been used to infer a length of fault rupture, which leads to an estimate of earthquake magnitude (e.g., Petersen et al., 2014).

Given the premise of a full-fault M9 CSZ earthquake, various ground-motion predictions have been made (e.g., USGS, 2017). These include – most recently – a suite of broadband synthetic seismograms (i.e., ground-motion time histories) (Frankel et al., 2018a; Wirth et al., 2018) that predict motions on a 1-km grid across the PNW. This suite includes 30 different realizations to reflect the uncertainty of key parameters (e.g., the down-dip limit of fault rupture; the slip distribution and location of asperities; and hypocenter location). Given that the last CSZ event occurred in 1700, these parameters are unknown for all such events. Shown in Fig. 1 are two such realizations. While both simulate a full-fault M9 CSZ rupture, it can be seen that predicted ground-motions vary significantly in some locales (e.g., peak ground velocities may vary by 400%). By corollary, the expected impacts on the built and living environments would also be very different (Marafi et al., 2019, 2020).

Notably, the actual ground motions experienced in 1700 (and in other paleo events) can be determined only through inverse analysis of coseismic evidence. But, among the many paleoseismic artifacts that have been documented, only soil liquefaction and landslides are presently capable of "recording" the intensity of ground motions. As summarized in Table 1, the date of an earthquake, and therefore recurrence-rate, can be derived from many types of evidence (although some are more likely to illuminate older records, owing to preservation potential or ease of discovery). Considering the spatial extent of such evidence (e.g., the length of coastline affected), an earthquake's location and magnitude may be estimated, at least crudely. However, most artifacts are only loosely correlated to the intensity of shaking, if at all. That is, the evidence may suggest that earthquake occurred, but do little or nothing to quantify the ground motions experienced. In this regard, soil liquefaction and landslides are more than just proxies of shaking, given that they could be used to quantitatively constrain its intensity across an affected region. This distinction arises because mechanistic models exist for predicting these phenomena as a function of ground-motion intensity measures (IMs).

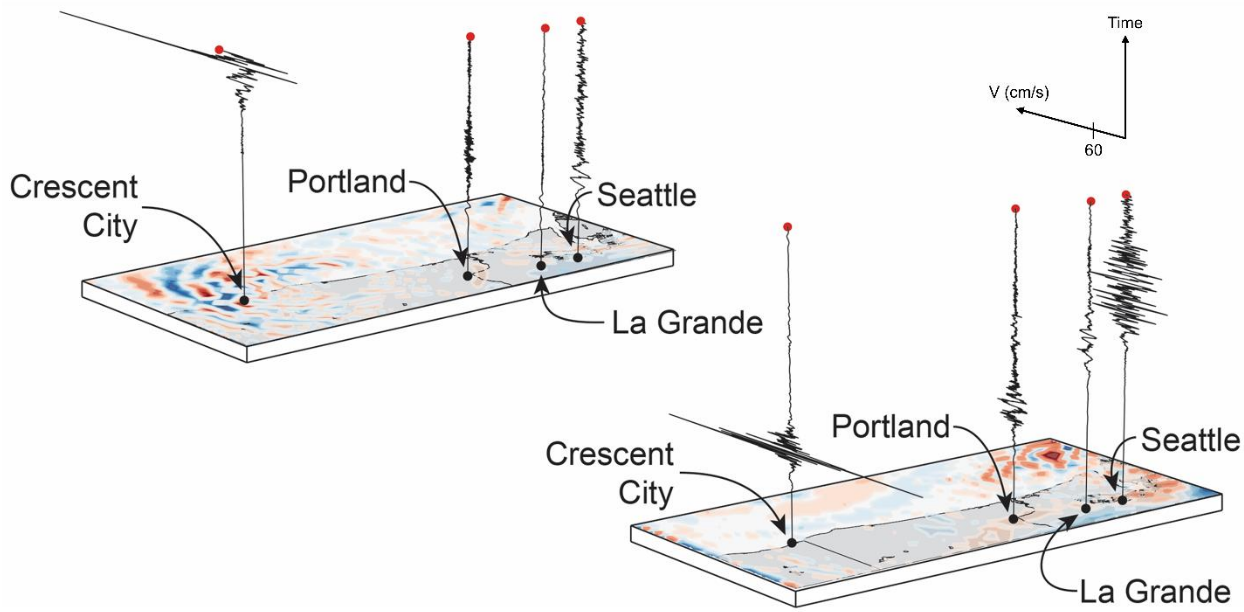


Fig. 1. Two ground-motion predictions of a M9 CSZ earthquake, reflecting the influence of salient modeling uncertainties (traces are ground velocity for the duration of the event; colors show a snapshot of seismic waves at a moment in time). The sensitivity of predictions to modeling uncertainties is readily apparent via the large discrepancies at some locales (e.g., in Seattle). Simulations by Frankel et al. (2018a) and Wirth et al. (2018).

Table 1. Synopsis of evidence from which paleoseismic parameters may be derived.

Paleoseismic Evidence	**Earthquake Parameters Obtainable from Evidence**		
	Rupture Date / Recurrence Rate	Rupture Location and Magnitude	Resultant Ground Motions
Dendrochronology	✓	✓	✕
Diatoms / Microfossils	✓	✓	✕
Other Subsidence Markers	✓	✓	✕
Tsunami Deposits / Impacts	✓	✓	✕
On- & Off-Shore Turbidite Records	✓	✓	✕
On-Fault Evidence	✓	✓	✕
Landslides / Rockfall	✓	✓	✓
Soil Liquefaction	✓	✓	✓

In conventional hazard analyses, wherein the seismic loading is given, these models are widely used to predict future outcomes (e.g., liquefaction). In paleoseismic studies, wherein the outcome is given (e.g., liquefaction did or did not occur), the models can be inverted to back-calculate the ground motions that likely would, and would not, produce the observation. It should be noted, however, that while paleoliquefaction features have been identified at over 180 locations in the PNW, including many that formed ca. 1700, paleolandslides with constrained ages are less common. As an example, landslides have yet to be linked to any CSZ earthquake (Struble et al. 2017). In addition, aseismic landslides occur frequently and are difficult to distinguish from their seismic counterparts, whereas aseismic ground failures resembling liquefaction are less common (e.g., Obermeier et al. 2011). Thus, liquefaction presently provides the best, and perhaps only, evidence from which the intensity of shaking in PNW paleoearthquakes may be quantified. This includes ruptures in the CSZ, but also on crustal faults in the Puget-Willamette Lowlands (e.g., the Seattle and Whidbey Island Faults) which have similarly not ruptured in modern history. The potential value of paleoliquefaction evidence has previously been proven in many other seismic zones (e.g., Tuttle and Hartleb, 2012; Bastin et al., 2016).

1.2 Motivation and Objectives

The long-term goal of this research is to use paleoliquefaction data to determine the strength of past ground motions in the PNW – both for CSZ events and others – and to further elucidate the locations, magnitudes, and recurrence rates of such ruptures. Towards that end, compilation of *regional* evidence is needed to answer questions of greatest interest (e.g., "When, on which fault, and of what magnitude was a paleoearthquake?" or "Which CSZ ground-motion simulations are plausible realizations of those experienced in 1700?"). This will be shown subsequently though a review of paleoliquefaction analytics. While analysis of an individual site could render ground motion IM values that likely would, or would not, produce an observation at that site, there will inevitably be an infinite number of these respective values. As an example, the minimum peak ground acceleration (*PGA*) requisite for liquefaction in a highly susceptible soil is ~ 0.1 g (de Magistris et al., 2013). If paleoliquefaction is observed at such a site, the *PGA* was therefore likely at least 0.1 g, but is otherwise unknown. A similarly loose upper-bound constraint is obtainable from a site where no paleliquefaction is observed, if the subsurface has very low susceptibility. Moreover, most *PGA* values could result from a small, nearby rupture, or from a large, distant rupture. The elucidation of regional ground-motion IM patterns thus requires spatially distributed study sites having a range of liquefaction susceptibilities, with and without observed manifestations of liquefaction. However, while many individual study sites have been documented in the PNW, these data exist across numerous publications and have not been compiled. Such an effort was undertaken nearly a decade ago in the Central-Eastern United States, where paleoliquefaction evidence was aggregated by Tuttle and Hartleb (2012), the results of which were used to inform seismic hazard analyses for nuclear facilities (NRC, 2012). The PNW has similarly ambiguous seismic records but lacks an analogous resource. As a result, regional scale paleoliquefaction studies have not been performed, and thus, the available field evidence has arguably not been exploited. Accordingly, the first objective of this paper is to compile existing paleoliquefaction evidence from the PNW into a community GIS database.

It is also critical that additional evidence be discovered, compiled, and analyzed, since better constraint of regional ground-motions in past earthquakes will result. However, a field search of the entire region would be extremely cost-prohibitive. Moreover, and using the CSZ as an example, there are infinite locales where the confirmed presence or absence of 1700 liquefaction would do little to inform or constrain ground-motion predictions (e.g., because various predictions lead to similar expectations of liquefaction). For field pursuits to be conducted more efficiently and strategically, it would be helpful to identify locations where uncertainties in ground-motion simulations (e.g., Frankel et al. 2018a; Wirth et al. 2018) give rise to significant differences in liquefaction predictions. Thus, the second objective of this paper is to develop maps – specific to the CSZ – that identify and prioritize where paleoliquefaction evidence should be searched for. This will be achieved, in part, using 30 different physics-based ground motions simulations of an M9 CSZ earthquake.

In the following, a summary of paleoliquefaction analytics is first presented. It will highlight: (i) the need to study evidence regionally; and (ii) prior applications of these analytics in the PNW, which while limited, hint at the potential for paleoliquefaction to provide new insights into persistent uncertainties. Next, the contents of the compiled paleoliquefaction database are described, and lastly, maps that guide future field expeditions for evidence in the CSZ are developed and discussed.

2. Summary of Paleoliquefaction Analysis Methods and Their Application in the PNW

While paleoliquefaction has been widely documented in the PNW, there are few studies in which seismic data has been derived from it (other than event age). Two of these studies are noted in the ensuing summary, to include emphasis of their shortcomings. It should be stressed, however, that these shortcomings are those of the tools then-available, which will serve to highlight recent advances. If not for these and other seminal studies, the database compiled herein would not be possible.

The study of paleoliquefaction has three phases: (i) field identification and interpretation; (ii) dating; and (iii) constraint of the earthquake magnitude and/or ground motion under which it formed. The reader is referred to the overviews of field interpretation by Obermeier et al. (2001; 2005), and to the field investigations of Obermeier and Dickenson (2000), Tuttle (2001), Talwani and Schaeffer (2001), Cox et al. (2007), and Tuttle et al. (2002a; 2002b; 2005), among others, for specific case studies. In addition, Sims and Garvin (1995), Quigley et al. (2013), Bastin et al. (2016), and Maurer et al. (2019) discuss field interpretation specific to spatiotemporally clustered earthquakes. Once identified, features may be dated via radiocarbon, optically stimulated luminescence, archeological or stratigraphic context, and soil development indicators, such as weathering and biologic activity. A comprehensive overview of paleoliquefaction dating methods is provided by Tuttle and Hartleb (2012). The techniques by which earthquake magnitude and/or shaking intensity are quantitatively constrained are generally called back- or inverse-analysis methods. The two most common to-date are the "magnitude-bound" method (e.g., Ambraseyes, 1988; Olson et al., 2005a; Maurer et al., 2015a) and the "site-specific geotechnical analysis," or for brevity, the "site-specific" method (e.g., Olson et al., 2005b; Rodriguez-Marek and Ciani, 2008; Green et al., 2005, 2014).

2.2 Magnitude-Bound Method

The magnitude-bound method uses a correlation relating earthquake magnitude to the site-to-source distance of the most distal site of liquefaction. Developed from observations in modern earthquakes, these correlations traditionally use data from variable geologic-tectonic settings and provide a lower-bound estimate of magnitude. As an example, Fig. 2 presents ten correlations from the literature.

One of these correlations was used by Bourgeois and Johnson (2001), who documented at least three episodes of paleoliquefaction in Washington's Snohomish River delta, with one episode dated ca. 910-990, coinciding with possible events on the Seattle Fault and CSZ. Bourgeois and Johnson (2001) deduced that: (1) a Seattle Fault rupture ≥ M_w7 would generate liquefaction 50 km away at the study site; and (2) other faults in the southern Puget Lowland, 75-120 km away, would require earthquakes ≥ M_w 6.5-7 to do so. Importantly, the maximum site-to-source distance of liquefaction is a function of numerous region- and site-specific factors. These include seismic source traits (e.g., focal depth and mechanism), transmission characteristics (e.g., ground motion attenuation and site effects), liquefaction susceptibility (e.g., density, fines-content, plasticity, and saturation); and subsurface stratigraphy (e.g., the quantity, depth, and thickness of all liquefiable strata, and the properties of overlying non-liquefiable strata), none which is directly accounted for by empirical magnitude-bound curves. Because these factors all vary (as reflected by the range of correlations in Fig. 2), region-specific correlations can provide more accurate estimates than those developed from global data (Olson et al., 2005a, 2005b; Maurer et al., 2015a). Shortcomings aside, the magnitude-bound method inherently requires compilation of regional evidence, since the liquefaction field resulting from an event must be fully delineated to properly ascribe a minimum rupture magnitude.

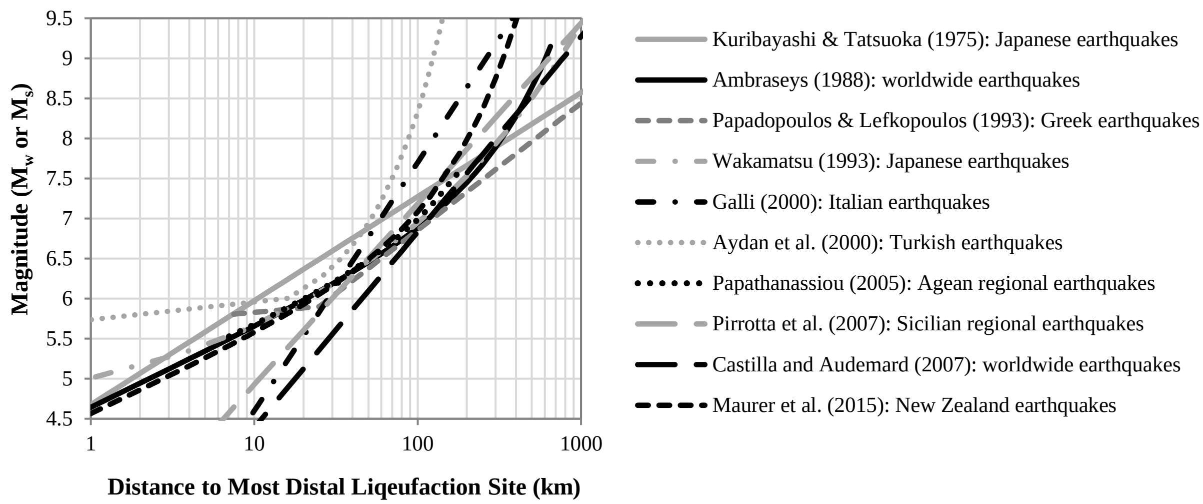


Fig. 2. Magnitude-bound curves for various geographic and tectonic settings, where site-to-source distance is quantified in terms of epicentral distance.

2.2 Site-Specific Geotechnical Analysis

The second, more technical site-specific method uses models based on in-situ geotechnical tests. In forward hazard analyses, wherein the seismic loading is given, these models predict the future triggering and manifestation of liquefaction (e.g., Green et al., 2019; Geyin and Maurer, 2020). In paleoliquefaction studies, wherein the outcome is given, the models are "inverted" to derive information about the causative earthquake. While implementations of the site-specific method have varied with time and place (e.g., Martin and Clough, 1990; Hayati and Andrus, 2008; Hu et al., 2002; Holzer et al., 2015; Gheibi et al., 2020; Rasanen and Maurer, 2021), variants can collectively: (i) identify combinations of rupture magnitude and ground-motion intensity likely to produce outcomes at individual sites (most applicable while the number of study sites is limited); (ii) probabilistically geolocate a seismic source from an evidence field and determine its magnitude (most applicable where faults are blind, or where prospective causative faults are plentiful); or (iii) compute the likelihood that a simulated ground-motion is a plausible realization of that experienced prior, as evidenced by the presence or absence of liquefaction (most applicable when competing forecasts are available).

With respect to the latter, and using the CSZ as an example, the likelihood that a ground-motion simulation represents a past event, given a set of field observations, may be computed as the product of the probabilities of those observations, conditioned on the simulation. Thus, the likelihood that a rupture had certain traits (e.g., location (L), geometry (G), and magnitude (M_w)…) given a set (x) of field observations at N different sites, can be computed as:

$$Likelihood\ (L, G, M_w \ldots | x) = P(X = x | L, G, M_w \ldots) = \prod_{i=1}^{N} P(X_i = x_i | L, G, M_w \ldots) \tag{1}$$

Where $P(X_i = x_i | L, G, M_w \ldots)$ is the probability of the observation at site i (liquefaction or no liquefaction) given an earthquake with parameters L, G, M_w, etc. By repeating for numerous simulations, the most plausible realizations of a past event can be probabilistically identified via the likelihood function (product of the probabilities of N observations), such that different combinations of L, G, M_w, etc. will be found more and less likely to produce the observed field evidence. Like the magnitude-bound approach, this method inherently relies on compilation of regionally distributed study sites. In Eq. (1), the probability of a field observation is computed as $P(\text{Liquefaction}|PGA, M_w)$ if manifestations of liquefaction are observed, and as 1 - $P(\text{Liquefaction}|PGA, M_w)$ otherwise. $P(\text{Liquefaction}|PGA, M_w)$ is the probability of observing liquefaction at a site, given ground-motion parameters PGA and M_w, as computed by a model of liquefaction triggering or, more ideally, liquefaction surface manifestation (embedded in which is a triggering model). As an example, Geyin and Maurer (2020) proposed fragility functions conditioned on three different liquefaction manifestation models (e.g., the liquefaction potential index, *LPI*, proposed by Iwasaki et al., 1978) computed using six different liquefaction triggering models (e.g., Green et al., 2019) such that users can select from, or average, 18 total functions. The results of such an analysis could in-turn inform CSZ modeling uncertainties such as: What were the extents of fault rupture? Where was the hypocenter? What was the direction of rupture propagation? Where were the rupture asperities?

The authors are aware of one prior application of the "site-specific" method in the PNW. Obermeier and Dickenson (2000) conducted field searches on six Columbia River islands, building upon and benefiting from many previous efforts (e.g., as compiled by Atwater, 1994). This resulted in the documented presence, or judged absence, of paleoliquefaction on each island, with features generally constrained to the year 1700 and decreasing in size and frequency moving inland. Obermeier and Dickenson (2000) used liquefaction models to back-calculate the *PGA*s that occurred on this transect in 1700. Notably, they suggest that *PGA*s were as low as 20% of those predicted by recent M9 simulations (Frankel et al., 2018a; Wirth et al., 2018). As stated by Obermeier and Dickenson (2000): "Our interpreted levels of shaking are considerably lower than current estimates that use theoretical and statistical models to predict ground motions of subduction earthquakes in the Cascadia region."

While these findings are provocative, they come with an important series of caveats. *First*, Obermeier and Dickenson (2000) utilized the liquefaction triggering model of Youd and Nobel (1997), to which standard penetration test (SPT) data from a site is input. However, the Youd and Nobel (1997) model is not used today - a sequence of major modifications and additions have since been made (e.g., to account for factors then unknown to be significant). *Second*, as new in-situ test methods have been developed, SPT-based liquefaction triggering models have fallen out of favor, with the cone penetration test (CPT) now recognized as the ideal (NRC, 2016). *Third*, because the study sites were located on islands, SPT equipment was not actually deployed. In its place, a hand-held variant was used (referred to today as a dynamic cone penetration test) and then correlated to SPT measurements by unknown means. The results of this test were then input to the Youd and Nobel (1997) model, which predicts liquefaction triggering at depth within a soil profile. *Fourth*, to predict whether liquefaction that triggered at depth should be expected to manifest at the surface, Obermeier and Dickenson (2000) utilized the Ishihara (1985) "H1-H2" chart (which predicts the thickness of a surficial crust needed to suppress surface manifestation). However, when tested on more recent earthquakes, this method has in some events exhibited prediction efficiencies similar to random guessing (van Ballegooy et al. 2015). Newer manifestation models informed by significantly larger datasets (e.g., van Ballegooy et al. 2014; Maurer et al. 2015b) are now available. *Fifth*, none of the methods adopted by Obermeier and Dickenson (2000) accounted for uncertainty (e.g., probabilistic liquefaction models were not available at the time). Notably, deterministic liquefaction models like Youd and Nobel (1997) traditionally have embedded conservatism, such that the binomial threshold for triggering corresponds to a relatively low probability of liquefaction (e.g., 15%). The *PGA*s constrained by Obermeier and Dickenson (2000) may thus correspond to the 15th percentile of what possibly occurred, rather than to a best estimate.

As stated by Obermeier and Dickenson (2000): "Our arguments are based largely on qualitative inferences of liquefaction susceptibility supported by preliminary geotechnical data." The true 1700 ground-motions thus remain enigmatic. Considering major advances in liquefaction analytics made over the last 20 years, CSZ paleoliquefaction features can and should be investigated using modern tools and methods. As concluded by Obermeier and Dickenson (2000): "More paleoliquefaction and geotechnical field studies are needed to bracket the strength of shaking."

3. Compilation of PNW Paleoliquefaction Evidence

Paleoliquefaction evidence in the PNW was compiled from 24 publications, as summarized in Table 2. The information compiled from each reference includes (when available): study site locations and feature morphologies (e.g., sand-blow thickness and dike width); dating information; important comments from the author; and citations to all original source documents, data, and figures.

Table 2. Literature reviewed and compiled in the paleoliquefaction database.

Reference	Study Region(s)	Reference ID
Atwater (1992)	Washington coast	1
Atwater (1994)	Columbia River	2
Atwater (2020)	Puget Sound (West Point), Washington	3
Bourgeois and Johnson (2001)	Puget Sound (Snohomish River delta), Washington	4
Briggs (1994)	Oregon coast	5
Clague et al. (1992)	Fraser River delta, British Columbia	6
Clague et al. (1997)	Fraser River delta (Annacis Island), British Columbia	7
Davis (2019)	Puget Sound (Duwamish River), Washington	8
Fiedorowicz (1997)	Oregon coast	9
Kelsey et al. (2002)	Oregon coast	10
Martin and Bourgeois (2012)	Hood Canal (Skokomish River delta and Lynch Cove), Washington; Lake Sammamish (Issaquah Creek), Washington	11
Obermeier (1995)	Chehalis River, Washington; Columbia River	12
Peterson and Madin (1997)	Columbia River; Washington coast; Oregon coast	13
Peterson et al. (2005)	Oregon coast	14
Peterson et al. (2008)	Oregon coast	15
Peterson et al. (2013)	Washington coast	16
Peterson et al. (2014)	Willamette River, Oregon; Oregon coast; Washington coast	17
Polenz et al. (2010)	Hood Canal (Skokomish Valley), Washington	18
Sherrod (2001)	Puget Sound, Washington	19
Sherrod et al. (2004)	Puget Sound, Washington	20
Sims (2002)	Calapooia River, Oregon	21
Takada and Atwater (2004)	Columbia River	22
Whistler et al. (2002)	Lake Sammamish (Issaquah Creek), Washington	23
Zehfuss (2005)	Puget Sound, Washington	24

Tuttle and Hartleb (2012), under the auspices of the Nuclear Regulatory Commission (NRC, 2012), previously developed an analogous GIS database for the Central and Eastern United States, compiling data for subsequent distribution to the research community. This seminal resource, which included multiple seismic zones but did not extend west of the Mississippi River, serves as an excellent guide for aggregating paleoliquefaction data elsewhere. To facilitate continuity between regions, the data fields proposed by Tuttle and Hartleb (2012) were adopted with minor additions for the PNW. These additions include fields to describe the thickness and lithology of the non-liquefiable capping layer, or "crust", as well as dendrochronological information, which field geologists have compiled at some sites in the PNW. A complete listing of the data-field names is given in Table 3, as are detailed descriptions of each attribute. Figure 3 illustrates several of these attributes, including: sand-blow thickness, width, and length; dike width; and sill thickness.

Table 3. Paleoliquefaction database field names and their descriptions.

Field Name	Description
SITE_NAME	Alphabetic designator of study area.
FEAT_ID	Unique alphabetic identifier for paleoliquefaction features within the same study area (e.g., Columbia River-01 and Columbia River-02). If a letter is present after the number, then there are multiple liquefaction events in the geologic record at the same location (e.g., Columbia River-06a and Columbia River-06b).
XCOORD	Numeric value of longitude, in decimal degrees.
YCOORD	Numeric value of latitude, in decimal degrees.
COORD_ORIG	Alphabetic description of reference from which study site coordinates are derived from. Some locations were digitized from maps, rather than obtained directly from coordinates. Site coordinates, as given in reports, may also have uncertainty due to limited measurement precision. For example, several Peterson and Madin (1997) study sites were in water (likely from GPS measurement error). In some cases, obvious errors were corrected by the authors (e.g., moving the coordinates to a riverbank near the original coordinates). In other cases, coordinates in water were left as-is.
OBS_TYPE	Alphabetic description of where/how paleoliquefaction was discovered (e.g., river cut bank, trench, borehole, geoslice).
FEAT_TYPE	Alphabetic description of feature type observed (e.g., sand blow, dike, sill, soft sediment deformation).
SSD_DESCR	Alphabetic description of seismic related soft sediment deformation (SSD) features (e.g., convolute beds, flame structure).
FEAT_REF	Alphabetic description of reference where paleoliquefaction feature information was obtained.
SB_THICK SB_WIDTH SB_LENGTH DK_WIDTH SILL_THICK	Numeric values of dimensions of sand blow thickness, sand blow width, sand blow length, dike width, and sill thickness, respectively. All dimensions are in cm and given as maximums. Of course, features may not be fully uncovered or delineated in the field, so maximum dimensions may be larger. Dimensions are either directly supplied by original authors or were inferred from to-scale figures. When available, sand blow thickness is measured adjacent to the vent. Sand blows are assumed circular, such that cross-sectional measurements of their size are assumed representative. The value "present" was assigned to sites where sand blows, dikes, or sills were documented but not measured.
CAP_THICK	Numeric value for the thickness of the capping layer, if provided.
CAP_LITH	Alphabetic description of type of soil the capping layer is mainly composed of, if provided.
DIM_REF	Alphabetic description of reference which provides paleoliquefaction feature dimensions either specifically in writing, from tables, or from to-scale figures.
MAX_CAL_2S	Alphanumeric description of two standard deviation maximum calibrated age range specified as cal AD or cal BC.
MIN_CAL_2S	Alphanumeric description of two standard deviation minimum calibrated age range specified as cal AD or cal BC.
MAX_CAL	Numeric maximum calibrated age in years AD (negative values indicate years BC). The single value given for maximum calibrated age is the two standard deviation limit. For example, if a site has a two standard deviation maximum calibrated age range of 1413-1642 AD, then the maximum calibrated age is 1413 AD.
MIN_CAL	Numeric minimum calibrated age in years AD (negative values indicate years BC). The single value given for minimum calibrated age is the two standard deviation limit. For example, if a site has a two standard deviation minimum calibrated age range of 1461-1878 AD, then the minimum calibrated age is 1878 AD.
C14_REF	Alphabetic description of reference which provides calibrated age or radiocarbon age. For sources which only provided radiocarbon ages, the Stuiver et al. (2020) CALIB program was used to convert radiocarbon ages to calibrated ages.
DENDRO_MAX	Alphanumeric maximum calibrated age in years AD from tree ring data.
DENDRO_MIN	Alphanumeric minimum calibrated age in years AD from tree ring data.
DENDRO_REF	Alphabetic description of reference which provides the dendrochronology data.
PREFAGEEST	Alphanumeric preferred age estimate in the format given by original authors. Some authors give an exact year, such as 1700 AD (Atwater, 1994), while others give qualitative descriptors, such as "slightly older than 130 BC" or "significantly younger than 1150 BC" (Obermeier, 1995). Preferred ages are based on radiocarbon, stratigraphy, archeology, or other dating methods. Additional details are provided in the COMMENT field where helpful.
PREFAGEREF	Alphabetic description of the source which provides the preferred age data.
STRAT	Alphabetic description of feature age based on stratigraphic relationships. For example, using the circa 1480 AD Mt St. Helen's ash layer to estimate a feature's age.
ARCHEO	Alphabetic description of feature age based on archaeological age data.
WEATHERING	Alphabetic description of degree of weathering of feature, or of surrounding sediments, to give an indication of the age of the feature or surrounding sediments.
COMMENT	Alphabetic description of salient comments/conclusions made by the original author(s), and other relevant information not captured in previous fields but extracted by the database compilers. There are three comment fields in total.

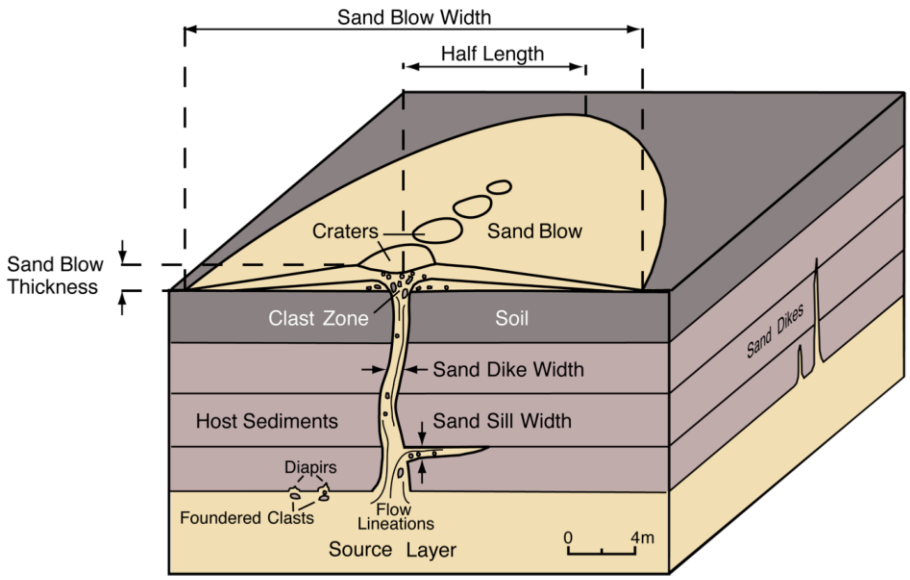


Fig. 3. Illustration of liquefaction-feature characteristics referenced in Table 3, including: sand-blow thickness, width, and length; dike width; and sill thickness. Figure from Tuttle and Hartleb (2012).

The resulting, curated dataset is available via the NHERI DesignSafe data depot at Rasanen et al. (2021) (https://doi.org/10.17603/ds2-fqkr-h615) and is provided as both a GIS map package and a flatfile spreadsheet. It should be noted, with respect to the attributes described in Table 3, that: (i) not all study sites contain information in all data fields (e.g., some lack dating information); (ii) many study sites have multiple liquefaction features, in which case features in a given locale are denoted by a single data point (e.g., Clague et al. (1992) documented 80 liquefaction features at a "site", but only provided a range of their sizes, in lieu of describing individual features); (iii) both radiocarbon dating methods of radiometric and accelerator mass dating are considered to give C14 dates; (iv) while all features in the database are judged to be of seismic origin based on conclusions made by the original investigators and reassessed by the current authors, the possibility of an aseismic cause nonetheless persists. In this regard, the criteria proposed by Obermeier (1996) for inferring seismic origin, further demonstrated by Obermeier et al. (2011), were used to provisionally rule out the possibility that features in the PNW were produced by aseismic geologic or climatic conditions.

In total, 185 study sites were compiled, as mapped in Fig. 4, and are respectively located in British Columbia (8), Oregon (43), and Washington (109). Additional sites reside along the banks, or on islands, of the Columbia River (25), which divides Oregon and Washington. Sites are otherwise concentrated in British Columbia's Fraser River delta, Oregon's Willamette River valley, Washington's Puget Sound, and in estuaries along the Pacific Coastline.

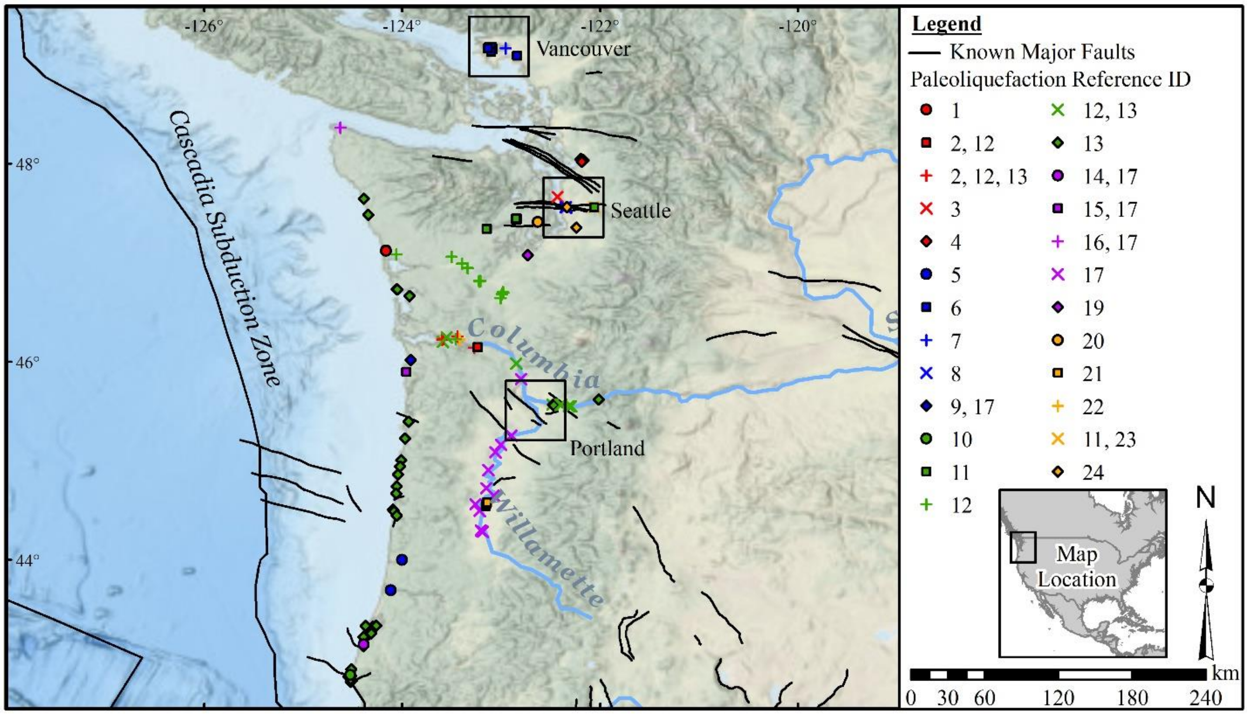
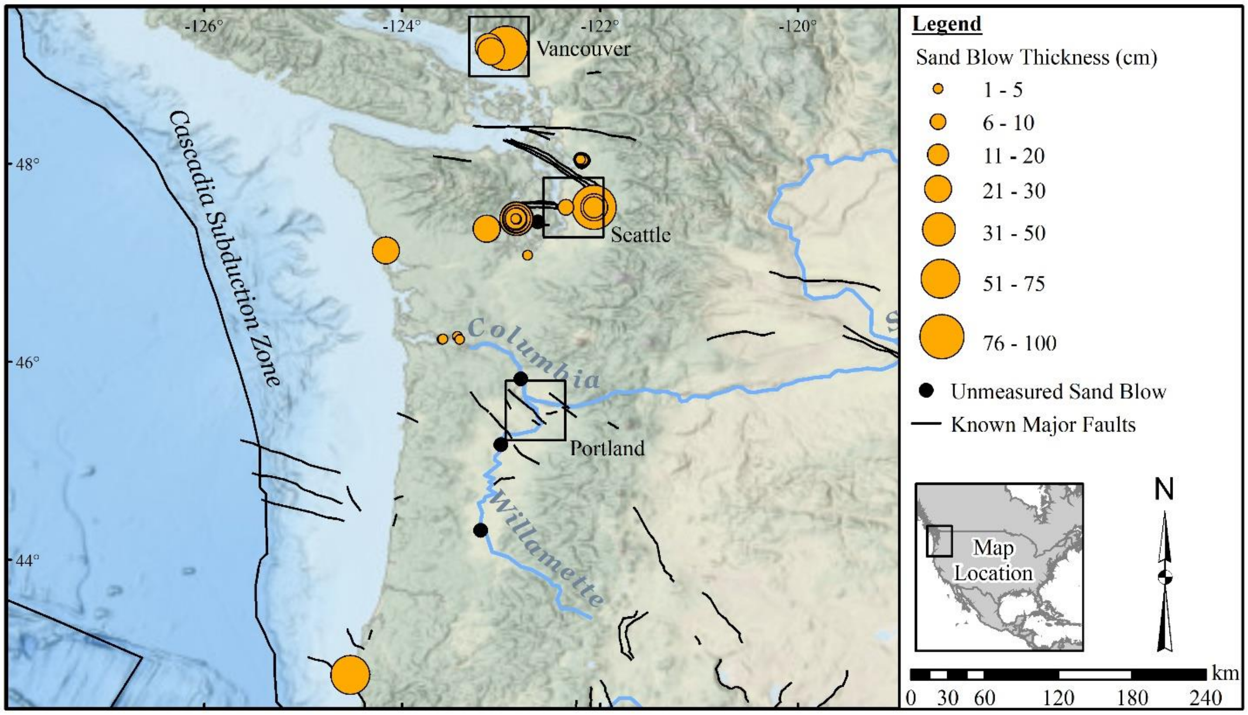


Fig. 4. Compilation of paleoliquefaction evidence. Reference IDs are given in Table 2.

The information compiled (see Table 3) allows for study sites to be symbolized by a variety of attributes (e.g., paleoliquefaction feature type, size, and age). Example products from the PNW database include the type of paleoliquefaction manifestation (e.g., sand blows or dikes) and the respective dimensions of the feature (e.g., sand blow thickness or dike width), as mapped in Fig. 5. In general, smaller, marginal features may more tightly constrain the intensity of past shaking, since this intensity was likely near the threshold for feature formation (i.e., for liquefaction triggering and manifestation). By corollary, larger features generally suggest prior loading far above such thresholds. However, because current liquefaction models predict the incidence of liquefaction manifestation more efficiently than the severity of manifestation (Maurer et al., 2015c), larger features may, at present, provide less quantitative constraint on the intensity of past shaking. The presence of large diameter features far from the coast (e.g., near the cities of Seattle and Vancouver) likely also indicates that not all features are associated with CSZ interface earthquakes, as will be further discussed. Additionally, as shown in Fig. 6, dating information – where available – can constrain the maximum calibrated age and/or minimum calibrated age (i.e., the dates which bound formation of a feature).



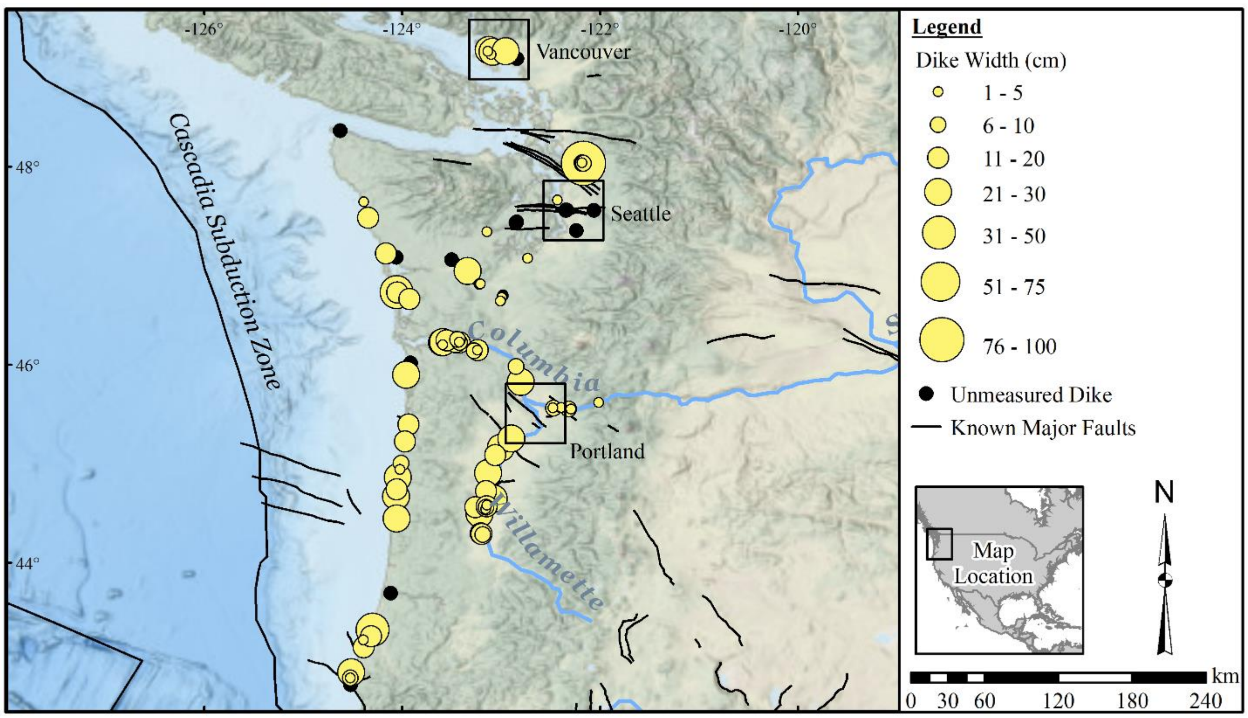


Fig. 5. Locations and sizes of paleoliquefaction features manifested as: (a) sand blows; and (b) dikes.

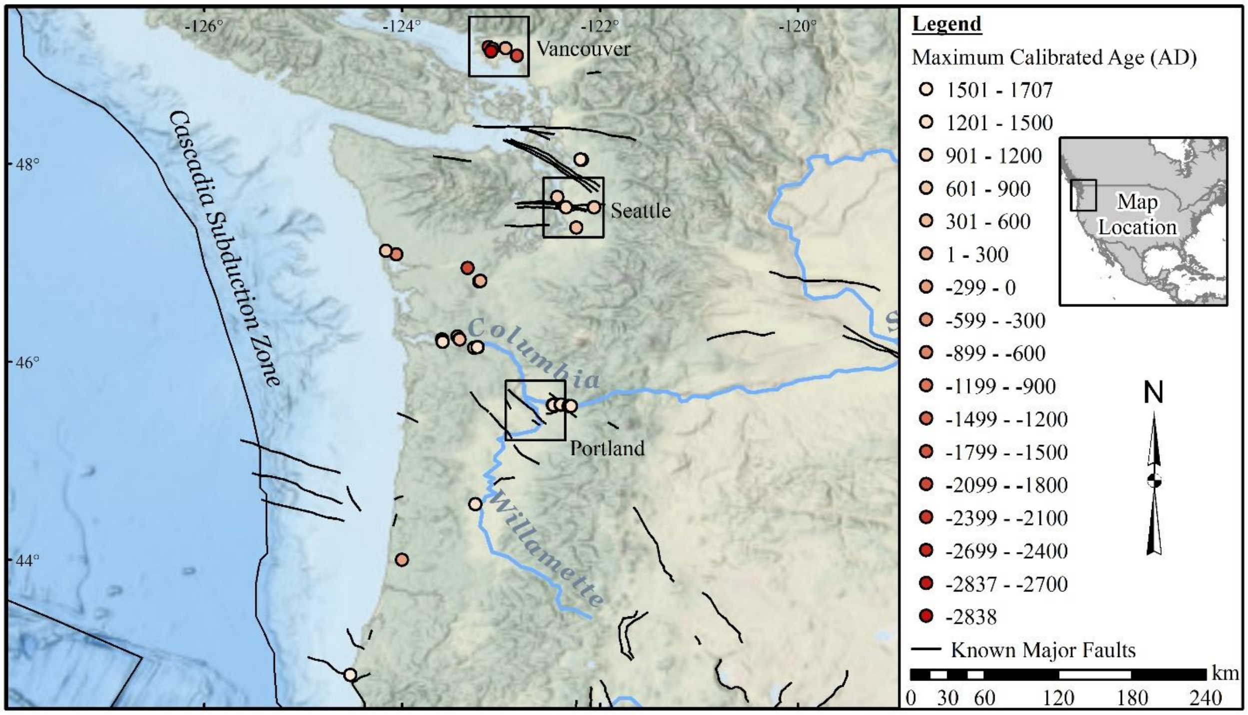
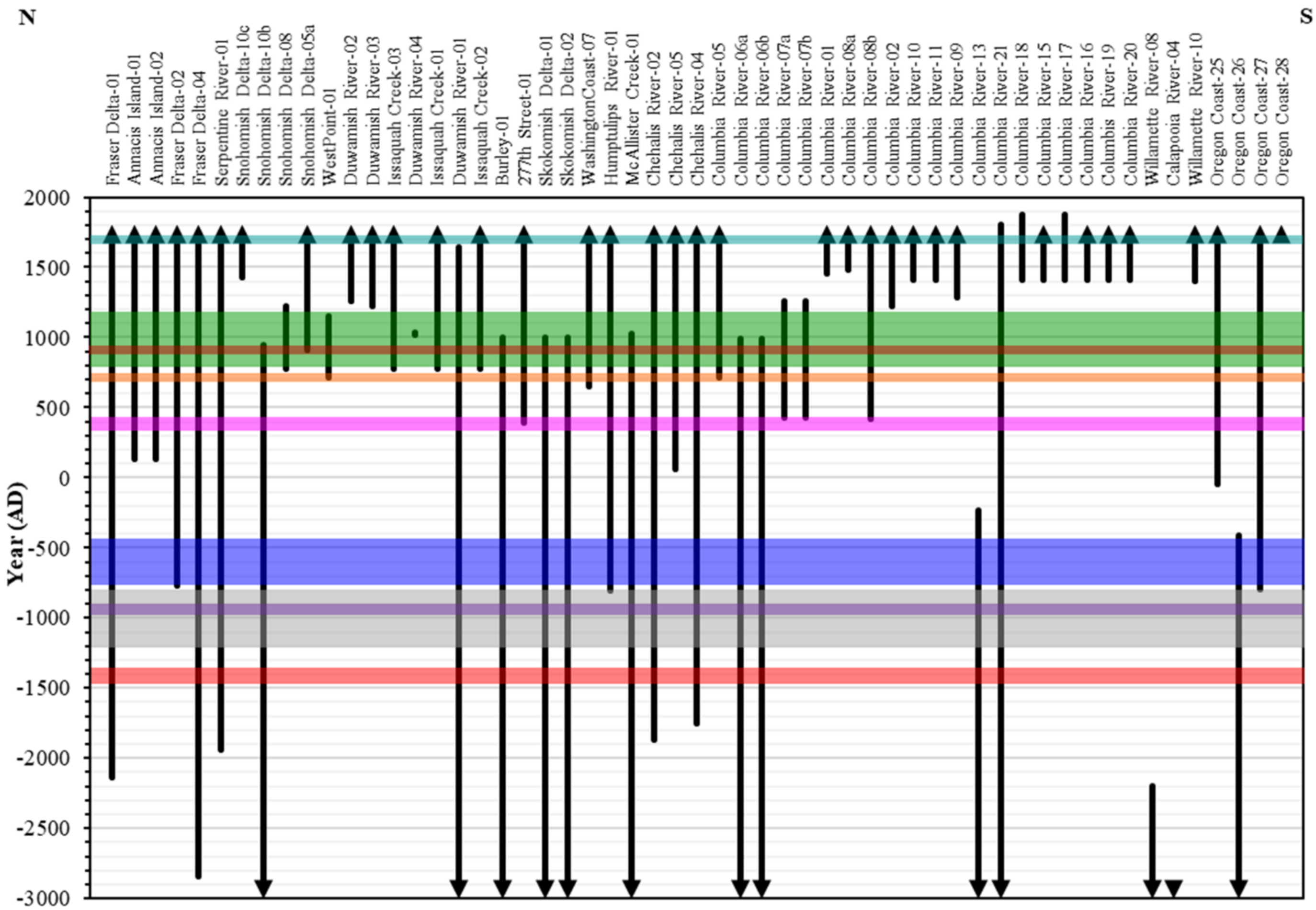


Fig. 6. Locations of paleoliquefaction features with dating information to constrain the maximum calibrated age (i.e., feature is younger than this date). Analogous figures can be developed to show the minimum calibrated age (i.e., feature is older than this date).

Of the 185 study sites compiled, 55 have liquefaction features with quantitatively constrained ages. Fig. 7 graphically depicts these constraints, juxtaposed against select historical earthquakes. Vertical lines denote the period in which a paleoliquefaction feature formed, with arrows denoting that either a minimum or maximum age is unknown. Superimposed are nine significant ruptures in the PNW, with emphasis on those which potentially could, or likely did, produce strong shaking where evidence was compiled. The constraints on these ruptures were collectively obtained from the paleoseismic studies of Atwater (1994, 1999), Satake et al. (1996), Atwater and Hemphill-Haley (1997), Atwater et al. (2003), and Kelsey et al. (2004). Other events proposed in the literature, such as the A.D. 770-1160 Tacoma fault rupture (Sherrod et al., 2004) and earlier ruptures in the CSZ, are not included. To this we also add the possibility of undocumented ruptures on faults known and unknown. While the age ranges of many features are relatively large (e.g., spanning multiple known earthquakes), several authors provide more narrow, preferred ages by holistically weighing the evidence at a site. In this regard, Fig. 7 does not attempt to convey the nuanced judgement of original investigators, but rather, serves as a preliminary guide for further reading. As an example, Davis (2019) investigated four neighboring sites on the Duwamish River in Seattle. While dating does not strictly bound the minimum ages (i.e., latest dates of formation) for many of the features investigated, Davis (2019) concluded:

“none of the dikes observed are likely to be as young as the 1700 Cascadia earthquake.” Other authors similarly reject, or adopt, the plausibility of certain events to have caused a given feature. Some identify a particular causative event, such as the 1700 CSZ rupture (e.g., Atwater 1994; Satake et al., 1996) while others qualify a quantitative constraint, such as “slightly older than 130 B.C.” or “significantly younger than 1150 B.C.” (Obermeier, 1995). Preferred age estimates, and the rationale supporting them, are provided in the data attributes (see Table 3).

Ultimately, the compiled database may be used to target study sites for: (i) dating (e.g., where features are undated, or otherwise loosely constrained); and (ii) in-situ geotechnical tests, which are needed to perform inverse-analysis of the shaking experienced (See *Section 2.2*). Towards this end, the database notably lacks observations classified as negative (i.e., well documented sites judged capable of preserving evidence, but where none is observed), from which upper bounds on the intensity of shaking may be computed. While the aphorism “absence of evidence is not evidence of absence” holds true for all paleoseismic research, the need nonetheless exists, where prudent, to explicitly judge sites as “negative.” While further field reconnaissance is undoubtedly needed in the PNW, to include focus on negative observations, such cases could at present be developed from the compiled database. That is, at sites observed to be susceptible to liquefaction and inferred to be capable of preserving evidence at a given time. As an example, liquefaction has yet to be tied to the 1700 CSZ rupture anywhere in the Puget Sound (i.e., the region between Seattle and Vancouver), despite there being numerous sites with recurrent liquefaction episodes. In total, 16 of the compiled study sites with age constraint – many located in the Puget Sound – are likely not to have liquefied in the 1700 CSZ rupture (or at least, no evidence has been found). In such cases, inverse-analyses would provide upper bounds on the intensity of shaking. Ultimately, regardless of whether sites are “positive” or “negative,” a majority lack dating and geotechnical tests, which are both costly to perform, but which are needed to fully exploit the evidence. Accordingly, study sites should be selected for subsequent research in a strategic manner, guided by the compiled database (e.g., such that meaningful constraint on shaking intensity might be obtained).



Event	Age Range	References
Cascadia Subduction Zone	AD 1700	Atwater (1994); Satake et al. (1996)
Cascadia Subduction Zone	AD 790 – 1180	Atwater and Hemphill-Haley (1997); Atwater et al. (2003)
Seattle Fault	AD 900 - 930	Atwater (1999)
Cascadia Subduction Zone	AD 700 - 730	Atwater and Hemphill-Haley (1997); Atwater et al. (2003)
Cascadia Subduction Zone	AD 350 - 420	Atwater and Hemphill-Haley (1997); Atwater et al. (2003)
Cascadia Subduction Zone	BC 760 - 450	Atwater and Hemphill-Haley (1997); Atwater et al. (2003)
Cascadia Subduction Zone	BC 980 - 900	Atwater and Hemphill-Haley (1997); Atwater et al. (2003)
Southern Whidbey Island Fault	BC 1200 - 800	Kelsey et al. (2004)
Cascadia Subduction Zone	BC 1450 - 1360	Atwater and Hemphill-Haley (1997); Atwater et al. (2003)

Fig. 7. Dating constraints on the formation of paleoliquefaction features at 55 study sites, juxtaposed against select, major earthquakes in the PNW. Study sites are identified on the horizontal axis and ordered from north to south; vertical lines denote the period in which a feature formed.

4. Forecasting CSZ Paleoliquefaction

In addition to further study of existing (i.e., published) paleoliquefaction sites, it is important that new sites be developed, to include geotechnical testing and analysis. As more sites are identified and analyzed, better constraint of regional ground-motions in past earthquakes will result. Yet, a field search of the entire PNW would be cost-prohibitive, and moreover, not all study sites will be of equal value. There are many locations where the presence or absence of liquefaction may do relatively little to constrain ground motions in a past event. For example, where: (i) the soil is unsusceptible to liquefaction, as is the case for most land; (ii) the soil is so susceptible to liquefaction that most ground-motion simulations predict liquefaction, even where motions differ; or (iii) most ground-motion simulations predict similar motions, and by corollary, lead to similar expectations of liquefaction. For field pursuits to be conducted more efficiently and strategically, it would be helpful to identify locations where uncertainties in ground-motion simulations (e.g., Frankel et al. 2018a; Wirth et al. 2018) give rise to differences in the forecasted liquefaction response. Accordingly, maps – specific to the CSZ – are next developed to prioritize where paleoliquefaction evidence should be searched for.

4.1 Geospatial Liquefaction Models

While liquefaction is most effectively forecasted using models based on in-situ geotechnical tests (e.g., Geyin and Maurer, 2020), such an approach is impractical at regional scale. As a result, regional maps of liquefaction hazard have traditionally been developed using geology maps, from which areal classifications of susceptibility are assumed. While this approach has been used in Cascadia (e.g., Palmer et al., 2007), it does not consider seismic loading, and thus, does not explicitly predict liquefaction. More recently, "geospatial" liquefaction models have been proposed for regional applications. In lieu of directly measuring subsurface traits, geospatial models attempt to infer below-ground conditions from above-ground parameters (e.g., metrics of surface slope, mineralogy, roughness, wetness, and reflectance; distance to and elevation above rivers, streams, and other water bodies; and various mapped or remotely sensed values describing geology, geomorphology, bedrock and water depth, hydrology, climate, etc.). In effect, these parameters serve as proxies of subsurface traits. The efficiency of one such model (Rashidan and Baise, 2020) has been shown, in some settings, to rival that of more sophisticated models based on subsurface test data (Geyin et al., 2020). In the 2001 Nisqually, Washington, earthquake, for example, it achieved 93% prediction efficiency when tested on all available field observations (Geyin et al., 2021). Accordingly, Rashidan and Baise (2020) will be adopted to compute the probability of liquefaction manifesting at the surface, defined as:

$$PoL(X) = \begin{cases} (1+e^{-X})^{-1} & if\ PGV > 3\frac{cm}{s}\ and\ \ PGA > 0.1\ g\ \ and\ \ V_{S30} < 620\ m/s \\ 0 & otherwise \end{cases} \quad (2)$$

Where: *PoL* is the probability of liquefaction manifesting within a 100 m by 100 m grid; *PGV* is peak ground velocity; *PGA* is peak ground acceleration; V_{S30} is the shear wave velocity time-averaged over the upper 30 m; and *X* is a series of model parameters and coefficients defined as:

$$X = 8.801 + 0.334\ln(PGV) - 1.918\ln(V_{s30}) + 0.0005408\,(precip)\; - 0.2054(d_w) - 0.0333(wtd) \quad (3)$$

In Eq. (3): *PGV* and V_{S30} are as previously defined; *precip* is the mean annual precipitation (mm) capped at 1700 mm; d_w is the closest distance (km) to a river or coastline; and *wtd* is the water table depth (m). Each of these inputs is readily available and was implemented exactly as described in Zhu et al. (2017) and Rashidan and Baise (2020). Specifically: V_{S30} was obtained from the Heath et al. (2020) global V_{S30} map; *precip* was obtained from the WorldClim database; *wtd* was obtained from the Fan et al. (2013) global water-depth map; and d_w was calculated using river and coastline data from the USGS HydroSHEDS database and the NOAA coastline dataset, respectively. These data sources are elaborated upon in the Data Availability section. In addition to *PoL*, Rashidan and Baise (2020) propose that the percent area of ground covered by liquefaction manifestation may be predicted as:

$$LSE(PoL) = 49.15\,(1 + 42.4e^{-9.165PoL})^{-2} \quad (4)$$

Where *LSE* is liquefaction spatial extent (%) and *PoL* was previously defined and calculated in Eqs. (2-3). Maps of both *PoL* and *LSE* will be generated for 30 different simulations of a CSZ earthquake.

4.2 CSZ Ground-Motion Simulations

Physics-based ground motion simulations, which explicitly model kinematic fault rupture, wave propagation, and the subsurface velocity structure, are especially useful for studying seismic sources that lack historic records, such as the CSZ. Physics-based simulations also help to elucidate and quantify complex phenomena that may go undetected by empirical observations at discrete stations (e.g., the effects of directivity, basins, and topography). Accordingly, Frankel et al. (2018a) and Wirth et al. (2018) developed 30 physics-based realizations of a full-fault M9 CSZ rupture. This suite of predictions includes variations in the hypocenter location, down-dip limit of fault rupture, slip distribution, and locations of asperities. The down-dip rupture extent was varied to be consistent with the logic tree adopted for the CSZ by Petersen et al. (2014) in the US National Seismic Hazard Map. For each scenario, a total of 500,000 motions were generated on a 1 km^2 spacing for a region ranging from Northern California to British Columbia, and from off the Pacific Coast to Central Washington and Oregon. The resulting motions can be retrieved from Frankel et al. (2018b) and have been used to forecast a variety of impacts on the built and living environment (e.g., Marafi et al., 2019, 2020).

4.3 Site-Response Analysis

Inherent to the Frankel et al. (2018a) and Wirth et al. (2018) suite of simulations is an assumed surficial shear-wave velocity (V_S) of 600 m/s. That is, the predictions are for rock conditions and do not consider the potential for local soil conditions to alter the amplitude and duration of incoming motions. Therefore, to predict liquefaction using the Rashidan and Baise (2020) model, the CSZ ground-motion simulations were first modified for surficial conditions using site-amplification factors derived from wave-propagation site-response analyses (i.e., to propagate the motions from rock through softer surficial materials, where present). Towards that end, profile measurements cannot feasibly be made everywhere, given the regional scale of the analyses. Accordingly, site amplification factors were computed, in part, using the Marafi et al. (2021) soil velocity model (SVM), which predicts profiles of *Vs* versus depth (*z*), and which is specific to the PNW. Using this approach, V_S (*z*) is defined as:

$$V_s(z) = \begin{cases} V_{S0}, & z < 2.5\,m \\ V_{S0} + 1000 \cdot \left(k\,\frac{z-2.5}{Z_{1.0}-2.5}\right)^{\frac{1}{n}}, & z \geq 2.5\,m \end{cases} \tag{5}$$

where: V_{S0} defines V_S at the surface; k controls the near-surface rate-of-change in V_s; n controls the rate-of-change in V_s at greater depths; $Z_{1.0}$ is the depth, in meters, where V_S = 1 km/s; V_s and V_{S0} have units of m/s; and *z* is depth in meters. To anchor the predicted V_s at $Z_{1.0}$, the parameter k is defined as:

$$k = \left(\frac{1000-V_{S0}}{1000}\right)^n \tag{6}$$

Thus, to apply Eqs. 5 and 6 at any given location requires input variable $Z_{1.0}$ and model parameters V_{S0} and *n*. Drawing from 909 V_S profiles measured in the Cascadia region, Marafi et al. (2021) modeled and trained expressions for V_{S0} and *n*:

$$V_{S0} = a_0 + a_1(V_{S30})^{a_2} \tag{7}$$

$$n = b_0(V_{S30})^{b_1}(Z_{1.0})^{b_2}(V_{S30}Z_{1.0})^{b_3} \tag{8}$$

where: a_0 = -629; a_1 = 434; a_2 = 0.122; b_0 = 0.00912; b_1 = 0.646; b_2= -0.201; b_3 = 0.136; and where other variables are as previously defined. While the training profiles were distributed throughout Washington, Oregon, and British Columbia, a majority were in the general vicinities of either Portland, Seattle, or Vancouver, consistent with both the extent and concentration of compiled paleoliquefaction study sites. For the present study, predictions of input variables V_S30 and $Z_{1.0}$ were obtained from the models of Heath et al. (2020) and Stephenson et al. (2017), respectively, as suggested by Marafi et al. (2021). The Stephenson et al. (2017) community velocity model provides mapping of deep geologic structure, but not of the near surface, and has a minimum V_S of 600 m/s. It thus provides the best means of estimating $Z_{1.0}$ but is otherwise unsuitable for predicting ground motions at the surface.

While V_S profile measurements are not yet available at paleoliquefaction sites, the adopted SVM was shown by Marafi et al. (2021) to provide significantly better predictions of profiles in the PNW across all site conditions, as compared to other regional or general SVMs. This may be attributable to Cascadia's numerous geologic basins and glaciated landscapes, which give rise to a wider range of V_{S30} and $Z_{1.0}$ combinations than is typically found in other data-rich seismic zones (e.g., California). Accordingly, the Marafi et al. (2021) SVM was used to estimate V_S profiles at 1 km^2 across a domain consistent with that of the CSZ ground-motion simulations.

Each of the 500,000 simulation motions, for each of 30 realizations, was modified via equivalent linear site-response analysis using pysra (Kottke, 2020), a Python implementation of the software Strata (Kottke and Rathje, 2008). Motions were input to the predicted profiles at a V_S of 600 m/s, consistent with the velocity at which they were computed. Nonlinear soil behavior was modeled per Darendeli (2001) with the following inputs: plasticity index = 30; unit weight = 19.6 kN/m^3; ground water depth = 5 m; coefficient of at-rest earth pressure = 1.0; and overconsolidation ratio computed per Wair et al. (2012). While systematic parameter variation was not undertaken, the most significant findings were relatively insensitive to the selected inputs, relative to other uncertainties (e.g., that of V_S versus depth). Ultimately, regional-scale analyses (i.e., analyses that do not use continuous subsurface measurements) have inevitable limitations and uncertainties. The Marafi et al. (2021) SVM is not a probabilistic model, nor are probabilistic estimates of its input parameters, V_{S30} and $Z_{1.0}$, available. The lack of uncertainty quantification within the methods for site-response and liquefaction should not be interpreted to mean that these and other uncertainties do not exist. As improved models of the subsurface velocity structure are developed, or as site-specific measurements are made, the accuracies of the site-response analyses performed herein will improve, and by corollary, so too will forecasts of consequent liquefaction.

Following the approach described above, ground surface time-histories were obtained with the same resolution and extents as the Frankel et al. (2018) and Wirth et al. (2018) simulations. For each of the 500,000 motions per realization, *PGV* and *PGA* were computed considering the geometric mean horizontal component of motion. These IMs were then linearly interpolated at a resolution of 100 m^2. Mapped in Fig. 8, considering all 30 realizations, is the 5-to-95 percentile variation in expected *PGA* and *PGV*. For many locations, these expected intensities vary significantly across the suite of realizations (e.g., differences in *PGA* exceeding 1.5 g) despite all modeling a M9 rupture. Some of the largest variations are seen: (i) on the Pacific Coast near the Juan de Fuca Strait in the north, and near Port Orford, OR, in the vicinity of the CSZ's terminus in the south (driven by differing assumed hypocenter locations, among other modeling variables); and (ii) where deep sediment basins have a propensity to amplify differences in incoming motions (e.g., in the vicinities of Portland, Seattle, and Vancouver). It can also be seen in Fig. 8 that some paleoliquefaction study sites are in areas of large ground-motion uncertainty. This hints at the possibility for the presence or absence of paleoliquefaction to constrain influential modeling variables, at least insofar as what occurred in A.D. 1700. In addition to the maps shown in Fig. 8, the median expected *PGA* and *PGV*, considering all 30 realizations, are mapped at full resolution in the Rasanen et al. (2021) GIS package, where the PNW paleoliquefaction database is also found. These predicted ground-motion intensities could be used to forecast a variety of regional-scale seismic impacts, in addition to liquefaction.

4.4 Results and Discussion

Liquefaction parameters *PoL* and *LSE* were forecasted across the PNW for each of the 30 modified ground-motion predictions resulting from *Section 4.3*. Given the uncertainties inherent to regional analyses, these products (i.e., liquefaction forecasts) should be viewed as preliminary tools: (i) to guide future field reconnaissance efforts and site-specific geotechnical studies, as will be further discussed; and (ii) to inform regional planning, policy, and science communication. As an example, *LSE* is mapped in Fig. 9 for one CSZ realization ("csz013"). It can be seen, for this realization, that some of the largest predicted *LSE* values are in the Fraser river delta of British Columbia. Notably, confirmed 1700 A.D. liquefaction has yet to be discovered anywhere in the vicinity, despite there being other, apparently older features at several study sites in the area. Thus, while site-specific geotechnical testing and analysis would be needed to confirm the geospatial model's prediction, the results in Fig. 9 might suggest that "csz013" is an unlikely realization of the motions experienced in 1700 A.D. Ultimately, geotechnical analyses at several regionally distributed study sites would be needed to conclude this.

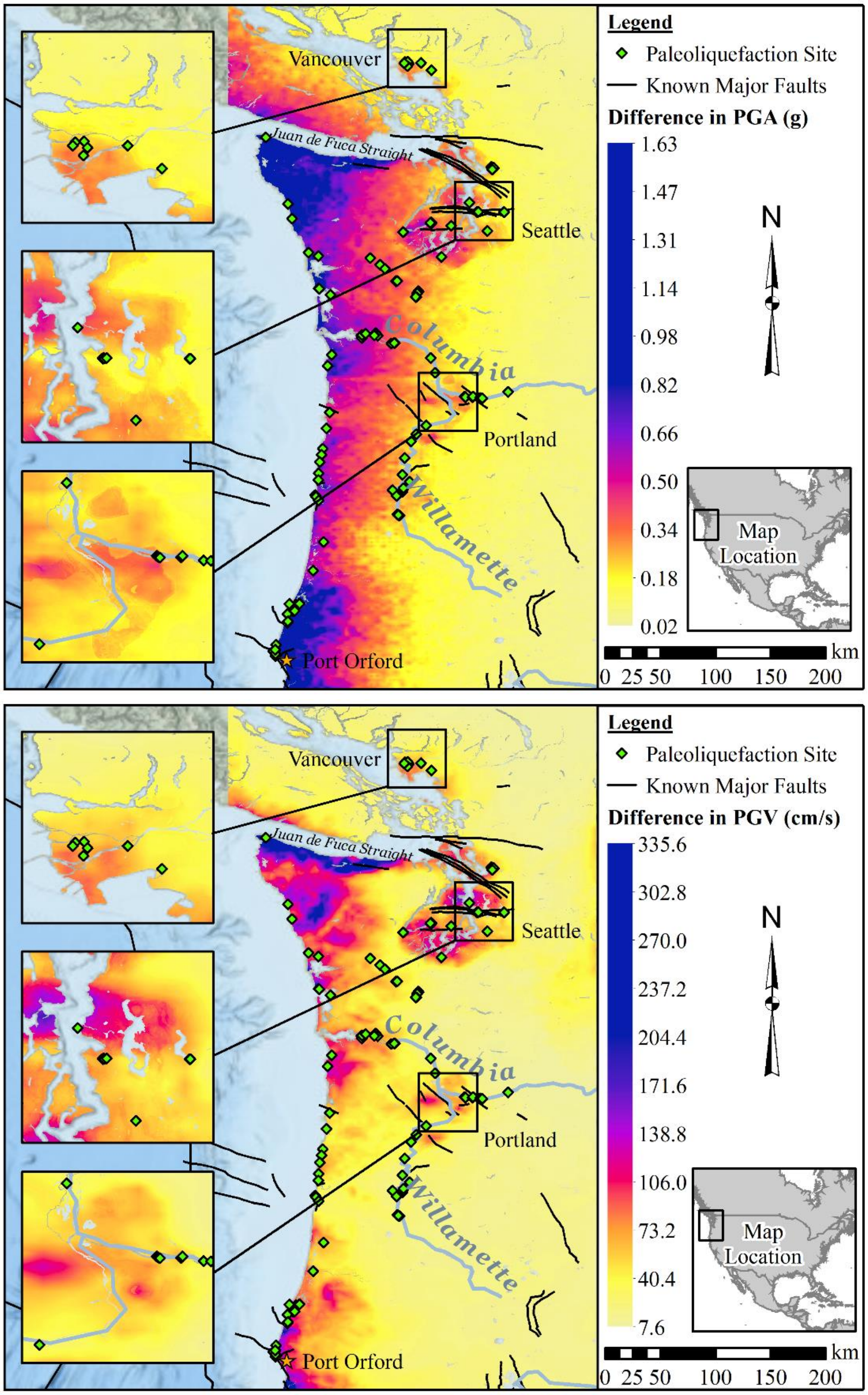


Fig. 8. Maps showing the 5-to-95 percentile variation in expected (a) *PGA* and (b) *PGV*, as computed from 30 different ground-motion simulations of a M9 CSZ earthquake. Expected shaking intensities vary significantly (e.g., by more than 1 g) depending on which simulation is adopted.

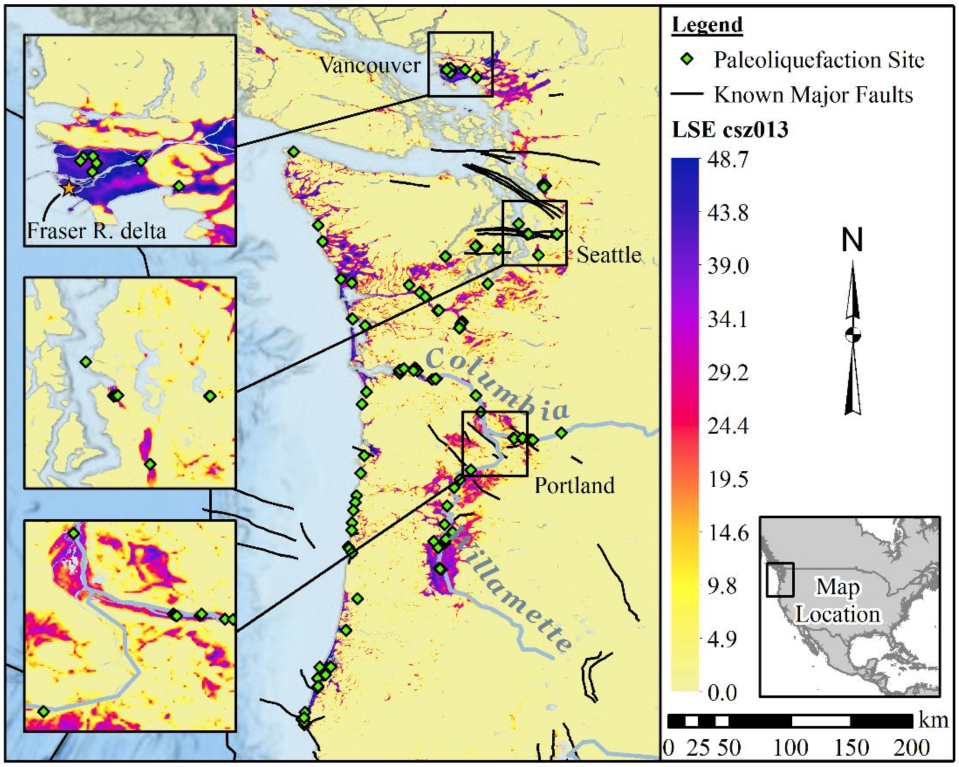


Fig. 9. Forecasted liquefaction spatial extent, *LSE* (%) computed for CSZ M9 ground-motion simulation "csz013" developed by Frankel et al. (2018a) and Wirth et al. (2018).

Analogous *PoL* and *LSE* forecasts are provided for each of the 30 CSZ M9 realizations, at full resolution, in the Rasanen et al. (2021) GIS package downloadable from https://doi.org/10.17603/ds2-fqkr-h615. Collectively, these forecasts may be used to identify locations: (i) where paleoliquefaction evidence is likely to be found; and (ii) where the expected liquefaction response differs across the suite of CSZ ground-motion realizations. As an example, the 5-to-95 percentile variation in *LSE*, considering all 30 realizations, is mapped in Fig. 10. Locations with large *LSE* differences are those where: (i) there are large differences in ground-motion intensity across realizations; and (ii) the subsurface is inferred to be susceptible to liquefaction. In this respect, differences in *LSE* increase as the inferred susceptibility to liquefaction increases, all else being equal. Fig. 10, which is included in the above GIS package, can thus be used to identify where the presence or absence of liquefaction would have greater potential to constrain the intensity of past shaking (i.e., where differences in *LSE* are greatest). In these locales, the presence or absence of features would be strongly at odds with some subset of the ground motion realizations, thereby diminishing their plausibility. Elsewhere, where differences in *LSE* are small, the presence or absence of features would potentially do little or nothing to constrain ground-motion models.

Regionally, some of the largest differences in *LSE* are (from south to north): near Klamath Lake, OR; along the Skagit River in the vicinity of Mt. Vernon, WA; and along the Fraser River in the vicinity of Chilliwack, BC. In each of these locales, predicted ground motions are sufficient for liquefaction in some realizations, assuming the subsurface is highly susceptible, but mechanistically insufficient in others, regardless of subsurface conditions. While paleoliquefaction has not been discovered in any of these areas, it is unknown whether field searches have been undertaken. Specific to the three urban centers highlighted in Fig. 10, relatively large *LSE* differences are predicted along the Columbia and Willamette Rivers in Portland, along the Duwamish River in Seattle, and in the Fraser River delta near Vancouver. In these urban areas, liquefaction manifestations are likely for some realizations of a CSZ rupture, but unlikely for others. Paleoliquefaction features have been documented in each of these urban areas, yet few, if any, of these were conceivably caused by the 1700 CSZ event (see Figure 7 and associated discussion). If these sites are judged not to have liquefied in 1700, the maximum shaking intensity could be provisionally constrained via geotechnical testing (i.e., if the intensity were any larger, liquefaction would be expected). Such analyses would also have the potential to inform regional ground-motion models, given that some subset of realizations would be less likely to represent that which occurred in 1700. Given the importance of geotechnical testing (e.g., CPTs) toward constraint of shaking intensity, an additional strategy is to search for the presence or absence of paleoliquefaction where testing has already been performed. That is, at locations where the subsurface is well documented, and where a significant cost of inverse analysis is already paid for. Towards this end, provided in the Rasanen et al. (2021) GIS package and mapped in Fig. 10, are the locations of many hundreds of CPTs presently available from the Washington Dept. of Natural Resources. Many of these CPTs, the data for which can be obtained from Jeschke et al. (2019), are in the vicinity of public lands and waterways where paleoliquefaction can be searched for.

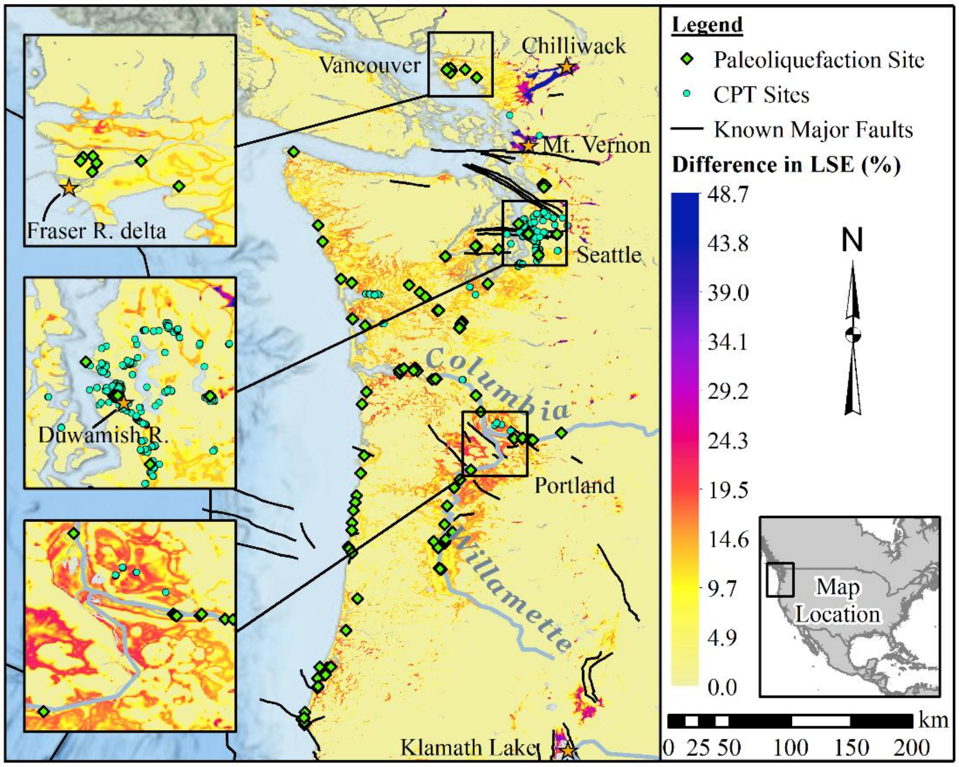


Fig. 10. The 5-to-95 percentile variation in expected *LSE*, as computed from 30 different ground-motion simulations of a M9 CSZ earthquake; paleoliquefaction and CPT sites are also mapped.

5. Conclusions

Paleoliquefaction provides the best, if not only, evidence from which the intensities of previous ground motions in the PNW may be constrained. Towards that ultimate goal, this paper first reviewed current paleoliquefaction inverse-analysis methods and their limited, prior applications in the PNW. Next, all existing PNW paleoliquefaction evidence was compiled from the literature into a GIS database for distribution to the engineering geology research community – the first such database compiled for the U.S. Pacific Northwest. This resulted in detailed data from 185 study sites (e.g., feature locations, types, sizes, and ages). Lastly, paleoliquefaction evidence was forecasted for 30 different physics-based ground motion simulations of a M9 CSZ earthquake. Collectively, these maps can be used to guide field explorations by engineering geologists for new evidence, considering both: (i) where evidence is likely to be found; and (ii) whether evidence is likely to provide meaning constraint of CSZ ground motion models. Of additional utility, this process resulted in the first ever suite of M9 CSZ ground-motion predictions for the Cascadia Region. In making these predictions, wave-propagation based site-response analyses were used to account for expected near-surface site effects. Together, the maps of expected shaking intensity and liquefaction may be useful in regional loss estimation, scenario planning, and science communication. Ultimately, as more paleoliquefaction evidence is identified and analyzed, better constraint of regional ground-motions in past earthquakes will result. This will undoubtedly require unified collaboration between geoscientists, geoengineers, and other research professionals. The products presented herein form a foundation for these efforts.

6. Data Availability

Available in digital format from NEHRI DesignSafe (https://doi.org/10.17603/ds2-fqkr-h615) (Rasanen et al., 2021) are the: (i) PNW paleoliquefaction database; (ii) maps showing the predicted median and variance of *PGA* and *PGV* across 30 simulations of a CSZ earthquake; (iii) maps of forecasted *PoL* and *LSE* for 30 simulated CSZ earthquakes; and (iv) locations of existing CPTs, as discussed in the text. In addition, all inputs required of the Rashidan and Baise (2020) geospatial liquefaction model are globally available. Distance to river was computed from the Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS) database (https://www.hydrosheds.org/page/hydrorivers, accessed July 2019). Distance to coast was computed from the National Oceanic and Atmospheric Administration coastline dataset (https://shoreline.noaa.gov/data/datasheets/medres.html, accessed July 2019). Mean annual precipitation was obtained from the WorldClim database (https://worldclim.org/, accessed July 2019).

7. Acknowledgements

The presented study is based on work supported by the National Science Foundation (NSF) under Grant No. CMMI-1751216, by the NSF Graduate Research Fellowship Program under Grant No. DGE-1762114, and by the University of Washington Royalty Research Fund (RRF). However, any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF or RRF.

8. References

Allen, T.I. and Wald, D.J. (2009). "On the use of high-resolution topographic data as a proxy for seismic site conditions (Vs30)." *Bulletin of the Seismological Society of America* 99(2A): 935-943.

Ambrayses, N.N. (1988). "Engineering seismology." *Earthquake Eng and Structural Dynamics* 17: 1-105.

Atwater, B.F. (1992). "Geologic Evidence for Earthquakes During the Past 2000 Years Along the Copalis River, Southern Coastal Washington." *Journal of Geophysical Research* 97(B2): 1901-1919.

Atwater, B.F. (compiler) (1994). "Geology of Holocene liquefaction features along the lower Columbia River at Marsh, Brush, Price, Hunting, and Wallace Islands, Oregon and Washington." *U.S. Geol. Surv. OpenFile Rept.* 94-209, 30 pp.

Atwater, B.F. (1999). "Radiocarbon dating of a Seattle earthquake to A.D. 900–930." *SRL* 70: 232.

Atwater, B.F., and Hemphill-Haley, E. (1997). “Recurrence Intervals for Great Earthquakes of the Past 3,500 Years at Northeastern Willapa Bay, Washington.” *U.S. Geological Survey Paper 1576.*

Atwater, B.F., Stuiver, M., and Yamaguchi, D.K. (1991). “Radiocarbon test of earthquake magnitude at the Cascadia subduction zone.” *Nature* 353: 156-158.

Atwater, B.F., Carson, B., Griggs, G.B., Johnson, H.P., and Salmi, M.S. (2014). “Rethinking turbidite paleoseismology along the Cascadia subduction zone.” *Geology* 42(9): 827-830.

Atwater, B.F., Tuttle, M.P., Schweig, E.S., Rubin, C.M., Yamaguchi, D.K., and Hemphill-Haley, E. (2003). “Earthquake recurrence inferred from paleoseismology.” *Developments in Quaternary Science* 1: 331-348.

Aydan, O., Ulusay, R., Kumsar, H., and Tuncay, E. (2000). “Site investigation and engineering evaluation of the Duzce-Bolu earthquake of November 12, 1999.” *Turkish Earthquake Foundation, Istanbul*. Report No. TDV/DR 09-51, 307 p.

Bastin, S., Bassett, K., Quigley, M.C., Maurer, B.W., Green, R.A., Bradley, B.A., and Jacobson, D. (2016). “Late Holocene liquefaction at sites of contemporary liquefaction during the 2010-2011 Canterbury Earthquake Sequence.” *Bulletin of the Seismological Society of America* 106(3): 881-903.

Bourgeois, J., and Johnson, S.Y. (2001). “Geologic evidence of earthquakes at the Snohomish delta, Washington, in the past 1200 yr.” *Geological Society of America Bulletin* 113(4): 482-494.

Briggs, G.G. (1994). "Coastal Crossing of the Elastic Strain Zero-Isobase, Cascadia Margin, South Central Oregon Coast.” *Dissertations and Theses*. Paper 4739.

Castilla, R.A. and Audemard, F.A. (2007). “Sand blows as a potential tool for magnitude estimation of pre-instrumental earthquakes.” *Journal of Seismology* 11: 473-487.

Clague, J.J., Naesgaard, E., and Sy, A. (1992). “Liquefaction features on the Fraser delta: evidence for prehistoric earthquakes?” *Canadian Journal of Earth Sciences* 29: 1734.1745.

Clague, J.J., Naesgaard, E., and Nelson, A.R. (1997). “Age and significance of earthquake-induced liquefaction near Vancouver, British Columbia, Canada.” *Can. Geotech. J.* 34: 53-62.

Cox, R.T., Hill, A.A., Larsen, D., Holzer, T., Forman, S.L., Noce, T., Gardner, C., and Morat, J. (2007). “Seismotectonic implications of sand blows in the southern Mississippi Embayment.” *Engineering* Geology 89: 278-299.

Darendeli, M. B. (2001). “Development of a new family of normalized modulus reduction and material damping curves.” PhD Dissertation, Dept. of Civil Engineering, University of Texas at Austin.

Davis, E. (2019). “Seattle liquefaction features along the Duwamish Waterway, Washington.” *Seismological Society of America Annual Meeting*, 23-26 April, Seattle, USA.

de Magistris, F. S., Lanzano, G., Forte, G., & Fabbrocino, G. (2013). “A database for PGA threshold in liquefaction occurrence.” *Soil Dynamics and Earthquake Engineering*, 54: 17-19.

Engelhart, S.E., Horton, B.P., Nelson, A.R., Hawkes, A.D., Witter, R.C., Wang, K., Wang, P.-L., and Vane, C.H. (2013). “Testing the use of microfossils to reconstruct great earthquakes at Cascadia.” *Geology* 41 (10): 1067–1070.

Fan, Y., Li, H., and Miguez-Macho, G. (2013). “Global patterns of groundwater table depth.” *Science* 339: 940 943.

Fiedorowicz, B.K. (1997). “Geologic Evidence of Historic and Preshistoric Tsunami Inundation at Seaside, Oregon.” *M.S. Thesis,* Portland State University, Portland.

Frankel, A., Wirth, E., Marafi, N., Vidale, J., and Stephenson, W. (2018a). “Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust Based On 3D Simulations and Stochastic Synthetics (Part 1): Methodology and Overall Results.” *Bulletin of the Seismological Society of America* 108 (5A): 2347-2369.

Frankel, A., Wirth, E., and Marafi, N. (2018b). "The M9 Project Ground Motions." DesignSafe-CI. https://doi.org/10.17603/DS2WM3W.

Galli, P. (2000). "New empirical relationships between magnitude and distance for liquefaction." *Tectonophysics* 324: 169-187.

Geyin, M. and Maurer, B.W. (2020). "Fragility functions for liquefaction-induced ground failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.

Geyin, M., Baird, A.J., and Maurer, B.W. (2020). "Field assessment of liquefaction prediction models based on geotechnical vs. geospatial data, with lessons for each." *Earthquake Spectra* 36(3): 1386–1411.

Geyin, M., Maurer, B.W., and Christofferson, K. (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Eng., In Review.*

Gheibi, E., Gassman, S., & Talwani, P. (2020). "Regional assessment of prehistoric earthquake magnitudes in the South Carolina Coastal Plain." *Bulletin of Eng Geology and the Environment 79*(3): 1413-1427.

Goldfinger, C., Nelson, C.H., Morey, A.E., Johnson, J.E., Patton, J.R., Karabanov, E., Gutiérrez-Pastor, J., Eriksson, A.T., Gràcia, E., Dunhill, G., Enkin, R.J., Dallimore, A., and Vallier, T., (2012). "Turbidite event history - methods and implications for Holocene paleoseismicity of the Cascadia subduction zone." *U.S. Geological Survey Professional Paper 1661–F*, 170 pg.

Green, R.A., Olson, S.M., and Obermeier, S.F. (2005). "Geotechnical analysis of paleoseismic shaking using liquefaction effects: field examples." *Engineering Geology* 76: 263-293.

Green, R.A., Maurer, B.W., Bradley, B.A., Wotherspoon, L., and Cubrinovski, M. (2014). "Implications from liquefaction observations in New Zealand for interpreting paleoliquefaction data in the central eastern United States." U.S. Geological Society Technical Report G12AP20002, 97pp.

Green, R.A., Bommer, J.J., Rodriguez-Marek, A., Maurer, B.W., Stafford, P.J., Edwards, B., Kruiver, P.P., De Lange, G. and Van Elk, J. (2019). "Addressing limitations in existing 'simplified' liquefaction triggering evaluation procedures: application to induced seismicity in the Groningen gas field." *Bulletin of Earthquake Engineering* 17(8): 4539-4557.

Hayati, H. and Andrus, R.D. (2008). "Liquefaction potential map of Charleston, South Carolina base on the 1886 earthquake." *Journal of Geotechnical and Geoenvironmental Engineering* 134(6): 815-828.

Heath, D., Wald, D. J., Worden, C. B., Thompson, E. M., and Smoczyk, G. (2020). "A Global Hybrid VS30 Map with a Topographic-Slope-Based Default and Regional Map Insets." *Earthquake Spectra* 36(3): 1570-1584.

Holzer, T.L., Noce, T.E., and Bennett, M.J. (2015). "Strong ground motion inferred from liquefaction caused by the 1811-1812 New Madrid, Missouri, Earthquakes." *Bulletin of the Seismological Society of America* 105(5): 2589-2603.

Hu, K., Gassman, S.L., and Talwani, P. (2002). "Magnitudes of prehistoric earthquakes in the South Carolina Coastal Plain from geotechnical data." *Seismological Research Letters* 73(6): 979-991.

Ishihara, K. (1985). "Stability of natural deposits during earthquakes." In *International conference on soil mechanics and foundation engineering. 11* (pp. 321-376).

Iwasaki, T., Tatsuoka, F., Tokida, K., and Yasud, S. (1978). "A practical method for assessing soil liquefaction potential based on case studies at various sites in Japan." *2nd Intl Conf. Microzonation*.

Jeschke, D. A.; Eungard, D. W.; Troost, K. G.; Wisher, A. P, (2019). "Subsurface database of Washington State - GIS data." *Washington Geological Survey Digital Data Series 11, version 2.1,* previously released August 2017 [http://www.dnr.wa.gov/publications/ger_portal_subsurface_database.zip].

Kelsey, H.M., Witter, R.C., Hemphill-Haley, E. (2002). "Plate-boundary earthquakes and tsunamis of the past 5500 yr, Sixes River estuary, southern Oregon. *GSA Bulletin* 114(3): 298-314.

Kelsey, H.M., Sherrod, B., Johnson, S.Y., and Dadisman, S.V. (2004). “Land-level changes from a late Holocene earthquake in the northern Puget Lowland, Washington.” *Geological Society of America* 32(6): 469-472.

Kottke, A.R. and Rathje, E.M. (2008). “Technical manual for Strata.” Report No.: 2008/10. Pacific earthquake engineering research center. Berkeley: University of California.

Kottke, A.R. (2020). “Site response analysis for Python.” https://doi.org/10.5281/ zenodo.3522104. https://github.com/arkottke/pysra. Accessed 1 Feb 2020.

Kuribayahsi, E. and Tatsuoka, F. (1975). “Brief review of liquefaction during earthquakes in Japan.” *Soils and Foundations* 15(4): 81-92.

Marafi, N.A., Eberhard, M.O., Berman, J.W., Wirth, E.A., and Frankel, A.D. (2019). “Impacts of simulated M9 Cascadia subduction zone motions on idealized systems.” *Earthquake Spectra 35*(3), 1261-1287.

Marafi, N.A., Makdisi, A.J., Eberhard, M.O., & Berman, J.W. (2020). “Impacts of an M9 Cascadia Subduction Zone Earthquake and Seattle Basin on Performance of RC Core Wall Buildings.” *Journal of Structural Engineering* 146(2), 04019201.

Marafi, N. A., Grant, A., Maurer, B. W., Rateria, G., Eberhard, M. O., & Berman, J. W. (2021). “A generic soil velocity model that accounts for near-surface conditions and deeper geologic structure.” *Soil Dynamics and Earthquake Engineering 140*: 106461.

Martin, J.R. and Clough, G.W. (1990). “Implications from a geotechnical investigation of liquefaction phenomena associated with seismic events in the Charleston, SC area.” Report to USGS Grant No. 14-08-001-G-1348, Virginia Tech, Blacksburg, VA, 414 p.

Martin, M.A., and Bourgeois, J. (2012). “Vented sediments and tsunami deposits in the Puget Lowland, Washington – differentiating sedimentary processes.” *Sedimentology* 59: 419-444.

Maurer, B.W., Green, R.A., Quigley, M.C. and Bastin, S. (2015a). “Development of magnitude-bound relations for paleoliquefaction analyses: New Zealand case study.” *Engineering Geology* 197: 253-266.

Maurer, B.W., Green, R.A. and Taylor, O.D.S. (2015b). “Moving towards an improved index for assessing liquefaction hazard: lessons from historical data.” *Soils and Foundations* 55(4): 778-787.

Maurer, B.W., Green, R.A., Cubrinovski, M., and Bradley, B. (2015c). “Assessment of CPT-based methods for liquefaction evaluation in a liquefaction potential index framework.” *Géotechnique* 65(5): 328-336.

Maurer, B.W., Green, R.A., Wotherspoon, L.M. & Bastin, S. (2019). “The stratigraphy of compound sand blows at sites of recurrent liquefaction: implications for paleoseismicity studies.” *Earthquake Spectra* 35(3): 1421-1440.

National Research Council (NRC) (2016). “State of the art and practice in the assessment of earthquake induced soil liquefaction and its consequences.” committee on earthquake induced soil liquefaction assessment (Edward Kavazanjian, Jr., Chair, Jose E. Andrade, Kandian “Arul” Arulmoli, Brian F. Atwater, John T. Christian, Russell A. Green, Steven L. Kramer, Lelio Mejia, James K. Mitchell, Ellen Rathje, James R. Rice, and Yumie Wang), The National Academies Press, Washington, DC.

Nuclear Regulatory Commission (NRC) (2012). “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” Tech.Rep.NUREG-2115, NRC, Rockville, MD.

O’Donnell III, R.J., Hawkes, A.D., Lane, C.S., Engelhard, S.E., Horton, B.P., Bobrowsky, P., Sawai, Y., Witter, R.C., Nelson, A.R., and Tanigawa, K. (2017). “Assessing the utility of δ13C and bulk geochemistry in estuaries along the Cascadia subduction zone for coastal paleoseismology.” *Geological Society of America Abstracts with Programs* 49(6): 58-2.

Obermeier, S. F. (1995). “Preliminary estimates of the strength of prehistoric shaking in the Columbia River Valley and the southern half of coastal Washington, with emphasis for a Cascadia Subduction Zone earthquake about 300 years ago.” *U.S. Geol. Surv.* Open-File Rept. 94-589, 46 pp.

Obermeier, S. F. (1996). “Use of liquefaction-induced features for paleoseismic analysis—an overview of how seismic liquefaction features can be distinguished from other features and how their regional distribution and properties of source sediment can be used to infer the location and strength of Holocene paleo-earthquakes.” *Engineering Geology, 44*(1-4): 1-76.

Obermeier, S.F. and Dickenson, S.E. (2000). “Liquefaction evidence for the strength of ground motions resulting from late Holocene Cascadia subduction earthquakes, with emphasis on the event of 1700 A.D.” *Bulletin of the Seismological Society of America* 90(4): 876-896.

Obermeier, S.F., Pond, E.C., Olsen, S.M. with contributions by Green, R.A, Mitchell, J.K., and Stark, T.D. (2001). “Paleoliquefaction studies in continental settings: geologic and geotechnical factors in interpretations and back-analysis.” U.S. Geological Survey Open-File Report 01–029.

Obermeier, S.F., Olson, S.M. and Green, R.A (2005). “Field occurrences of liquefaction-induced features: a primer for engineering and geologic analysis of paleoseismic shaking.” *Engineering Geology* 76: 209-234.

Obermeier, S.F., Olson, S.M., Green, R.A., Schweig, E., and Williams, R. (2011). “Clastic dikes and ground fractures; seismic or not?” *Seismological Research Letters* 82: 335.

Olson, S.M., Green, R.A. and Obermeier, S.F. (2005a). “Revised magnitude bound relation for the Wabash valley seismic zone of the central United States.” *Seismological Research Letters* 76(6): 756–771.

Olson, S.M., Green, R.A. and Obermeier, S.F. (2005b). “Geotechnical analysis of paleoseismic shaking using liquefaction features: a major updating.” *Engineering Geology* 76: 235–261.

Palmer, S.P., Magsino, S.L., Bilderback, E.L., Poelstra, J.L., Folger, D.S.; Niggemann, R.A. (2007). “Liquefaction susceptibility and site class maps of Washington State, by county.” *Washington Division of Geology and Earth Resources Open File Report 2004-20*.

Papadopoulos, G.A. and Lefkopoulos, G. (1993). “Magnitude-distance relations for liquefaction in soil from earthquakes.” *Bulletin of the Seismological Society of America* 83(3): 925-938.

Papathanassiou, G., Pavlides, S., Christaras, B. and Pitilakis, K. (2005). “Liquefaction case histories and empirical relations of earthquake magnitude versus distance from the broader Aegean region.” *Journal of Geodynamics* 40: 257-278.

Peters, R., Jaffe, B., and Gelfenbaum, G. (2007). “Distribution and sedimentary characteristics of tsunami deposits along the Cascadia margin of western North America.” *Sed Geo* 200: 372-386.

Petersen, M.D., Moschetti, M.P., Powers, P.M., Mueller, C.S., Haller, K.M., Frankel, A.D., Zeng, Y., Rezaeian, S., Harmsen, S.C., Boyd, O.S., Field, N., Chen, R., Rukstales, K.S., Luco, N., Wheeler, R.L., Williams, R.A. and Olsen, A.H. (2014). “Documentation for the 2014 update of the United States national seismic hazard maps.” USGS Open-File Report 2014–1091, 243 p.

Peterson C.D. and Madin, I.P. (1997). “Coseismic Paleoliquefaction Evidence in the Central Cascadia Margin, USA.” *Oregon Geology* 59, 51-74.

Peterson, C.D., Cruikshank, K.M., Jol, H.M. and Schilichting, R.B. (2008). “Minimum Runup Heights of Paleotsunami from Evidence of Sand Ridge Overtopping at Cannon Beach, Oregon, Central Cascadia Margin, USA.” *Journal of Sedimentary Research* 78: 390-409. http://dx.doi.org/10.2110/jsr.2008.044

Peterson, C.D., Cruikshank, K.M., Darienzo, M.E., Wessen, G., Butler, V. and Sterling, S. (2013). “Coseismic Subsidence and Paleotsunami Runup Records from Latest Holocene Deposits in the Waatch Valley, Neah Bay, Northwest Washington, USA: Links to Great Earthquakes in the Northern Cascadia Margin.” *Journal of Coastal Research*, 29, 157-172. http://dx.doi.org/10.2112/JCOASTRES-D-12-00031.1

Peterson, C.D., et al. (2014). “Large-Scale Fluidization Features from Late Holocene Coseismic Paleoliquefaction in the Willamette River Forearc Valley, Central Cascadia Subduction Zone, Oregon, USA.” *Open Journal of Earthquake Research*, 3: 82-99

Peterson, C.D., Jol, H.M. and Alldritt, K. (2005). "Reconnaissance Subsurface Investigation of Bandon Marsh, Oregon." U.S. Fish and Wildlife Service, Newport.

Pirrotta, C., Barbano, M.S., Guarnieri, P. and Gerardi, F. (2007). "A new dataset and empirical relationships between magnitude/intensity and epicentral distance for liquefaction in central-eastern Sicily." *Annals of Geophysics* 50(6): 763-774.

Polenz, M., Czajkowski, J.L., Paulin, G.L., Contreras, T.A., Miller, B.A., Martin, M.E., Walsh, T.J., Logan, R.L., Carson, R.J., Johnson, C.N., Skov, R.H., Mahan, S.A., Cohan, C.R. (2010). "Geological Map of the Skokomish Valley and Union 7.5-minute Quadrangles, Mason County, Washington. *Washington Division of Geology and Earth Resources* Open File Report 2010-3, 1 sheet, scale 1:24,000, with 21 p. text.

Quigley, M.C., Bastin, S., Bradley, B.A. (2013). "Recurrent liquefaction in Christchurch, New Zealand, during the Canterbury earthquake sequence." *Geology* 41: 419–422.

Rasanen, R., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." DesignSafe-CI. https://doi.org/10.17603/ds2-fqkr-h615.

Rasanen, R., and Maurer, B.W. (2021). "Probabilistic seismic source location and magnitude via inverse analysis of liquefaction evidence." *Earthquake Spectra*, In Review, Elsevier.

Rashidian, V., and Baise, L.G. (2020). "Regional efficacy of a global geospatial liquefaction model." *Engineering Geology* 272: 105644.

Rodriguez-Marek, A., and Ciani, D. (2008). "Probabilistic methodology for the analysis of paleoliquefaction features." *Engineering Geology* 96(3): 159-172.

Satake, K., Shimazaki, K., Tsuji, Y. & Ueda, K. (1996). "Time and size of a giant earthquake in Cascadia inferred from Japanese tsunami record of January 1700." *Nature* 379, 246–249.

Sherrod, B.L. (2001). "Evidence for earthquake-induced subsidence about 1100 yr ago in coastal marches of southern Puget Sound, Washington." *Geo Soc of America Bull* 113(10): 1299-1311.

Sherrod, B., Brocher, T.M., Weaver, C.S., Bucknam, R.C., Blakely, R.J., Kelsey, H.M., Nelson, A.R., Haugerud, R. (2004). "Holocene fault scarps near Tacoma, Washington, USA." *Geological Society of America* 32(1): 9-12.

Sims, J.D. & Garvin, C.D. (1995). "Recurrent liquefaction induced by the 1989 Loma Prieta earthquake and 1990 and 1991 aftershocks: implications for paleoseismicity studies." *Bulletin of the Seismological Society of America* 85(1): 51-65.

Sims, J.D. (2002). "Paleoliquefaction Study of the Willamette River Valley, Portland to Corvallis, Oregon." *US Geological Survey*, Reston, Element II under Contract 02HQGR0021

Stephenson, W.J., Reitman, N.G., and Angster, S.J., (2017). "P- and S-wave velocity models incorporating the Cascadia subduction zone for 3D earthquake ground motion simulations, version 1.6." *U.S. Geological Survey Open-File Rep*ort 2017–1152, 17 p., https://doi.org/10.3133/ofr20171152.

Struble, W., Roering, J., Burns, W.J., Black, B., Calhoun, N., and Wetherell, L.R. (2017). "Propensity for deep-seated landslides in the Oregon coastal ranges during Cascadia megathrust earthquake through dendrochronological dating of landslide-dammed lakes." *GSA Abstracts* 49(6): 58-7.

Stuiver, M., Reimer, P.J., and Reimer, R.W. (2020). CALIB 7.1 [WWW program] at http://calib.org, accessed 2020-3-18.

Takada, K. and Atwater, B.F. (2004). "Evidence for Liquefaction Identified in Peeled Slices of Holocene Deposits along the Lower Columbia River, Washington." *BSSA* 94: 550-575.

Talwani, P. and Schaeffer, W.T. (2001). "Recurrence rates of large earthquakes in the South Carolina coastal plain based on paleoliquefaction data." *Journal of Geophysical Research* 106: 6621-6642.

Thrush, C., and Ludwin, R.S. (2007). “Finding fault: Indigenous seismology, colonial science, and the rediscovery of earthquakes and tsunamis in Cascadia.” *American Indian Culture and Res J* 31(4): 1-24.

Tuttle, M.P. (2001). “The use of liquefaction features in paleoseismology: lessons learned in the New Madrid seismic zone, central United States.” *Journal of Seismology* 5: 361-380.

Tuttle, M. and Hartleb, R. (2012). “CEUS Paleo-liquefaction Database, Uncertainties Associated with Paleo-liquefaction Data, and Guidance for Seismic Source Characterization: Central and Eastern United States Seismic Source Characterization for Nuclear Facilities, Appendix E.” Report NUREG-2115, U.S. Nuclear Regulatory Commission (NRC), Rockville, MD.

Tuttle, M.P., Dyer-Williams, K., and Barstow, N.L. (2002a). “Paleoliquefaction study of the Clarendon-Lindon fault system, western New York State.” *Tectonophysics* 353: 263-286.

Tuttle, M.P., Schweig, E.S., Sims, J.D., Lafferty, R.H., Wolf, L.W. and Haynes, M.L. (2002b). “The earthquake potential of the New Madrid Seismic Zone.” *Bull of the Seismological Society of America* 92(6): 2080-2089.

Tuttle, M.P., Schweig, E.S., Campbell, J., Thomas, P.M., Sims, J.D. and Lafferty, R.H. (2005). “Evidence for New Madrid earthquakes in A.D. 300 and 2350 B.C.” *Seismological Research Letters* 76: 489-501.

USGS (2017). “M 9.3 scenario earthquake – Cascadia megathrust – whole CSZ characteristic M branch.” *USGS National Earthquake Information Center*, available from < https://earthquake.usgs.gov/scenarios/eventpage/bssc2014cascadia_sub0_m9p34_se#executive>.

van Ballegooy, S., Malan, P., Lacrosse, V., Jacka, M.E., Cubrinovski, M., Bray, J.D., O’Rourke, T.D., Crawford, S.A. and Cowan, H. (2014). “Assessment of liquefaction-induced land damage for residential Christchurch.” *Earthquake Spectra* 30(1): 31-55.

Wair, B. R., DeJong, J. T., & Shantz, T. (2012). “Guidelines for estimation of shear wave velocity profiles.” Pacific Earthquake Engineering Research Center.

Wakamatsu, K. (1993). “History of soil liquefaction in Japan and assessment of liquefaction potential based on geomorphology.” A Thesis in the Department of Architecture Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Engineering, Waseda University, Tokyo, Japan.

Watt, J.T., Brothers, D.A., Bennett, S.E.K., Kluesner, J.W., Roland, E., Conrad, J.E., Sliter, R.W., Dartnell, P., Goldfinger, C., Patton, J., Michalak, M.J., Sherrod, B.L., Gomberg, J., and Wells, R.E. (2017). “Toward a systematic characterization of Cascadia upper plate morphology, structure, and Quaternary deformation history: an integrated onshore-offshore approach.” *GSA Abstracts* 49(6): 58.

Whistler, J.E., Atwater, B.F. and Montgomery, D.R. (2002). “Holocene liquefaction near the Seattle fault at the Issaquah Creek delta”. *Eos* 83: S22B–1036.

Wirth, E.A., Frankel, A.D., Marafi, N., Vidale, J. E., and Stephenson, W. J. (2018). “Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust based on 3-D Simulations and Stochastic Synthetics (Part 2): Rupture Parameters and Variability.” *Bulletin of the Seismological Society of America*, 108 (5A): 2370-2388.

Youd, T.L., and Noble, S.K. (1997). "Liquefaction criteria based on statistical and probabilistic analyses." In *Proc., NCEER Workshop on Evaluation of Liquefaction Resistance of Soils, NCEER Technical Rep. No: NCEER-97* 22:201-205.

Zehfuss, Z.H. (2005). “Distal records of sandy Holocene lahars from Mount Rainier, Washington”. Ph.D. dissertation, University of Washington, Seattle, Washington.

Zhu, J., Baise, L.G., and Thompson, E.M. (2017). “An updated geospatial liquefaction model for global application.” *Bulletin of the Seismological Society of America*, 107(3): 1365–1385.

APPENDIX F

Information describing publication 6 (Dataset 1), highlighted below, is provided on the following pages. Specifically, the dataset's posting on NHERI DesignSafe is shown below.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.
[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

PRJ-3102 | **Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)**

Download Dataset

Author(s)	**Rasanen, Ryan; Marafi, Nasser; Maurer, Brett**
Data Type	**Database**
Awards	**CAREER: Decoding the Enigmas of U.S. Seismic Hazard Via Multi-Scale, Multi-Physics Approaches to Paleoliquefaction Analysis - CMMI-1751216**
Keywords	**Paleoliquefaction, Paleoseismology, Seismic Hazard, Pacific Northwest, Washington, Oregon**
Date of Publication	**04-15-2021**
DOI Citation	**10.17603/ds2-fqkr-h615**
License(s)	**Open Data Commons Attribution** **Creative Commons Attribution Share Alike**

View Data Metrics | Leave Feedback

Description | In the U.S. Pacific Northwest (PNW), the historic earthquake record is often insufficient to provide inputs to seismic-hazard analyses or to inform ground-motion predictions for certain seismic sources (e.g., the Cascadia Subduction Zone, CSZ). As a result, paleoseismic studies are commonly used to infer information about the seismic hazard. However, among the many forms of coseismic evidence, soil liquefaction provides the best, if not only, evidence from which the intensities of previous ground motions may be constrained. Accordingly, the overarching goal of this research…
Show More

PRJ-3102

Name	Size	Last modified
Data Paper Manuscript.pdf ***Report***	2.6 MB	4/14/21 5:11 PM
Geospatial_Liquefaction_PNW.mpk ***Map***	2.6 GB	4/14/21 5:11 PM
Paleoliquefaction_Database_PNW.mpk ***Data***	524.4 kB	4/14/21 5:11 PM
Paleoliquefaction_Database_PNW.xlsx ***Data***	73.0 kB	4/14/21 5:11 PM
Quick ReadME.pdf ***README***	28.5 kB	4/14/21 5:11 PM
Washington_State_CPTs.mpk ***Map***	102.5 kB	4/14/21 5:11 PM

APPENDIX G

Publication 7 (Journal Paper 5), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.
[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

Select Liquefaction Case Histories from the 2001 Nisqually, Washington Earthquake: A Digital Dataset and Assessment of Model Performance

Ryan A. Rasanen,[a] **M.EERI, Mertcan Geyin,**[b] **and Brett W. Maurer,**[a] **M.EERI**

While soil liquefaction is common in earthquakes, the case history data required to train and test state-of-practice prediction models remains comparatively scarce, owing to the breadth and expense of data that comprise a single case history. The 2001 Nisqually, Washington, earthquake, for example, occurred in a metropolitan region and induced damaging liquefaction in the urban cores of Seattle and Olympia, yet case history data has not previously been published. Accordingly, this paper compiles 24 cone-penetration-test (CPT) case histories from free-field locations. The many methods used to obtain and process the data are detailed herein, as is the structure of the digital dataset. The case histories are then analyzed by 18 existing liquefaction response models to determine whether any is better, and to compare model performance in Nisqually against global observations. While differences are measured, both between models and against prior global case histories, these differences are often statistically insignificant considering finite-sample uncertainty. This alludes to the general inappropriateness of championing models based on individual earthquakes or otherwise small datasets, and to the ongoing needs for additional case history data and more rigorous adherence to best practices in model training and testing.

Introduction

The 28 February 2001 M_w6.8 Nisqually, Washington, earthquake damaged infrastructure throughout the Puget Sound Region of the U.S. Pacific Northwest. This included damage due to soil liquefaction, which was concentrated in Holocene alluvium and artificial fills in the urban

[a] Department of Civil and Environmental Engineering, University of Washington
[b] Norwegian Geotechnical Institute

areas of Olympia and Seattle (e.g., Bray et al., 2001; Seattle Office of Emergency Management, 2019). Intraslab earthquakes associated with the Cascadia Subduction zone, such as the 2001 Nisqually earthquake, are a common occurrence deep beneath the Puget Sound. Prior to the 2001 event, M_w7.1 and M_w6.7 ruptures occurred in 1949 and 1965, respectively, causing widespread ground failures (Chleborad and Shuster, 1990), often in areas where liquefaction was again observed in 2001. But unlike these prior events, the most recent was recorded by a dense network of strong motion stations across the affected region and occurred at a time in which in-situ geotechnical tests were commonplace. Although manifestations of liquefaction, or the lack thereof, were documented by reconnaissance teams in 2001, liquefaction case histories from the event have not been published.

While several "tiers" of liquefaction prediction model exist, including "geospatial" (e.g., Zhu et al., 2017; Rashidian and Baise, 2020; Geyin et al., 2021), "semi-empirical stress based" (e.g., Robertson and Wride, 1998; Moss et al., 2006; Boulanger and Idriss, 2016), and "constitutive, fully mechanistic" (e.g., Boulanger and Ziotopoulou, 2015), the second of these is the most common in routine practice. All models within this tier are based on the same modeling framework originally proposed by Seed and Idriss (1971) and Whitman (1971). This framework is semi-empirical in that the soil's resistance to liquefaction is correlated to an in-situ geotechnical test measurement, which foremost serves as a proxy of soil density. Among the in-situ tests for which models have been developed, the cone penetration test (CPT) is widely perceived to offer important advantages (NRC, 2016). The training of a CPT-based liquefaction triggering model requires several types of field data that collectively form a "case history." These include: (i) observations of liquefaction, or the lack thereof, which are almost always made at the ground surface; (ii) a profile of CPT data from the site of observation; (iii) knowledge of the groundwater table (*GWT*) depth, both at the time of the earthquake and at the time of CPT testing; and (iv) knowledge of the peak ground acceleration (*PGA*) at the site.

Owing to the breadth and expense of these data, liquefaction case histories have remained relatively scarce. Aside from the 2010-2011 Canterbury sequence, from which a large quantity of data was compiled in the city of Christchurch (e.g., Geyin et al., 2021), fewer than 300 high quality CPT case histories have been compiled from all other global earthquakes combined (e.g., Boulanger and Idriss, 2016). The lack of data from a wide variety of seismologic, geologic, and geomorphic settings inherently constrains the development of better prediction models, given the semi-empirical nature of the “stress-based” framework. The absence of case-history data from a 21st-century earthquake in a major U.S. city is especially notable, given that case histories have been compiled from relatively rural settings without many seismic instruments, and from events that predate modern reconnaissance tools and the CPT itself. Accordingly, this paper compiles a digital dataset of 24 CPT-based liquefaction case histories from the Nisqually earthquake. The cases are compiled from the free field on ground that is generally level, and the data is provided in a widely accessible format. These cases are among very few associated with subduction zone earthquakes, and potentially the only cases from an intraslab rupture. This compilation is also the second largest from any U.S. earthquake and the fourth largest from any earthquake outside of Canterbury, New Zealand. Only the 1989 Loma Prieta (U.S), 1999 Chi-Chi (Taiwan), and 2012 Emilia (Italy) earthquakes have produced more CPT case histories to date. Moreover, the newly compiled cases are from a region underlaid by unusually deep sedimentary basins, with $Z_{1.0}$ values (i.e., depths to a shear-wave velocity of 1 km/s) exceeding 800 m. The Nisqually case histories thus provide a useful addition for model training and testing. Following compilation and discussion of these new cases, the data are used to test the efficacies of 18 existing liquefaction response models that predict the triggering and manifestation of liquefaction. These results are compared to prior analyses of global cases to preliminarily assess whether, and how, model performance differs for the Nisqually dataset.

The paper is presented in two parts. The first pertains to the compilation of liquefaction case histories from the 2001 Nisqually earthquake. The various methods used to process and populate the database are detailed, as are the structure and formatting of the resulting dataset. The second pertains to the performance evaluation of eighteen popular liquefaction response models using the newly compiled case histories.

Compilation of Case Histories

Twenty-four liquefaction case histories are herein compiled from the 2001 Nisqually earthquake, which was epicentrally located in the vicinity of the Nisqually River’s confluence with the Puget Sound, and which had an estimated depth to top of rupture (Z_{TOR}) of 50 km (e.g., Bustin et al., 2004). The event was recorded by a network of more than 60 strong ground-motion stations located within 100 km of the fault rupture. A regional overview of the case-history sites, ground-motion recording stations, and location of fault rupture is shown in Figure 1. Of the 24 case-history sites, 17 map to artificial fills and the remainder are generally in Holocene alluvium (Washington Division of Geology and Earth Resources, 2016). Of relevance to the development of "magnitude-bound" relationships (e.g., Rasanen et al., 2021), the most distal observation of liquefaction had an epicentral, Joyner Boore, and rupture distance of R_{EPI} = 74 km, R_{JB} = 73 km, and R_{RUP} = 90 km, respectively. A succinct summary of the compiled case histories is given in Table 1. The details presented therein, in addition to many others, are discussed subsequently, and an extended version of the table is provided in the appendix.

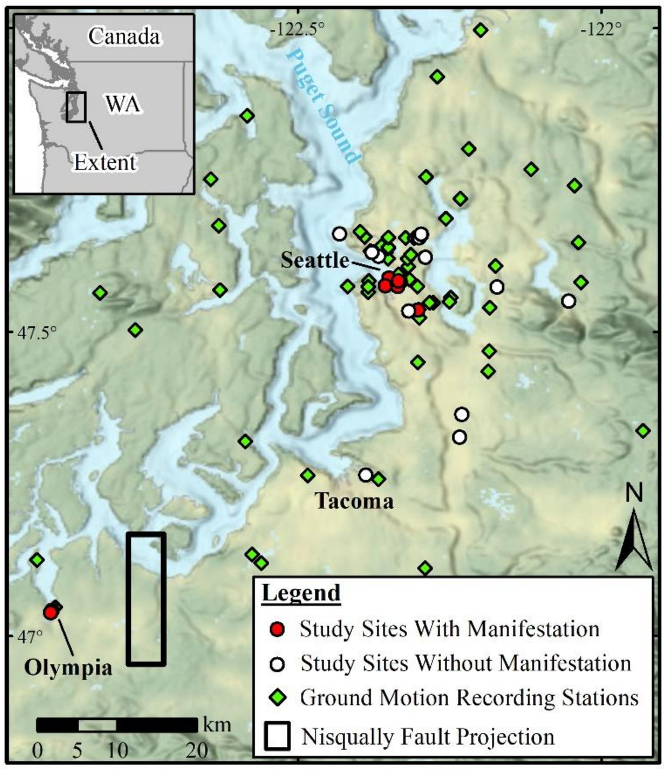


Figure 1. Regional overview of case-history locations and ground motion recording stations.

Liquefaction Response

Case histories were compiled with emphasis on free-field level-ground sites. Nine of the twenty-four case history sites are herein classified as "positive," meaning that surficial manifestations of liquefaction were observed. These consisted of: (i) seven sites with varying amounts of liquefaction ejecta, and in some cases, ground cracks and perceptible vertical settlements; (ii) one site (Capitol Lake-02) having both traces of liquefaction ejecta and a crack with ~5 cm of horizontal displacement, suggesting the possibility of minor lateral spreading toward a nearby lake; and (iii) one site (SODO-02) that lacked ejecta, but where cracks and differential settlements were observed in pavements and sidewalks on the perimeter of an improved area. The remaining fifteen sites are classified as “negative,” meaning that no surficial evidence of liquefaction was observed. In general, these latter sites are heavily trafficked, high visibility locations where focused reconnaissance efforts failed to identify any evidence of liquefaction from observations of the ground surface, pavements, infrastructure, etc. In each of these cases, CPTs and borings suggest the presence of soils susceptible to liquefaction and estimated ground motions (discussed subsequently) were sufficient to trigger liquefaction in soils with low cyclic resistance (e.g., de Magistrals, 2013). In addition, a large majority of the case-history sites – both with and without manifestations – are locations where ground failures were previously observed in the 1949 and/or 1965 earthquakes, as documented by Chleborad and Schuster (1990). Sources of reconnaissance information included Bray et al. (2001), Seattle Office of Emergency Management (2019), Pacific Northwest Seismic Network (2001), Hausler and Koelling (2004), Walsh (2001), Geoengineers, Inc. (2003), and personal records and communication with individuals from the City of Seattle, regional utility companies, the University of Washington, and those who previously published reconnaissance reports in the public domain. The imbalance of negative over positive cases is judged to be appropriate. While historical sampling practices have produced datasets that are heavily biased towards positives, such datasets do not accurately represent the population of field observations, wherein liquefaction is typically observed at a small fraction of sites where susceptible soils are present.

Table 1. Summary of compiled case histories. An extended version appears in the appendix.

No.	Site Name	Longitude (WGS84)	Latitude (WGS84)	GWT Depth (m)	Manifestation	Conditional Median PGA (g)	Conditional Lognormal σ of PGA
1	Capitol Lake-01	-122.904273	47.041971	0.61	Yes	0.221	0.192
2	SODO-01	-122.330874	47.582869	3.05	Yes	0.234	0.079
3	Capitol Lake-02	-122.906716	47.038174	1.52	Yes	0.217	0.224
4	Harbor Island-01	-122.349801	47.588704	3.29	Yes	0.180	0.144
5	Harbor Island-02	-122.355378	47.575863	2.44	Yes	0.184	0.167
6	SODO-02	-122.335805	47.578781	2.13	Yes	0.199	0.145
7	SODO-03	-122.334837	47.575020	2.74	Yes	0.198	0.164
8	SODO-04	-122.333936	47.584171	0.99	Yes	0.197	0.099
9	Boeing Field-01	-122.301620	47.536125	2.13	Yes	0.201	0.100
10	Duwamish-01	-122.316733	47.534451	2.29	No	0.192	0.188
11	Discovery Park-01	-122.430819	47.661330	3.96	No	0.084	0.244
12	Elliott Ave-01	-122.366639	47.624508	1.83	No	0.142	0.142
13	Smith Cove-01	-122.376982	47.631682	2.74	No	0.127	0.120
14	Smith Cove-02	-122.377416	47.630920	3.66	No	0.128	0.123
15	Emerald Downs-01	-122.232961	47.327071	1.00	No	0.158	0.353
16	Bush School-01	-122.289623	47.622844	3.05	No	0.119	0.205
17	Factoria-01	-122.171810	47.573752	1.52	No	0.102	0.245
18	UW-01	-122.301101	47.654810	0.76	No	0.129	0.156
19	UW-02	-122.299341	47.655123	0.76	No	0.129	0.166
20	UW-03	-122.295665	47.660272	2.44	No	0.127	0.189
21	UW-04	-122.296197	47.660901	2.44	No	0.126	0.190
22	Issaquah-01	-122.053685	47.550644	2.44	No	0.125	0.279
23	Blair Waterway-01	-122.387559	47.264867	3.05	No	0.104	0.236
24	Green River-01	-122.229279	47.364461	5.18	No	0.166	0.335

Subsurface Data

Case histories were compiled utilizing existing geotechnical tests available in the Jeschke et al. (2019) database, which presently contains more than 100,000 subsurface tests and observations from Washington State, including CPTs, standard penetration tests (SPTs), borings, shear-wave velocity (V_S) measurements, laboratory tests, and well logs. While significantly more CPT case-histories could potentially be compiled from the Nisqually earthquake, the reliance on existing tests in the present effort resulted in 24 sites where the liquefaction response was reliably documented in 2001. A future effort could compile additional case histories, particularly cases based on SPTs, which in the Jeschke et al. (2019) database greatly outnumber CPTs in the affected region. Moreover, these data collectively form the basis for a very large case-history compilation whenever a large earthquake next impacts the region. In compiling case histories, CPTs for which the pre-drill depth exceeded the GWT depth were excluded, given the additional uncertainty associated with such cases. For the remaining cases, CPT data were infilled in the pre-drill zone by averaging the data measured over a 10 cm interval beneath the depth of pre-drill and applying that average uniformly to the pre-drill interval. In the absence of this correction, the recorded data is that of noise as the cone penetrates the open boring. The purpose of this correction, which results in each CPT having an initial series of repeating measurements, is simply to provide reasonable data for approximating soil unit weights and in-situ stresses below the pre-drill zone, should analysts estimate unit weights via CPT correlations. It should thus be understood that data above the reported pre-drill depth is not an actual measurement. Cross-correlation (Buck et al., 2002) was used to ensure that tip and sleeve measurements were properly aligned in accordance with ASTM D5778. For 15 of the 24 CPTs, pore pressure measurements (u_2) are available, and in these cases, the reported cone tip resistance is q_t (i.e., corrected for u_2 effects). Analysts should therefore not make additional corrections when u_2 is provided. For the remainder of CPTs without u_2 measurements, the reported cone tip resistance is q_c. This discrepancy (i.e., combining q_c and q_t measurements) has historically been ignored in liquefaction modeling, given that u_2 measurements are often unavailable and that u_2 effects are negligible in soils susceptible to liquefaction. GWT levels were sourced from the geotechnical reports that documented CPTs while also considering proximal borings, well logs, and monitoring wells. In general, multiple GWT measurements were available in the immediate vicinity of each case-history site. For case histories where surface manifestations were observed, CPTs were typically performed within 10-20 m of the observed sand boil or spreading crack, and occasionally much closer. This is consistent with the methodology adopted by Geyin et al. (2021) to classify case histories in New Zealand and is believed to be consistent with case histories elsewhere, although many authors have not disclosed the locations of in-situ tests relative to liquefaction observations. Nonetheless, and like all liquefaction case histories, it is possible that CPTs do not well represent the subsurface profiles that gave rise to the observed responses. CPTs might differ from these "true" CPTs due to measurement errors or subsurface spatial variability, or because the surficial manifestation was produced by soil that liquefied at some lateral distance from the manifestation, and not directly beneath it (e.g., Bassal and Boulanger, 2021). Thus, the classified manifestation is itself uncertain, just as with any other case history.

Ground Motion Intensities (PGAs)

The Nisqually earthquake was recorded by at least 60 strong-motion records having R_{RUP} < 100 km. On average, the 24 case-history sites are located 1.9 km from the nearest station, with 13 sites located within 1 km of a recording. Inherent to the stress-based framework for predicting liquefaction, the imposed cyclic stress is computed from PGA, and thus, all such existing models use this intensity measure (IM). While future liquefaction response models may use IMs other than PGA, or vector IMs that utilize multiple measures of intensity, PGA continues to be among the most efficient, sufficient, and predictable IMs for the initiation of liquefaction (e.g., Sideras, 2019; Wu et al. 2022). Considering this, and that IMs other than PGA have not been estimated for many prior case histories, it is unlikely that PGA will be supplanted in the very near future.

For the Nisqually earthquake, *PGAs* were estimated at case-history sites using the approach implemented by Bradley (2014) in Canterbury, New Zealand, which has since been used in numerous analyses of Canterbury liquefaction (e.g., Maurer et al., 2015a; Geyin and Maurer, 2021; Rateria and Maurer, 2022). Using this approach, a "conditional" distribution of *PGA* is computed at any given site, wherein predictions from a ground-motion model (GMM), which provides the unconditional distribution, are conditioned on ground-motion stations (SMS). In effect, instrumental *PGA*s and GMM predictions are merged considering site conditions at SMS locations and case-history sites, as well as the spatial correlation of intra-event residuals. Our implementation is analogous to that in Canterbury, except that models appropriate for the Cascadia Subduction Zone are adopted. The *PGA* at any SMS, *i*, may be expressed as:

$$\ln(PGA_i) = \mu_{\ln PGA_i} + \eta + \varepsilon_i, \quad (1)$$

where $\ln(PGA_i)$ is the natural logarithm of the observed *PGA* at SMS *i*; $\mu_{\ln PGA_i}$ is the mean of the natural logarithm of *PGA* at SMS *i* predicted by a GMM, which is a function of various site, path, and source parameters; η is the inter-event residual, which by definition is the same at all locations; and ε_i is the intra-event residual, which varies from site to site but is spatially correlated. Within Eq. 1, the GMM predicts a *PGA* distribution generically defined as:

$$\ln(PGA_i) \sim N(\mu_{\ln PGA_i}, \sigma_\eta^2 + \sigma_\varepsilon^2), \quad (2)$$

where $X \sim N(\mu_X, \sigma_X^2)$ indicates that X has a normal distribution with mean, μ_X, and variance, σ_X^2. With knowledge of *PGA*, η, and ε at SMS sites, conditional *PGA* distributions can be computed at the locations of liquefaction case histories.

First, the Kuehn et al. (2020) subduction zone GMM, which has Cascadia regional adjustment factors, was used to compute unconditional *PGA* distributions at approximately 70 SMS sites in the Puget Sound region. SMS records from dams and those elevated in buildings were first removed. The Kuehn et al. (2020) GMM was implemented for an M_w6.8 Cascadia intraslab rupture, with R_{RUP} computed from a strike of 360º, dip of 70º, and Z_{TOR} of 50 km based on Bustin et al. (2004), Ichinose et al. (2004), and Kao et al. (2008). Regional basin flags were assigned as mapped in Ahdi et al. (2020) and Parker et al. (2020), except for the Seattle basin, for which the updating of Wirth et al. (2018) was used. $Z_{1.0}$ and the depth to a V_S of 2500 m/s ($Z_{2.5}$) were computed from the Stephenson et al. (2017) regional velocity model. Lastly, for a large majority of SMS sites, the time-averaged V_S over the upper 30 m (V_{S30}) was adopted from the site-specific measurements of Wong et al. (2011) or McPhillips et al. (2020). For the remainder, and for most case-history sites, V_{S30} was assigned from the region-specific map of Palmer et al. (2007). Site-specific GMM inputs are provided in Table S1 for each case-history site.

A mixed-effects regression was next used to compute η and the ε_i's for each SMS (e.g., Pinheiro et al., 2008). The covariance matrix of intra-event residuals was then computed by accounting for the spatial correlation between SMS locations and a test site of interest. In this regard, the joint distribution of intra-event residuals at a site of interest and the SMS is:

$$\begin{bmatrix} \varepsilon^{site} \\ \boldsymbol{\varepsilon^{SMstation}} \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon^{site}} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{bmatrix}\right), \quad (3)$$

where $X \sim N(\boldsymbol{\mu_X}, \boldsymbol{\Sigma})$ indicates that X has a multivariate normal distribution with mean $\boldsymbol{\mu_X}$ and covariance matrix $\boldsymbol{\Sigma}$ (i.e., similar to Eq. 2, but in vector form), and where $\sigma^2_{\varepsilon^{site}}$ is the variance of the intra-event residual at a case history site. The components of the covariance matrix were computed from:

$$\boldsymbol{\Sigma}(i, j) = \rho_{i,j}\, \sigma_{\varepsilon i}\sigma_{\varepsilon j} \quad (4)$$

where $\rho_{i,j}$ is the spatial correlation of intra-event residuals between locations *i* and *j*, and $\sigma_{\varepsilon i}$ and $\sigma_{\varepsilon j}$ are the standard deviations of the intra-event residual at those respective locations. This spatial correlation was computed using the model of Goda and Atkinson (2009), which is specific to subduction zones. Based on the joint distribution of intra-event residuals (Eq. 3), the conditional distribution of ε^{site} was computed (e.g., Johnson and Wichern, 2007):

$$\left[\varepsilon^{site}|\boldsymbol{\varepsilon^{SMstation}}\right] = N(\boldsymbol{\Sigma_{12}} \cdot \boldsymbol{\Sigma_{22}^{-1}} \cdot \boldsymbol{\varepsilon^{SMstation}}, \sigma^2_{\varepsilon^{site}} - \boldsymbol{\Sigma_{12}} \cdot \boldsymbol{\Sigma_{22}^{-1}} \cdot \boldsymbol{\Sigma_{21}})$$

$$= N(\mu_{\varepsilon^{site} | \varepsilon^{SMstation}}, \sigma^2_{\varepsilon^{site} | \varepsilon^{SMstation}}) \quad (5)$$

Using the conditional distribution of ε^{site} and substituting into Eq. 2, the conditional distribution of PGA at a case history site, PGA_{site}, is:

$$[ln\,PGA_{site} \mid ln\,PGA_{SMstation}] = N(\mu_{\ln PGA_i} + \eta + \mu_{\varepsilon^{site}| \varepsilon^{SMstation}}, \sigma^2_{\varepsilon^{site}| \varepsilon^{SMstation}}) \quad (6)$$

This conditional distribution is lognormal and defined by the conditional median and conditional uncertainty (i.e., lognormal standard deviation). When a case history is located far from an SMS, the conditional and unconditional distributions are more similar (i.e., the *PGA* distribution assigned to the site is similar to that predicted by the GMM). Conversely, when a case history is very near to an SMS, the conditional distribution of *PGA* approaches the recorded value. The conditional median and conditional uncertainty of *PGA* are provided in the dataset for each case history and are mapped for the Seattle area in Figures 2 and 3, respectively, where 17 of 24 case histories are located. The complete PGA distribution may be generated by:

$$PGA_x = PGA_{50} * \exp\,(n\sigma_{lnPGA}) \quad (7)$$

where PGA_x is the value of *PGA* for the x^{th} percentile of the conditional distribution, PGA_{50} is the conditional median, σ_{lnPGA} is the conditional lognormal standard deviation, and *n* is the "z-value" of the standard normal distribution for the x^{th} percentile. The 16^{th} and 84^{th} percentiles of the distribution, for example, are computed with *n* values of -1 and 1, respectively. While many prior compilations and analyses of liquefaction case histories (e.g., those training probabilistic triggering models) have assigned uncertainties to *PGA*s in a largely subjective manner, the adopted approach computes uncertainty explicitly and objectively. These advantages also apply to the estimates of median *PGA*s, which for other case histories have often been computed with less rigor and more subjectivity (e.g., assuming that the *PGA* at a site of interest is equal to that at the nearest recording station, despite potential differences in path and site effects).

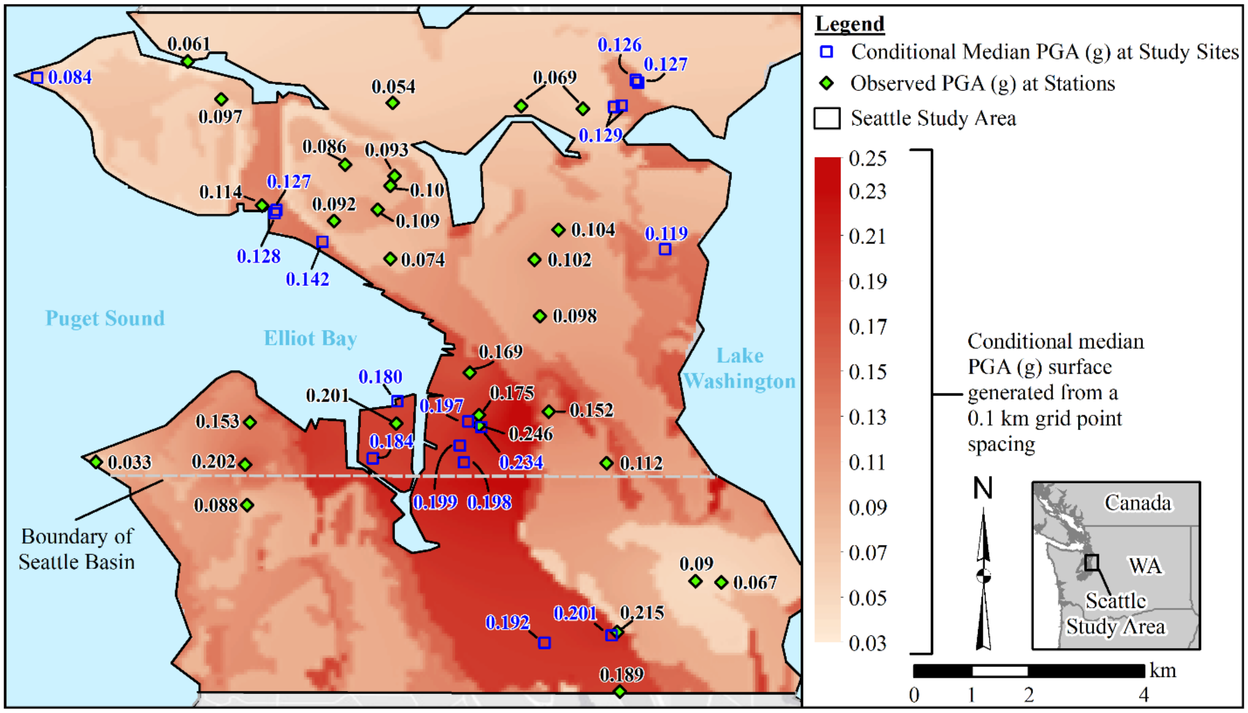


Figure 2. Recorded and computed conditional median *PGA*s in the Seattle area.

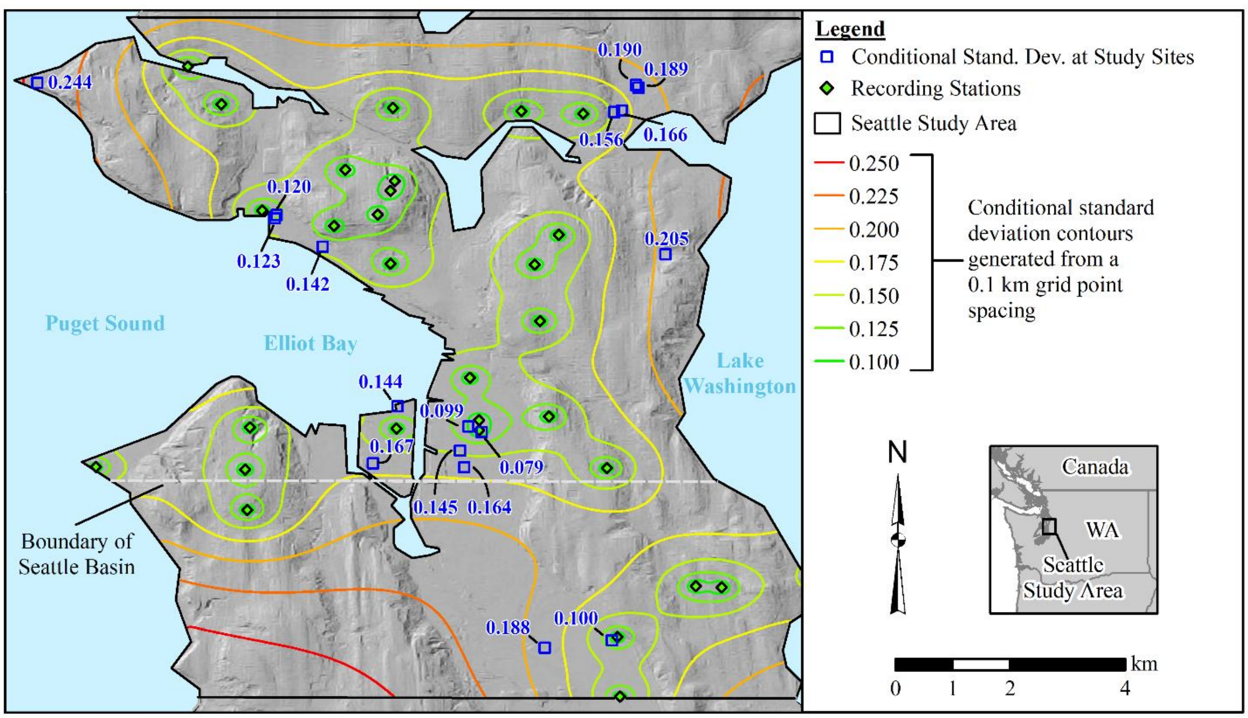


Figure 3. Computed conditional lognormal standard deviations of *PGA* in the Seattle area.

Data Structure

Provided in Table S1 of the appendix is a comprehensive summary of the case-history metadata. In addition to the data given in Table 1, Table S1 provides: citations for the geotechnical testing and post-earthquake reconnaissance; GMM inputs for each case-history site (e.g., V_{S30}, $Z_{1.0}$, $Z_{2.5}$, basin flags); notes and commentary specific to each case history; and the CPT pre-drill depth and test date. The CPT data for each case history is also provided in the appendix in 24 separate files (.xlsx format), where the most pertinent metadata specific to each case is again given (e.g., G*WT* and pre-drill depths, median and sigma of conditional *PGA*, etc.). The structure of each of these 24 files, which are named after our site identifiers (e.g., "Bush School-01.xlsx), is illustrated in Table 2 for case history No. 16. The dataset can also be accessed via this paper's appendix or the DesignSafe data depot at Rasanen et al. (2022) (https://doi.org/10.17603/ds2-nsf8-7944).

Table 2. Case-history data structure (shown is the data for case history 16: Bush School-01).

Depth (m)	**q (MPa)**	**f_s (MPa)**	**u_2 (MPa)**	**Site Name**	Bush School-01
0	2.208	0.090	0.000	**Latitude (WGS84)**	47.6228442
0.01	2.208	0.090	0.000	**Longitude (WGS84)**	-122.289624
0.02	2.208	0.090	0.000	**Geotechnical Reference**	ZZA, Inc. (2002)
0.03	2.208	0.090	0.000	**Test Date**	6/11/2002
0.04	2.208	0.090	0.000	**Manifestation Classification**	No
0.05	2.208	0.090	0.005	**Water Table Depth (m)**	3.05
⋮	⋮	⋮	⋮	**Pre-drill (m)**	0.61
				Conditional Median PGA (g)	0.119
				Conditional Standard Deviation	0.205

Limitations and Uncertainties

The methods used to compile Nisqually case histories are consistent with past precedent and, in some respects, are more rigorous (e.g., the measurement and estimation of ground motions). Nonetheless, and like existing case histories, there are limitations and uncertainties that analysts should be aware of. First, it is generally acknowledged that soil shear strength, as measured by penetration resistance, slowly increases with time (e.g., Kulhaway and Mayne, 1990), meaning that the time of in-situ testing relative to the time of the earthquake may be relevant. Despite this possibility, there has been no standard or best practice for the relative timing of in-situ testing when compiling case-history data. Inherent to existing compilations of global case histories (e.g., Boulanger and Idriss, 2016), in-situ testing has been performed both well in advance and well after the earthquake of interest, often with multiple other earthquakes occurring in the meantime. This issue is discussed in detail by Geyin et al. (2021). In the dataset compiled herein, CPTs were performed in advance of the Nisqually earthquake at 22 sites and approximately two years after the earthquake at the remaining two sites. Unlike some prior compilations in the literature, we include the date of CPT testing, should analysts wish to consider it.

Second, while the depth of CPT penetration is assumed to be the limiting depth of potentially liquefiable soil, the possibility of premature refusal (i.e., termination on shallow gravels that overlay susceptible materials) is difficult to refute. This possibility was mitigated by making use of the plentiful borings in the affected area, which were used to identify and eliminate provisional case histories with a higher likelihood of premature refusal. One exception, for which some degree of additional uncertainty exists, is case #4 (Harbor Island-01). From several adjacent borings, the CPT at this site is interpreted to have terminated on gravel near a depth of 8 m, which in the adjacent borings is underlain by sandy gravels and gravelly sands of reported medium density. Hence there is some possibility of additional, potentially liquefiable soil, albeit the SPT data indicate that the materials least resistant to liquefaction reside within the extent of the CPT.

Third, an estimate of the *GWT* depth is needed both at the time of in-situ testing and at the time of the earthquake. The former is required for stress-normalization of CPT data (i.e., routine processing), while the latter is used to infer saturation and compute the imposed cyclic stress. In this regard, the Nisqually earthquake impacted a region characterized by relatively wet winters and relatively dry summers, which could give rise to temporally fluctuating *GWT* depths. In the present effort, and like most prior case histories, the *GWT* at the time of the earthquake can only be estimated from measurements at other times. Of the 24 CPTs, 18 were performed during what we interpret to be the wet season, consistent with the Nisqually earthquake's occurrence in February. Analyses of the remaining six, which were performed during what we interpret to be drier months, give no indication of bias (i.e., the computed liquefaction hazards, as subsequently computed and presented herein, do not appear systematically suppressed, which could occur if the *GWT* depths at the time of the earthquake were overestimated). Nonetheless, because one *GWT* depth is used to represent conditions at the times of both testing and shaking, this depth should be viewed as an uncertain quantity.

Fourth, *PGA* was estimated at study sites by conditioning GMM predictions on ground-motion recordings. Inherent to this approach, and to the fundamental way in which "stress-based" liquefaction models are trained, this *PGA* should be that which occurs in the absence of liquefaction. Notably, four strong-motion instruments (stations SDS, SDW, HAR, and BOE) – all in the Duwamish River valley immediately south of downtown Seattle – were in areas where liquefaction was observed during the Nisqually earthquake (e.g., Frankel et al., 2002). It is well established that liquefaction can influence ground-motion records, typically in the form of high-frequency acceleration pulses due to the soil's cyclic mobility/dilation response, and a subsequent reduction in high frequency motion due to softening from liquefaction (e.g., Kramer et al., 2016). Each of these effects could influence a recorded *PGA* (i.e., augment or suppress it), and in turn, influence the conditional *PGA*s computed at case-history sites, especially if proximal to an affected instrument. Of the four stations above, Zhan and Chen (2021) reported little or no evidence of liquefaction in the records from stations SDS and HAR, which may be attributable to the generally minor liquefaction observed in these areas (i.e., SODO district and Harbor Island), or otherwise indicate that soil directly beneath the instruments did not liquefy. As noted by Zhan and Chen (2021), however, the records from stations BOE and SDW do show possible evidence of liquefaction, especially for the latter. If the *PGA*s recorded at these stations are associated with a high frequency dilation pulse, this *PGA* could exceed that which would have occurred in the absence of liquefaction. It has been suggested in such cases (e.g., Upadhyaya et al., 2019) that taking the *PGA* prior to any inferred evidence of liquefaction may be more appropriate. However, whereas a dilation pulse could artificially increase the *PGA*, selecting a peak value prior to evidence of liquefaction may artificially reduce it. Thus, the "true" *PGA* – sans liquefaction – cannot be known with confidence. For this study, station record SDW was removed from the analyses, given its proximity to case histories in Seattle's SODO district and the interpreted strong evidence of liquefaction. Conversely, the record from BOE was retained despite this uncertainty. While case history #9 (Boeing Field-01) shares the namesake of the potentially dubious BOE station, it is located more than 1 km from BOE and much nearer to a station without inferred evidence of liquefaction. Additionally, the recorded *PGA* at BOE was not a spatial outlier (i.e., not unexpectedly high or low). Thus, while the possibility of unwanted influence persists in the dataset, this influence is judged to be quite minimal, given the large number of recording stations in the region and fact that very few are believed to have been impacted by liquefaction. Nonetheless, analysts should be aware of this issue.

Fifth, and following from above, conditional *PGA*s were estimated at study sites using several site and region-specific inputs and components, such as the Goda and Atkinson (2009) spatial correlation model. While this model was developed specifically for subduction zone settings, it may be possible to instead develop and use an event-specific correlation model, given the large number of records, albeit most case histories compiled herein are relatively close to recording stations, so *PGA* estimates would likely be very similar. Similarly, selecting other GMMs, basin maps, V_{S30} values, etc. would invariably change the *PGA* estimates at study sites.

Performance Assessment of Liquefaction Response Models

Evaluation Methodology

Six CPT-based liquefaction triggering models will be tested using the newly compiled data. These include Robertson and Wride (1998) [RW98], Architectural Institute of Japan (2001) [AIJ01], Moss et al. (2006) [Mea06], Idriss and Boulanger (2008) [IB08], Boulanger and Idriss (2016) [BI16], and Green et al. (2019) [Gea19]. However, because these models predict the incidence of liquefaction as a function of depth within a profile, whereas the field observations of liquefaction response were made at the ground surface, the triggering models cannot be directly evaluated. This discrepancy is not unique to the Nisqually earthquake, but rather, is characteristic of essentially all existing liquefaction case histories, given the extreme difficulty and expense of determining which strata, at which depths, did or did not liquefy. While past studies have evaluated triggering models by selecting one "critical layer" from the profile, this selection is highly subjective and invites confirmation bias. Different analysts will invariably select different critical strata and different representative properties for those strata, meaning that nearly any *a priori* assumption of model superiority may be corroborated. Accordingly, to objectively compare predictions to field observations, each of the six triggering models will be used in series with three separate models that predict surficial manifestations of liquefaction: (i) the liquefaction potential index, or *LPI*, proposed by Iwasaki et al. (1978); (ii) a modified *LPI*, termed LPI_{ISH}, proposed by Maurer et al. (2015b), and (iii) the liquefaction severity number, or *LSN*, of van Ballegooy et al. (2014). For brevity, and because *LPI*, LPI_{ISH}, and *LSN* are well known in the literature, their definitions are omitted here but are exactly as described in Geyin et al. (2020). For consistency with how *LPI*,

LPI_{ISH}, and *LSN* are defined and routinely computed in practice, liquefaction triggering was computed using the deterministic version of each triggering model. Given the six triggering models and three manifestation models, a total of 18 prediction models will be evaluated.

Of course, this approach to model evaluation also calls attention to the paradoxical way in which triggering models have been developed and used. Specifically, and as discussed by Geyin et al. (2020) and Upadhyaya et al. (2022), the development of a triggering model inherently requires use of a manifestation model, given that surficial evidence presents an inverse problem. Liquefaction could conceivably occur in different strata, at different depths, and to differing degrees (or not at all) but produce a manifestation that is perceptibly the same. The developers of triggering models, however, have to-date used manifestation models defined by personal and generally unknown judgements, rather than by analytical expressions (e.g., *LPI*). This makes it impossible to evaluate the triggering models in a manner consistent with their respective developments, given that the manifestation model used to develop a triggering model, and that subsequently used with the model to make forward predictions, are not the same. Our evaluation is thus acknowledged to be less than completely rational but completely consistent with the current state-of-practice.

To make predictions of surface manifestation, and to facilitate comparison with other global case histories, the "global" fragility functions of Geyin and Maurer (2020), which are conditioned on *LPI*, LPI_{ISH}, and *LSN*, will be used to compute probabilities of surface manifestation for each of the 24 Nisqually case histories. These functions were previously trained on a large compilation of existing global case histories and are specific to each of the 18 models evaluated herein (i.e., the function coefficients are specific to the adopted triggering and manifestation model). Prior to using any of the six triggering models, liquefaction susceptibility was inferred from the CPT soil behavior type index (I_c) (Robertson and Wride, 1998). However, because the relationship between I_c and susceptibility is uncertain, the I_c – susceptibility model of Maurer et al. (2019) was used to probabilistically predict susceptibility, as defined by Boulanger and Idriss (2006), from the measured I_c. While this model was trained on data from New Zealand, it is the only known model of its type and its applicability elsewhere has not been disproven. Ultimately, our accounting for the uncertain relationship between I_c and susceptibility is generally inconsequential. Nonetheless, to include this uncertainty, as well as that of the estimated *PGA* for each case history, the probability of manifestation is computed as:

$$P(\text{Manifestation}) = \int_{PGA} \int_{I_c} P(\text{Manifestation}|PGA, I_c\ threshold)\, f(I_c) f(PGA) \cdot \mathrm{d}I_c \cdot \mathrm{d}PGA \quad (8)$$

Where $f(I_c\, threshold)$ is the probability density function (PDF) of the I_c threshold for discriminating susceptibility (Maurer et al., 2019); $f(PGA)$ is the PDF of conditional PGA, as computed herein for each case history (e.g., Eq. 7); and $P(\text{Manifestation}|PGA, I_c\ threshold)$ is the probability of surficial manifestation, which is conditioned on *PGA* and the threshold I_c for determining susceptibility, among numerous other inputs, and which is computed via the fragility functions of Geyin and Maurer (2020):

$$F_M(LMM) = \Phi\left(\frac{ln(LMM) - \ln(\theta)}{\beta}\right) \quad (9)$$

Where $F_M(LMM)$ is the probability of surface manifestation conditioned on a liquefaction manifestation model (*LMM*) index value; Φ is the Gaussian cumulative distribution function, and θ and β are the distribution's median and logarithmic standard deviation, respectively, and are obtained from Table 4 of Geyin and Maurer (2020) for each of the 18 models.

Model efficiency will be evaluated using receiver operating characteristic (ROC) analyses, which are ubiquitous in science, engineering, medicine, etc. (e.g., Fawcett, 2006) and widely used in earthquake engineering. Specifically, the area under the ROC curve (*AUC*) will be used to quantify performance. *AUC* is a particularly attractive measure of efficiency when working with class-imbalanced data, as is common for liquefaction datasets. Whereas measures like overall accuracy are sensitive to this sampling bias and may thus give the impression that biased models are better than unbiased models, *AUC* is not (Fawcett, 2006). In the context of this study, *AUC* is the probability that sites with manifestations have higher computed probabilities of manifestation than sites without manifestations. Models with higher *AUC* are thus more efficient. However, to evaluate whether differences in *AUC* could happen by chance (i.e., due to the finite sampling of case histories), tests of statistical significance will be performed using the technique of DeLong et al. (1988), which is specific to ROC analyses. Each model will be compared against all others to determine which, if any, is significantly better. Finally, the optimum operating point, or *OOP*, will be computed from ROC analyses of each model. In the context of this study, the *OOP* is the computed probability of manifestation that most optimally separates cases with and without observed manifestations. An *OOP* much less than 50% would suggest that liquefaction manifestations occurred more readily than expected in the Nisqually dataset, as compared to the

Geyin and Maurer (2020) analyses of other global case histories. In other words, this would suggest that the liquefaction hazard is underpredicted in the Nisqually event based on the calibrated median behavior of the model in prior earthquakes. Conversely, an *OOP* greater than 50% would indicate that manifestations occurred less readily than expected and that the hazard is to some degree overpredicted by the Geyin and Maurer (2020) fragility functions. Similar to the interpretation of *AUC*, confidence intervals will be computed for each model's *OOP* using bootstrap sampling, thereby accounting for the finite-sample uncertainty of *OOP* values. From these confidence intervals, it can be determined whether any deviations of the *OOP*s from 50% are statistically significant.

Model Evaluation Results and Discussion

Using the data and methodology above, 18 models were used to predict probabilities of surface manifestation for 24 case histories. These predicted probabilities are presented in Table 3, where the prediction averaged across the 18 models is also shown for each case history. Considering all cases, sites with observed manifestations have average probabilities ranging from 0.12 to 0.79 whereas sites without observed manifestations have average probabilities ranging from 0.00 to 0.40. It may be observed that the predictions made by the 18 models often vary considerably for a given case history. The difference between the highest and lowest predicted probability, for example, exceeds 0.50 for 6 cases and exceeds 0.30 for 12 cases. These large differences indicate that the cases may be useful for resolving discrepancies amongst the models (i.e., that the soil is near the modeled threshold for triggering over several meters of the subsurface). To facilitate case-history selection and interpretation for various uses, profiles of the cone-tip resistance normalized for overburden and corrected for fines content (q_{c1ncs}), cyclic stress ratio (*CSR*), I_c, and factor of safety against liquefaction triggering (*FoS*) are provided in the appendix for each case history. These figures utilize the Boulanger and Idriss (2016) model to compute q_{c1ncs}, *CSR*, and *FoS* and an example is shown in Figure 4 for case-history no. 12.

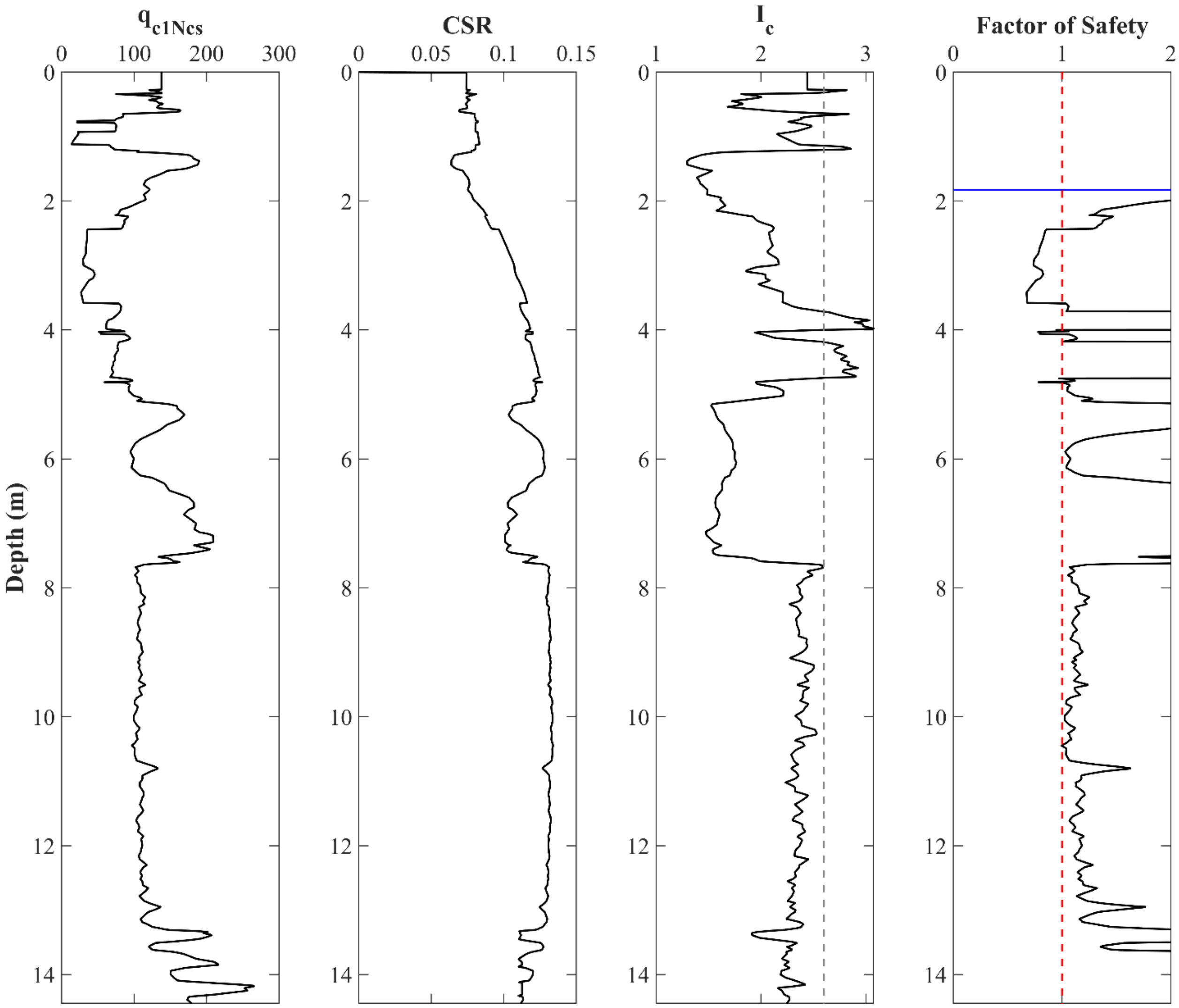


Figure 4. Computed q_{c1ncs}, *CSR*, I_c, and *FoS* for case history no. 12 (Elliot Ave-01), as described in the text; analogous figures are provided in the appendix for all 24 case histories.

Several observations can be made from the plots like that in Figure 4. First, and as suggested by the large variance across model predictions, many profiles contain soils very near the expected threshold for liquefaction. The case history in Figure 4 is one such example, in that ~6 m of the profile has computed *FoS* between 0.8 and 1.2. Second, all "no manifestation" case-history sites are inferred to contain deposits susceptible to liquefaction (e.g., soils with $I_c < 2.6$) that would be expected to trigger under sufficient seismic loading; these inferences were also confirmed by proximal borings in each case. Third, it can also be seen that many "no manifestation" cases contain deposits expected to liquefy by one or more triggering model. The case in Figure 4 is one of several examples where triggering is predicted but no manifestation is observed. These cases may help to resolve discrepant predictions amongst triggering models and/or further elucidate the complex relationship between the triggering and surficial manifestation of liquefaction.

To quantify model efficiency, ROC analyses were performed and are summarized in terms of *AUC* in Table 4; additionally, ROC plots are provided in the appendix for each of the 18 models. The top-performing models include BI16-*LPI* (i.e., the BI16 triggering model used with the *LPI* manifestation model) and RW98-*LPI*, which each have an *AUC* of 0.926. For context, an *AUC* of 0.5 corresponds to random guessing (i.e., there is a 50% probability that sites with manifestations have higher model predictions than sites without manifestations). In the case of a perfectly efficient model, this probability is 100% (i.e., *AUC* is 1). Nearly as efficient as these models are $BI16\text{-}LPI_{ISH}$ and Gea-*LPI* (*AUC* = 0.919). The least efficient model for this dataset is Mea06-*LSN* (*AUC* = 0.793), albeit its performance is still nearer to a perfect model than to random guessing. Notably, the computed *AUC* values have relatively large finite-sample uncertainties, given the limited number of compiled case histories. If these 24 cases were hypothetically selected from a much larger population of case histories, then different samples of 24 cases would likely result in different *AUC* values, and in turn, different assessments of model superiority. In this way, *AUC* values are more uncertain when based on a small sample of the population, all else being equal, and there is less confidence that any one sample (e.g., an observed difference in model performance) accurately reflects the overall population. This uncertainty also increases further as the sample becomes more heterogeneous, indicating that a larger sample is required to accurately represent the population at large.

Table 3. Probabilities of surficial liquefaction manifestations (i.e., “ground failure”), as predicted by 18 models for 24 case histories.

Case History No.	Surface Manifestation	RW98			AIJ01			Mea06			IB08			BI16			Gea19			Average
		LPI	LPI_{ISH}	*LSN*	*LPI*	LPI_{ISH}	*LSN*	*LPI*	LPI_{ISH}	*LSN*	*LPI*	LPI_{ISH}	*LSN*	*LPI*	LPI_{ISH}	*LSN*	*LPI*	LPI_{ISH}	*LSN*	
1	Yes	0.75	0.82	0.90	0.68	0.74	0.65	0.80	0.86	0.88	0.73	0.79	0.89	0.73	0.78	0.89	0.74	0.77	0.90	0.79
2	Yes	0.35	0.31	0.46	0.42	0.34	0.31	0.37	0.35	0.46	0.34	0.32	0.45	0.46	0.45	0.50	0.45	0.46	0.50	0.41
3	Yes	0.08	0.14	0.12	0.05	0.07	0.06	0.14	0.19	0.20	0.13	0.19	0.18	0.13	0.22	0.18	0.14	0.29	0.19	0.15
4	Yes	0.06	0.08	0.19	0.04	0.04	0.11	0.05	0.09	0.18	0.08	0.20	0.18	0.05	0.17	0.18	0.06	0.25	0.18	0.12
5	Yes	0.34	0.22	0.68	0.52	0.36	0.65	0.38	0.30	0.77	0.30	0.29	0.65	0.43	0.35	0.76	0.44	0.39	0.75	0.48
6	Yes	0.11	0.03	0.32	0.25	0.09	0.33	0.42	0.31	0.62	0.12	0.06	0.32	0.22	0.08	0.43	0.22	0.08	0.45	0.25
7	Yes	0.09	0.03	0.23	0.03	0.01	0.10	0.41	0.34	0.59	0.17	0.10	0.30	0.15	0.07	0.28	0.15	0.07	0.29	0.19
8	Yes	0.50	0.34	0.63	0.59	0.51	0.51	0.61	0.56	0.69	0.38	0.32	0.58	0.57	0.53	0.74	0.60	0.59	0.76	0.56
9	Yes	0.25	0.22	0.45	0.28	0.22	0.41	0.44	0.45	0.60	0.28	0.27	0.43	0.32	0.31	0.51	0.32	0.35	0.53	0.37
10	No	0.05	0.02	0.15	0.04	0.01	0.11	0.21	0.20	0.38	0.11	0.08	0.23	0.12	0.07	0.25	0.11	0.06	0.25	0.14
11	No	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	No	0.16	0.16	0.64	0.17	0.18	0.44	0.37	0.40	0.73	0.28	0.39	0.71	0.21	0.30	0.70	0.25	0.42	0.72	0.40
13	No	0.05	0.03	0.24	0.04	0.00	0.18	0.17	0.14	0.47	0.08	0.08	0.27	0.04	0.05	0.25	0.04	0.11	0.27	0.14
14	No	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.05	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00
15	No	0.07	0.11	0.17	0.05	0.05	0.07	0.17	0.29	0.44	0.08	0.15	0.18	0.10	0.17	0.19	0.11	0.22	0.23	0.16
16	No	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00
17	No	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	No	0.08	0.04	0.20	0.05	0.01	0.08	0.29	0.33	0.55	0.14	0.12	0.28	0.10	0.06	0.23	0.11	0.11	0.27	0.17
19	No	0.05	0.12	0.17	0.01	0.03	0.07	0.19	0.22	0.40	0.11	0.22	0.27	0.07	0.19	0.24	0.08	0.30	0.26	0.17
20	No	0.01	0.00	0.02	0.00	0.00	0.00	0.08	0.05	0.25	0.03	0.01	0.08	0.01	0.00	0.04	0.01	0.00	0.05	0.04
21	No	0.00	0.00	0.03	0.00	0.00	0.01	0.07	0.07	0.24	0.02	0.01	0.08	0.01	0.00	0.04	0.01	0.00	0.05	0.04
22	No	0.16	0.10	0.43	0.20	0.13	0.35	0.29	0.22	0.61	0.20	0.18	0.48	0.18	0.16	0.46	0.19	0.21	0.50	0.28
23	No	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.04	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	No	0.01	0.00	0.02	0.01	0.00	0.01	0.04	0.01	0.11	0.03	0.01	0.05	0.03	0.01	0.04	0.02	0.00	0.04	0.02

Table 4. Summary of liquefaction model performance, as quantified by *AUC* (the area under a receiver operating characteristic (ROC) curve).

Model	**RW98**	**AIJ01**	**Mea06**	**IB08**	**BI16**	**Gea19**
LPI	0.926	0.896	0.874	0.889	0.926	0.919
LPI_{ISH}	0.881	0.911	0.859	0.852	0.919	0.867
LSN	0.859	0.859	0.793	0.830	0.852	0.852

Accordingly, to evaluate whether the differences in *AUC* reported in Table 4 are statistically significant, p-values were computed per DeLong et al. (1988) to compare each model's *AUC* to all others. These values are given in Table 5 and are the probabilities that *AUC* values from two models came from the same distribution. That is, the probability that an observed difference in model performance is due to finite sampling (as opposed to real differences in performance). We adopt the significance level of 0.1, such that p-values below this are deemed statistically significant (i.e., there is at least 90% confidence the models are different). Smaller p-values are generally expected when: (i) the dataset is large and relatively homogeneous; and (ii) the prediction models have large observed differences in performance, which is to say the predictions are not strongly correlated. Based on this threshold, Table 5 compares all model pairs and indicates whether either model is better – to a statistically significant degree – via the cell shading. If the cell is shaded grey, the model in the top row is better, whereas if the cell is shaded orange, the model in the left column is better. If the cell is unshaded, the models are not significantly different. The most salient findings from Table 5 are as follows. First, no model is significantly better than all others. Second, most of the measured differences in model performance (i.e., Table 4) are not statistically significant. Of the 153 model comparisons made in Table 5, just 13 are significant. While this number would increase if the 0.1 threshold were relaxed (e.g., to 29 using the unusual threshold of 0.2), the overall conclusion is that little can be confidently said about model superiority. Third, the models with significantly different performance are largely models based on *LSN*, which when compared to the best-performing models (e.g., BI16-*LPI*), are significantly less efficient.

Table 5. Summary of p-values computed to compare the performance of each model, as quantified via *AUC*, to all others.

p-value matrix[1]		***LPI***						***LPI_{ISH}***						***LSN***					
		RW98	**AIJ01**	**Mea06**	**IB08**	**BI16**	**Gea19**	**RW98**	**AIJ01**	**Mea06**	**IB08**	**BI16**	**Gea19**	**RW98**	**AIJ01**	**Mea06**	**IB08**	**BI16**	**Gea19**
LPI	**RW98**		0.509	0.432	0.122	1.000	0.734	0.339	0.573	0.193	0.234	0.856	0.311	0.151	0.174	**0.054**	**0.044**	0.114	0.114
	AIJ01			0.746	0.880	0.409	0.564	0.802	0.705	0.558	0.582	0.709	0.673	0.545	0.539	0.174	0.310	0.446	0.446
	Mea06				0.802	0.302	0.400	0.935	0.642	0.724	0.833	0.616	0.942	0.814	0.833	**0.056**	0.448	0.577	0.577
	IB08					0.227	0.207	0.880	0.580	0.448	0.557	0.562	0.718	0.521	0.580	0.117	0.138	0.328	0.328
	BI16						0.480	0.446	0.707	0.150	0.338	0.889	0.409	0.223	0.252	**0.036**	**0.054**	**0.094**	**0.094**
	Gea19							0.483	0.840	0.168	0.338	1.000	0.432	0.244	0.286	**0.045**	**0.050**	0.104	0.104
LPI_{ISH}	**RW98**								0.430	0.757	0.462	0.161	0.625	0.751	0.754	0.355	0.496	0.675	0.675
	AIJ01									0.418	0.283	0.810	0.337	0.322	0.300	0.142	0.178	0.305	0.305
	Mea06										0.924	0.404	0.924	1.000	1.000	0.184	0.491	0.810	0.810
	IB08											**0.075**	0.530	0.914	0.917	0.562	0.774	1.000	1.000
	BI16												**0.087**	0.358	0.351	0.180	0.224	0.333	0.333
	Gea19													0.912	0.913	0.466	0.641	0.846	0.846
LSN	**RW98**														1.000	0.176	0.385	0.811	0.811
	AIJ01															0.235	0.482	0.852	0.852
	Mea06																0.364	**0.094**	**0.094**
	IB08																	0.414	0.414
	BI16																		1.000
	Gea19																		

[1]Cell values are the probabilities that *AUC* samples for two models could have come from the same distribution. Values less than 0.1 are deemed "significant," in which case the model with significantly better performance is indicated via the cell shading.

Each of these findings is generally consistent with analogous analyses previously performed for global CPT-based case histories (Geyin et al., 2020). This calls attention to the general inappropriateness of championing prediction models based on observations in individual earthquakes, and to the ongoing needs for additional case-history data and more rigorous adherence to best practices in model training and testing. With few exceptions (e.g., the Canterbury earthquake sequence of the 2010's), case histories compiled from individual earthquakes – or even from many earthquakes – are insufficient to establish statistically significant differences in model performance. Yet numerous papers have trained, calibrated, or tested liquefaction response models using small datasets, with many deriving claims of model superiority. Tests of statistical significance are exceedingly rare in the liquefaction literature, however, and if presented, would likely change the perception of prominent works. Additionally, published models are in many instances never tested against existing models. None of the six liquefaction triggering models evaluated herein, for example, was tested against any other model when originally proposed. The same is true for many models that predict manifestations of liquefaction. While such practices would be objectionable in many fields, they have passed as acceptable in geotechnical engineering.

Finally, the *OOP*, or computed probability of manifestation that most optimally separates cases with and without observed manifestations, was extracted from the ROC analyses. An *OOP* of 0.19 was computed for the probability of manifestation averaged across all 18 models, suggesting that liquefaction manifestations may have occurred more readily than expected by the Geyin and Maurer (2020) fragility functions. That is, the liquefaction hazard may have been underpredicted based on calibrations of the 18 models against global case histories. If so, this apparent bias could have several causes. First, the observed manifestations of liquefaction were largely "minor" except for cases #9 (Boeing Field-01) and #1 (Capital Lake-01), where manifestations could be classified as more severe per the criteria of Maurer et al. (2014a). If the criteria for selecting "positive" case histories were historically more stringent (i.e., requiring more severe manifestations to draw attention and warrant documentation), this might impart a tendency in the Geyin and Maurer (2020) fragility functions to underpredict minor manifestations, being that the functions are based on global and generally older case histories. This is plausible, given that relatively less severe liquefaction might be overlooked if not in a major city such as Seattle. As an example, case #6 (SODO-02) lacked ejecta but was interpreted to be a site of liquefaction based on cracking and differential settlements observed in pavement. In this case and others, where the observed manifestations were relatively minor, it might be expected that the computed liquefaction hazards would be on the lower end of the spectrum for "positive" sites. It has similarly been observed that lateral spreading may occur at low *LSN*, *LPI*, and LPI_{ISH} values (e.g., Maurer et al., 2015c), given that it can result from a thin liquefied stratum that would not otherwise produce ejecta or large settlements. Case #3 (Capitol Lake-02), which had a low computed probability of manifestation, involved a possible spreading crack and a trace amount of ejecta. In the absence of the crack, however, it might be conjectured that liquefaction would not have manifested.

Second, 17 of the 24 case histories map to artificial fill. While this is not unusual in the context of global case histories, it nonetheless may explain the greater-than-expected liquefaction response, which could be expected to increase with decreasing deposit age, all else being equal. CPT penetration resistance does fully capture small-strain soil aging effects (e.g., Maurer et al., 2014b) and correction factors to account for this shortcoming in young deposits (e.g., Bwambale and Andrus, 2019) were not used. Third, many of the case histories were compiled from the Seattle basin (as evidenced by large $Z_{1.0}$ and $Z_{2.5}$ values – see Table S1), which could result in more cycles of seismic loading than expected. While liquefaction models account for the influence of near-surface conditions on cyclic demand at depth, no model explicitly considers the effects of deep sediment basins. As such, liquefaction might occur more readily where strong basin effects are observed, given that these effects can increase the duration of loading. Similarly, there are relatively few subduction zone events in the datasets used to train existing liquefaction models and model components, which may result in predictions with more uncertainty, or even bias, in subduction zone settings. This is particularly the case for large-magnitude interface ruptures, such as that which occurred in the Cascadia Subduction Zone 323 YBP and many times prior. In this regard, further research is needed to assess or improve the applicability of state-of-practice liquefaction models to subduction zone settings, which might entail the development of region-specific models (e.g., Green et al., 2020). Of course, we must add to the above various other possible causes for the apparent underprediction of liquefaction response, including other fundamental shortcomings in the models used to predict liquefaction triggering and/or its surficial manifestation.

However, like *AUC* values the finite-sample uncertainty of *OOP*s should be considered, given the limited number of case histories. To account for this uncertainty, bootstrap sampling of the 24 cases was used to compute a 95% confidence interval for the *OOP* averaged across all models. This interval ranged from 0.12 to 0.56, and as such, it cannot be confidently concluded that the *OOP* is less than 0.5 (i.e., that the Nisqually case histories deviate from median global behavior). Thus, while liquefaction manifestations may have occurred more readily than expected for the Nisqually dataset, additional case histories would be needed to confirm this, or to draw any definitive conclusion about performance and bias. The large uncertainty of the *OOP* again alludes to the dangers of training or calibrating a liquefaction model with limited data (e.g., from a single earthquake), as researchers have opted for, or been resigned to, throughout the existing literature.

Conclusions

While liquefaction is observed in nearly all moderate to large earthquakes, the data needed to train state-of-practice liquefaction prediction models has remained relatively scarce, owing to its cost and variety. This lack of data inhibits the training and testing of better models. Accordingly, this paper compiled and preliminarily analyzed 24 CPT-based case histories from the 2001 Nisqually, Washington, earthquake, which affected the Puget Sound of the U.S. Pacific Northwest. The cases were compiled from the free field on ground that is generally level, are among very few in the literature associated with subduction zone earthquakes, and collectively represent one of the largest compilations from a U.S. earthquake to date. Subsequent analyses of the data using 18 different prediction models revealed differences in model efficiency, albeit these differences were rarely statistically significant when considering finite-sample uncertainty. While these measurements used a particular metric of efficiency (*AUC*), it is our judgement that this overall result would stand independent of the metric chosen. This conclusion alludes to the general inappropriateness of championing models based on individual earthquakes or otherwise small datasets, as is frequently done in the literature without statistical backing. This also calls attention to the obvious and ongoing need for more data. The Nisqually case histories provide a useful addition for model training and testing but must be combined with existing case histories (e.g., Brandenberg et al., 2020; Montalva et al., 2021; Geyin et al., 2021) and others not yet collected.

Data Availability

The compiled dataset is available in digital format through both the DesignSafe Data Depot at https://doi.org/10.17603/ds2-nsf8-7944 and in an electronic appendix to this paper, where supplemental figures are also provided.

Funding

The presented study is based on work supported by the National Science Foundation (NSF) under Grant No. CMMI-1751216 and by the NSF Graduate Research Fellowship Program under Grant No. DGE-1762114. However, any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

Acknowledgements

The authors gratefully acknowledge the many individuals who contributed to post-earthquake reconnaissance following the Nisqually earthquake as well as Silvia Mazzoni of the John Garrick Institute for Risk Sciences at UCLA for coding the ground motion model utilized in this study.

References

AGRA Earth & Environmental, Inc. (1996) *Subsurface Exploration and Preliminary Geotechnical Engineering Report ICA Soccer and Baseball Field Development Seattle, Washington.* May 28, 1996. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

Ahdi S.K., Ancheta T.D., Contreras V., Kishida T., Kwak D.Y., Kwok O.L., Parker G.A., Stewart J.P. 2020. *Chapter 5: Site Condition Parameters.* PEER Report No. 2020/03, J.P. Stewart (ed.), Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Bassal, PC, and Boulanger, RW. 2021. System response of an interlayered deposit with spatially preferential liquefaction manifestations. *Journal of Geotechnical and Geoenvironmental Engineering*, *147*(12): 05021013.

Boulanger RW and Idriss IM (2006) Liquefaction susceptibility criteria for silts and clays. *Journal of Geotechnical and Geoenvironmental Engineering* 132(11): 1413-1426.

Boulanger RW and Idriss IM (2016) CPT-based liquefaction triggering procedure. *Journal of Geotechnical and Geoenvironmental Engineering*, *142*(2): 04015065-04015065.

Boulanger RW and Ziotopoulou K (2015) PM4Sand (Version 3): A sand plasticity model for earthquake engineering applications. *Center for Geotechnical Modeling Report No. UCD/CGM-15/01, Department of Civil and Environmental Engineering, University of California, Davis, Calif*.

Bradley BA (2014) Site-specific and spatially-distributed ground motion intensity estimation in the 2010-2011 Christchurch earthquakes. *Soil Dynamics and Earthquake Engineering* 48: 35-47.

Brandenberg SJ, Zimmaro P, Stewart JP, Kwak DY, Franke KW, Moss RE, ... and Kramer SL (2020) Next-generation liquefaction database. *Earthquake Spectra*, *36*(2): 939-959.

Bray JD, Sancio R, Kammerer AM, Merry S, Rodriguez-Marek A, Khazai B, ... and Nykamp M (2001) *Some observations of the geotechnical aspects of the February 28, 2001, Nisqually earthquake in Olympia, South Seattle, and Tacoma, Washington.* Report sponsored by NSF, PEER, UCB, University of Arizona, Washington State University, Shannon and Wilson Inc., and Leighton and Associates.

Buck JR, Daniel MM and Singer AC (2002) *Computer Explorations in Signals and Systems Using MATLAB*. 2nd ed. Upper Saddle River, NJ: Prentice Hall.

Bustin A, Hyndman RD, Lambert A, Ristau J, He J, Dragert H, and Van der Kooij M (2004) Fault parameters of the Nisqually earthquake determined from moment tensor solutions and the surface deformation from GPS and InSAR. *Bulletin of the Seismological Society of America*, *94*(2): 363-376.

Bwambale B and Andrus RD (2019) State of the art in the assessment of aging effects on soil liquefaction. *Soil Dynamics and Earthquake Engineering*, 125: 105658.

Chleborad AF and Schuster RL (1990) *Ground failure associated with the Puget Sound region earthquakes of April 13, 1949 and April 29, 1965*. United States Geological Survey Open-File Rept.

CH2M Hill Engineering Inc. (1989) *West Point Treatment Plant Secondary Treatment Facilities ST 08 Liquids Stream: Supplemental Plant Site Geotechnical Data Report.* April 1989. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

Dames and Moore (1997) *Report of Geotechnical Investigation King Street Yard Maintenance Facility Seattle, Washington.* November 12, 1997. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 14, 2022.

DeLong ER, DeLong DM and Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44: 837-845.

de Magistris, FS, Lanzano, G, Forte, G, and Fabbrocino, G (2013) A database for PGA threshold in liquefaction occurrence. *Soil Dynamics and Earthquake Engineering 54*: 17-19.

Earth Consultants Inc. (1992) *Geotechnical Engineering Study Home Depot SODO Center 1st Avenue South and South Forest Street, Seattle, Washington.* October 23, 1991. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 14, 2022.

Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874.

Frankel AD, Carver DL, & Williams RA (2002) Nonlinear and linear site response and basin effects in Seattle for the M 6.8 Nisqually, Washington, earthquake. *Bulletin of the Seismological Society of America*, *92*(6): 2090-2109.

GeoEngineers, Inc. (1990) *Geotechnical Report 1135 S. Webster Street, Seattle, Washington.* September 11, 1990. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

GeoEngineers, Inc. (1993) *Geotechnical Engineering Services Proposed Thoroughbred Racing Facility Auburn, Washington.* June 16, 1993. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 12, 2022.

GeoEngineers, Inc. (2003) *Summary Geotechnical Report Ground Improvement and Foundation Support Old Seattle Chocolate Factory Site, Seattle, Washington.* July 22, 2003. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 17, 2022.

Geyin M and Maurer BW (2020) Fragility functions for liquefaction induced ground failure. *Journal of Geotechnical and Geoenvironmental Engineering,* 146(12): 04020142.

Geyin M, Baird AJ, and Maurer BW (2020) Field assessment of liquefaction prediction models based on geotechnical vs. geospatial data, with lessons for each. *Earthquake Spectra* 36(3): 1386–1411.

Geyin M and Maurer BW (2021) Evaluation of a cone penetration test thin-layer correction procedure in the context of global liquefaction model performance. *Engineering Geology*, *291*, 106221.

Geyin M, Maurer BW, and Christofferson K (2022) An AI driven, mechanistically grounded geospatial liquefaction model for rapid response and scenario planning. *Soil Dynamics and Earthquake Engineering*, *159*, 107348.

Geyin M, Maurer BW, Bradley BA, Green RA, and van Ballegooy S (2021). CPT-based liquefaction case histories compiled from three earthquakes in Canterbury, New Zealand. *Earthquake Spectra*, *37*(4), 2920-2945.

Goda K and Atkinson GM (2009) Probabilistic Characterization of Spatially Correlated Response Spectra for Earthquakes in Japan. *Bulletin of the Seismological Society of America* 99(5): 3003-3020.

Green RA, Bommer JJ, Rodriguez-Marek A, Maurer BW, Stafford PJ, Edwards B, Kruiver PP, De Lange G and Van Elk J (2019) Addressing limitations in existing 'simplified' liquefaction triggering evaluation procedures: application to induced seismicity in the Groningen gas field. *Bulletin of Earthquake Engineering* 17(8): 4539-4557.

Green RA, Bommer JJ, Stafford PJ, Maurer BW, Kruiver PP, Edwards B, ... & van Elk J (2020) Liquefaction hazard in the Groningen region of the Netherlands due to induced seismicity. Journal of Geotechnical and Geoenvironmental Engineering, 146(8): 04020068.

Hart-Crowser & Associates, Inc. (1978) *Subsurface Exploration Report South 266th Street and South Central Avenue, Kent, Washington.* November 1978. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 14, 2022.

Hart-Crowser & Associates, Inc. (1981) *Field Exploration and Geotechnical Engineering Study Proposed J.A. Jones Graving Dock Alexander Avenue and South Lincoln Avenue on Blair Waterway Port of Tacoma, Washington.* March 26, 1981. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 14, 2022.

Hart-Crowser & Associates Inc. (1985) *Subsurface Explorations and Geotechnical Engineering Study Proposed Additions to Factoria Square Shopping Center Bellevue, Washington.* December 24, 1985. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

Hausler EA and Koelling M (2004) Performance of Improved Ground During the 2001 Nisqually Earthquake. *International Conference on Case Histories in Geotechnical Engineering*, New York, New York 18. https://scholarsmine.mst.edu/icchge/5icchge/session03/18

Ichinose GA, Thio HK, and Somerville PG (2004) Rupture process and near-source shaking of the 1965 Seattle-Tacoma and 2001 Nisqually, intraslab earthquakes. *Geophysical Research Letters*, *31*(10).

Iwasaki T, Tatsuoka F, Tokida K, and Yasuda SA (1978) Practical method for assessing soil liquefaction potential based on case studies at various sites in Japan. *2nd Int. Conf. on Microzonation*, San Francisco.

Jeschke DA, Eungard DW, Troost KG, Wisher AP (2019) Subsurface Database of Washington State GIS data. Washington Geological Survey Digital Data Series 11, version 2.1, previously released Aug 2017 http://www.dnr.wa.gov/publications/ger_portal_subsurface_database.zip.

Kao H, Wang K, Chen RY, Wada I, He J, and Malone SD (2008) Identifying the rupture plane of the 2001 Nisqually, Washington, earthquake. *Bulletin of the Seismological Society of America*, *98*(3): 1546-1558.

Kramer SL, Sideras SS and Greenfield MW (2016) The timing of liquefaction and its utility in liquefaction hazard evaluation. *Soil Dynamics and Earthquake Engineering* 91: 133-146.

Kuehn N, Bozorgnia Y, Campbell KW, and Gregor N (2020) *Partially non-ergodic ground-motion model for subduction regions using NGA-subduction database.* Peer Report 2020/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Maurer BW, Green RA, Cubrinovski M and Bradley BA (2014a) Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand. *Journal of Geotechnical and Geoenvironmental Engineering* 140(7): 04014032.

Maurer BW, Green RA, Cubrinovski M, and Bradley BA (2014b) Assessment of aging correction factors for liquefaction resistance at sites of recurrent liquefaction. In *10th national conference proceedings on earthquake engineering.* Earthquake Engineering Research Institute, Anchorage, AK.

Maurer BW, Green RA, Cubrinovski M and Bradley BA (2015a) Calibrating the liquefaction severity number (LSN) for varying misprediction economies: a case study in Christchurch, New Zealand. *6th International Conference on Earthquake Geotechnical Engineering*, Paper No. 491.

Maurer BW, Green RA and Taylor ODS (2015b) Moving towards an improved index for assessing liquefaction hazard: lessons from historical data. *Soils and Foundations* 55(4): 778-787.

Maurer BW, Green RA, Cubrinovski M, & Bradley BA (2015c) Assessment of CPT-based methods for liquefaction evaluation in a liquefaction potential index framework. *Géotechnique* 65(5): 328-336.

Maurer BW, Green RA, van Ballegooy S and Wotherspoon L (2019) Development of region-specific soil behavior type index correlations for evaluating liquefaction hazard in Christchurch, New Zealand. *Soil Dynamics and Earthquake Engineering* 117: 96-105.

McPhillips DF, Herrick JA, Ahdi S, Yong AK, and Haefner S (2020) Updated Compilation of VS30 Data for the United States. *United States Geological Survey* data release, https://doi.org/10.5066/P9H5QEAC.

Milbor-Pita, Inc. (1996) *Geotechnical Report, Heritage Park, Olympia, Washington.* January 13, 1996. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 14, 2022.

Moss RES, Seed RB, Kayen RE, Stewart JP, Der Kiureghian A and Cetin KO (2006) CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential. *Journal of Geotechnical and Geoenvironmental Engineering* 132(8): 1032-1051.

National Academies of Sciences, Engineering, and Medicine (2016) State of the art and practice in the assessment of earthquake-induced soil liquefaction and its consequences. The National Academies Press. https://doi.org/10.17226/23474.

Pacific Northwest Seismic Network (2001) Preliminary observations of ground deformation, Nisqually earthquake. Available at: https://assets.pnsn.org/notable/WEBDIR_01022818543p/maps/gd.jpg (Accessed Jan. 31, 2022).

PacRim Geotechnical, Inc. (1999) *Geotechnical Report Pier 17 Renovation Port of Seattle, Washington.* July 29, 1999. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed June 1, 2022.

Palmer SP, Magsino SL. Bilderback EL, Poelstra JL, Folger DS, Niggemann RA (2007) *Liquefaction susceptibility and site class maps of Washington State, by county.* Open File Report 2004-20, Washington Division of Geology and Earth Resources.

Parker GA, Stewart JP, Boore DM, Atkinson GM, and Hassani B (2020) *NGA-Subduction Global Ground-Motion Models with Regional Adjustment Factors.* PEER Report 2020/03, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Pinheiro J, Bates DM, DebRoy S, Sarkar D and the R Core Team (2008) NLME: Linear and nonlinear mixed effects models, R Core package, Ver. 3.1, http://www.r-project.org (accessed Mar. 26, 2014).

Rasanen R, Geyin M, Maurer B (2022) Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake. DesignSafe-CI. https://doi.org/10.17603/ds2-nsf8-7944 v1.

Rasanen RA, Marafi NA, and Maurer BW (2021) Compilation and forecasting of paleoliquefaction evidence for the strength of ground motions in the US Pacific Northwest. *Engineering Geology 292*: 106253.

Rashidian V, and Baise LG (2020) Regional efficacy of a global geospatial liquefaction model. *Engineering geology*, *272*: 105644.

Rateria G and Maurer BW (2022) Evaluation and updating of Ishihara's (1985) model for liquefaction surface expression, with insights from machine and deep learning. *Soils and Foundations 62*(3): 101131.

Rittenhouse-Zeman & Associates (1987) *Subsurface Exploration and Geotechnical Engineering Report Leiendecker Building, Seattle, Washington.* August 11, 1987. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 9, 2022.

Robertson PK and Wride CE (1998) Evaluating cyclic liquefaction potential using cone penetration test. *Canadian Geotechnical Journal* 35(3): 442-459.

RZA AGRA, Inc. (1993) *Subsurface Exploration and Geotechnical Engineering Report Proposed Costco Warehouse Facility Issaquah, Washington.* March 10, 1993. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

Seattle Office of Emergency Management (2019) *Seattle Hazard Identification and Vulnerability Analysis (SHIVA).* April 10, 2019. Available from: https://www.seattle.gov/emergency-management/hazards, last accessed May 10, 2022.

Seed HB and Idriss IM (1971) Simplified procedure for evaluating soil liquefaction potential. *Journal of the Soil Mechanics and Foundations division*, *97*(9): 1249-1273.

Shannon & Wilson, Inc. (1997) *Final Geotechnical Data Report Mercer Street Tunnel Contract Denny Way/Lake Union CSO Project.* October 1997. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

Shannon & Wilson, Inc. (1997) *Geotechnical Report Terminal 18 Project Off-Terminal Improvements Harbor Island, Seattle, Washington.* November 1997. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 26, 2022.

Shannon & Wilson, Inc. (1998) *Geotechnical Report, West Galer Street Ramp, Seattle, Washington.* June 1998. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

Sideras S (2019) *Evolutionary intensity measures for more accurate and informative evaluation of liquefaction triggering.* PhD Dissertation, University of Washington, Seattle, WA

Stephenson WJ, Reitman NG, and Angster SJ (2017) *P-and S-wave velocity models incorporating the Cascadia subduction zone for 3D earthquake ground motion simulations, Version 1.6*—Update for Open-File Report 2007–1348 (No. 2017-1152). United States Geological Survey.

Terra Associates, Inc. (2000) *Geotechnical Report Adams News Building Renovation 1555 West Galer Street Seattle, Washington.* March 10, 2000. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

Terra Associates, Inc. (2000) *Report Geotechnical Engineering Services Proposed Water System Improvements King County International Airport Seattle and Tukwila, Washington.* April 6, 2001. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed June 11, 2022.

Upadhyaya S, Green RA, Rodriguez-Marek A, Maurer BW, Wotherspoon L, Bradley BA and Cubrinovski M (2019) Influence of corrections to recorded peak ground accelerations due to liquefaction on predicted liquefaction response during the 2010-2011 Canterbury, New Zealand, earthquake sequence. In *Proc. 13th Australia New Zealand Conference on Geomechanics (13ANZCG)*.

van Ballegooy S, Malan P, Lacrosse V, Jacka ME, Cubrinovski M, Bray JD, O'Rourke TD, Crawford SA and Cowan H (2014a) Assessment of liquefaction-induced land damage for residential Christchurch. *Earthquake Spectra* 30(1): 31-55.

Walsh, TJ (2001). Working a geologic disaster. *Washington Geology* 28(3): 6-18.

Washington Division of Geology and Earth Resources (2016) Surface geology, 1:100,000--GIS data, November 2016: Digital Data Series DS-18, version 3.1, previously released June 2010

Whitman RV (1971) Resistance of soil to liquefaction and settlement. *Soils and Foundations*, 11(4): 59-68.

Wirth EA, Chang SW, and Frankel AD (2018) *2018 report on incorporating sedimentary basin response into the design of tall buildings in Seattle, Washington.* US Geol. Surv. Open-File Rept. 2018-1149.

Wong IG, Stokoe KH, Cox BR, Lin YC, and Menq FY (2011) Shear-wave velocity profiling of strong motion sites that recorded the 2001 Nisqually, Washington, earthquake. *Earthquake Spectra*, 27(1)" 183-212.

Wu Q, Li DQ, and Du W (2022) Identification of optimal ground-motion intensity measures for assessing liquefaction triggering and lateral displacement of liquefiable sloping grounds. *Earthquake Spectra*: 87552930221094344.

Zhu J, Baise LG, and Thompson EM (2017) An updated geospatial liquefaction model for global application. *Bulletin of the Seismological Society of America*, *107*(3): 1365-1385.

Zipper Zeman Associates, Inc. (2002) *Geotechnical Engineering Design Study Bush School Project Phase 1 Seattle, Washington.* June 28, 2002. Available from Washington State Department of Natural Resources Washington Geologic Information Portal, last accessed May 10, 2022.

APPENDIX H

Information describing publication 8 (Dataset 2), highlighted below, is provided on the following pages. Specifically, the dataset's posting on NHERI DesignSafe is shown below.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press*, 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.
[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

PRJ-3758 | **Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake**

Download Dataset

Cite this Data:

Rasanen, R., M. Geyin, B. Maurer. (2022) "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." DesignSafe-CI. https://doi.org/10.17603/ds2-nsf8-7944 v1

Download Citation: DataCite XML | RIS | BibTeX

Author(s)	**Rasanen, Ryan; Geyin, Mertcan; Maurer, Brett**
Data Type	**Database**
Natural Hazard Type	**Earthquake**
Date of Publication	**11-25-2022**
Awards	**CAREER: Decoding The Enigmas Of U.S. Seismic Hazard Via Multi-Scale, Multi-Physics Approaches To Paleoliquefaction Analysis | CMMI-1751216**
Keywords	**Nisqually, Earthquake, Liquefaction, Case History, Cone Penetration Test**
License(s)	**Creative Commons Attribution**

View Data Diagram | View Project Metrics | Leave Feedback

Description | While soil liquefaction is common in earthquakes, the case history data required to train and test state-of-practice prediction models remains comparatively scarce, owing to the breadth and expense of data that comprise a single case history. The 2001 Nisqually, Washington, earthquake, for example, occurred in a metropolitan region and induced damaging liquefaction in the urban cores of Seattle and Olympia, yet case history data has not previously been published. Accordingly, we compile 24 cone-penetration-test (CPT) case histories from free-field locations. The many…
Show More

PRJ-3758

Name	Size
Case History Data ***Data***	--
Paper Manuscript.pdf ***Report***	1.7 MB
Quick ReadME.pdf ***README***	82.9 kB
Summary Table S1.xlsx ***Data, Spreadsheet***	16.7 kB

APPENDIX I

Publication 9 (Journal Paper 6), highlighted below, is provided on the following pages in full (as a manuscript that is currently being prepared for submission) once completed.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Conference Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Seismic Source Parameters from Regional Paleoseismic Evidence." *ASCE Geo-Congress 2022*, 10.1061/9780784484043.040	Published
2	Journal Paper 1	2	Rasanen, R.A. and Maurer, B.W. (2022). "Probabilistic Seismic Source Location and Magnitude Via Inverse Analysis of Paleoliquefaction Evidence." *Earthquake Spectra*, 10.1177/87552930211056355, EERI.	Published[1]
3	Journal Paper 2	2	Rasanen, R.A. and Maurer, B.W. (2021). "Probabilistic Seismic Source Inversion from Regional Landslide Evidence." *Landslides* 19: 407-419 10.1007/s10346-021-01780-9, Springer.	Published[2]
4	Journal Paper 3	3	Rasanen, R.A. and Maurer, B.W. (2023). "Probabilistic Seismic Source Inversion of the 1886 Charleston, South Carolina, Earthquake from Macroseismic Evidence: A Major Updating." *Engineering Geology*, 312: 106958, 10.1016/j.enggeo.2022.106958, Elsevier.	Published
5	Journal Paper 4	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest." *Engineering Geology* 292: 106253, 10.1016/j.enggeo.2021.106253, Elsevier.	Published
6	Dataset 1	3	Rasanen, R.A., Marafi, N.A., and Maurer, B.W. (2021). "Compilation and Forecasting of Paleoliquefaction Evidence for the Strength of Ground Motions in the U.S. Pacific Northwest: A Digital Dataset (Version 2)." *DesignSafe-CI*, 10.17603/ds2-fqkr-h615	Published
7	Journal Paper 5	3	Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). "Select liquefaction case histories from the 2001 Nisqually, Washington, earthquake: A digital dataset and assessment of model performance." *Earthquake Spectra, In Press,* 10.1177/87552930231174244, EERI.	Published
8	Dataset 2	3	Rasanen, R.A., Geyin M., and Maurer, B.W. (2022). "Select CPT-Based Liquefaction Case Histories from the 2001 Nisqually, Washington, Earthquake." *DesignSafe-CI*, 10.17603/ds2-nsf8-7944 v1.	Published
9	Journal Paper 6	3	Rasanen, R.A. et al. (2023). "Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards." *In Preparation*.	In Preparation

[1]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2020 ASCE GeoCongress Poster Competition.
[2]A presentation by R. Rasanen based on this paper was awarded *2nd Place* in the 2022 ASCE GeoCongress Poster Competition.

Implications of Physics-Based M9 Ground Motions on Liquefaction-Induced Damage in the Cascadia Subduction Zone: Looking Forwards and Backwards

Rasanen et al.

Abstract: Given the premise of an M9 Cascadia Subduction Zone (CSZ) earthquake, various predictions of the resultant, regional ground motions have been made, including a suite of physics-based simulations that reflect key modeling uncertainties. However, because the last CSZ interface rupture occurred in 1700, the shaking expected in such an event is especially uncertain, as are the impacts to the built and living environments. In this regard, coseismic soil liquefaction poses a significant threat and must be considered in any scenario study used to inform planning and response, or to focus mitigation resources. Liquefaction is also notable for its potential to "ground truth," or validate, ground-motion predictions, given that its presence or absence in the geologic record can elucidate the intensities of shaking in past events. Liquefaction is thus an important phenomenon, both looking forward for damage prediction, and looking backward to constrain the characteristics of past earthquakes. Accordingly, this study: (i) predicts liquefaction in M9 CSZ ruptures using cone-penetration-test data from 400 locations in Oregon, Washington, and British Columbia; and (ii) constrains the ground motion intensities experienced in 1700, to the degree permitted, by regional paleoliquefaction evidence. Looking forward, the predictions suggest that liquefaction could be both pervasive in the region and affect numerous significant hubs of population, transportation, and commerce, with the potential for damage across hundreds of square kilometers. Looking backward, the analyses suggest that 1700 ground motion intensities may have been less in northern portions of the CSZ (e.g., Seattle and Vancouver) than prevailing forecasts indicate, given the peculiar paucity of 1700 evidence throughout the Puget Sound. This possibility has important implications, but also significant caveats and uncertainties, and is discussed in detail. Ultimately, further discovery and analysis of CSZ paleoliquefaction, or lack thereof, will confirm or modify the conclusions drawn herein.

1.0 INTRODUCTION

1.1 Significance of M9 Physics-Based Ground Motion Simulations in Cascadia

Physics-based ground motion simulations are increasingly used to supplement empirical ground motion models, particularly where seismic hazards are historically quiescent, as in the Cascadia Subduction Zone (CSZ). In addition to populating gaps in ground-motion datasets, physics-based simulations help to elucidate and quantify complex effects (e.g., those of directivity, topography, basins, and other subsurface velocity structures) via explicit modeling of fault rupture, wave propagation, nonlinear behavior, and other physical phenomena. Given the premise of a full-margin M9 CSZ earthquake, physics-based seismograms (i.e., ground-motion time histories) were recently developed (Frankel et al., 2018; Wirth et al., 2018) to predict motions across the Pacific Northwest. This suite of simulations included 30 realizations that reflect the uncertainties of key modeling parameters (e.g., the down-dip limit of fault rupture; the slip distribution and location of asperities; the rupture direction; and the hypocenter location). Given that the last CSZ interface rupture occurred in 1700 CE, the actual motions are unknown for all such events, and consequently, are quite uncertain. Shown in Figure 1 are two such realizations. While both simulate a full-margin M9 rupture, the predicted ground-motion intensities vary significantly in many locations (e.g., Seattle). Although uncertain, these motions provide the best-available predictions from which the expected coseismic impacts can be identified, quantified, prepared for, and mitigated across the region.

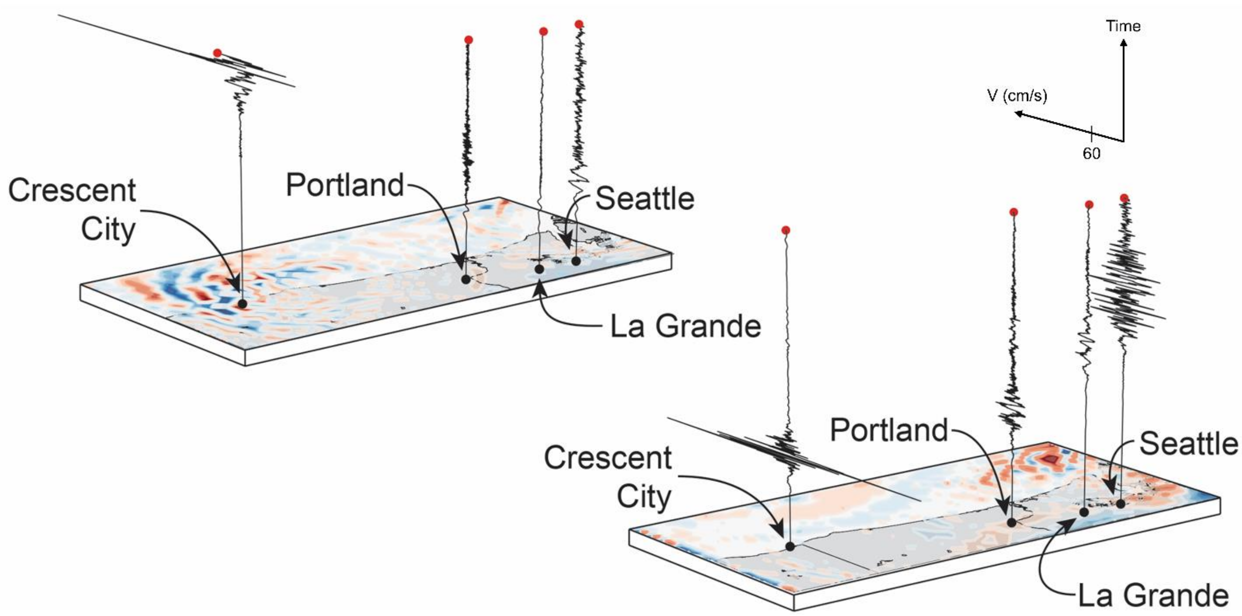


Figure 1. Results of two simulations of an M9 CSZ earthquake, reflecting the influence of salient modeling uncertainties (traces are ground velocity; colors show a snapshot of seismic waves at a moment in time). The sensitivity of predictions to modeling uncertainties is readily apparent via large discrepancies (e.g., in Seattle). Simulations by Frankel et al. (2018) and Wirth et al. (2018); Figure by Marafi, 2018.

1.2 Significance of Soil Liquefaction in Cascadia

Among the many downstream impacts worthy of investigation in CSZ earthquakes (e.g., structural damage, tsunamis, landslides), the most consequential could conceivably be soil liquefaction. Liquefaction involves the sudden loss of soil shear strength, results in permanent ground deformations, and routinely damages an array of infrastructure, including buried lifelines, buildings, port structures, roads, levees, and bridges. And, because these impacts threaten mobility across all modes of transportation, liquefaction can also severely hinder post-event response and recovery. In the CSZ, for example, the Department of Homeland Security recently predicted that 1712 miles of road, 1128 miles of rail, 837 bridges, and 8 ports will be unavailable after an M9 earthquake due predominantly to liquefaction (DHS, 2019). While these predictions employed many simple assumptions and beg further research, they highlight the potentially staggering impact of liquefaction in large CSZ earthquakes. Any serious analysis to inform earthquake planning and response, or to focus mitigation resources, must consider it. In fact, liquefaction is in some earthquakes responsible for more economic loss than any other effect, as exemplified in Canterbury, NZ, where liquefaction induced ~$10B in damage (Parker and Steenkamp, 2012).

1.3 Prior Predictions of Liquefaction in CSZ Earthquakes

Mechanistic models based on in-situ geotechnical tests were first developed to predict liquefaction in the early 1970's (e.g., Seed and Idriss, 1971). Significant expansions in training data and in the understanding of liquefaction mechanics has today led to an ensemble of models trusted worldwide in engineering practice (e.g., as summarized by Geyin and Maurer, 2021a). These models predict the occurrence and damage-potential of liquefaction and have been shown to have global prediction efficiencies of ~75-85% (e.g., Maurer and Sanger, 2023). In the 2001 Nisqually, WA, earthquake, for example, an ensemble of popular models was shown to be ~90% efficient (Rasanen et al., 2023). Although other earthquake impacts, such as landslides, tsunamis, and structural damage, have been predicted in CSZ ruptures, we are unaware of any study that has used state-of-practice models to predict liquefaction response at sites throughout the region.

In contrast, beneath this modeling tier are first-order geologic and "geospatial" models that infer soil traits from geology maps and other freely available parameters. The DHS (2019) study, for example, used the "HAZUS" program (FEMA, 2019), which is popular with planners and emergency managers, and wherein geologic units are inferred to have broad areal classifications of liquefaction hazard (e.g., "low" vs. "high"). Regional-scale liquefaction predictions for an M9 event were also made by Rasanen et al. (2021a) using the geospatial model of Rashidian and Baise (2020). Similar to HAZUS, this model infers subsurface conditions from several geospatial variables (estimates of topographic slope, annual precipitation, groundwater depth, and subsurface shear-wave velocity (V_s)), but does not learn from, or anchor to, subsurface geotechnical test data, even though such data is plentiful in the CSZ. While these empirical, first-order predictions provide some semblance of where liquefaction could occur and are useful in some applications, these predictions are unaware of actual subsurface conditions and would not be used in engineering practice to design infrastructure. Global tests of such models (e.g., Geyin et al., 2020) have highlighted that while they can be efficient in certain events and in certain locations, the prediction of subsurface conditions without subsurface tests is exceedingly difficult, especially across diverse landforms and climates. For this reason, mechanistic models based on in-situ test data tend to have significantly higher prediction efficiencies. Accordingly, neither the expected spatial distribution of liquefaction in a large CSZ event, nor its damage-potential for civil infrastructure, are known with any confidence. Amidst considerable investments made to better understand the CSZ and its downstream consequences, comparatively less investment has been made to understand the liquefaction hazard.

1.4 Significance of Paleoliquefaction Evidence

In regions of infrequent moderate-to-large earthquakes, the historic earthquake catalog may be insufficient to provide accurate inputs to seismic-hazard analyses (i.e., the locations, magnitudes, and recurrence-rates of fault ruptures) or to inform ground-motion predictions for certain earthquake scenarios. The 1700 CSZ interface rupture, for example, was likely much larger than any subsequent event in the region, yet there is little-to-no eyewitness account of it (Thrush and Ludwin 2007), let alone ground-motion records. Paleoseismic evidence must therefore be relied on to elucidate and constrain the seismic hazard. To that end, information about CSZ earthquakes and their effects has been inferred from dendrochronology (e.g., Atwater et al., 1991), turbidites (e.g., Goldfinger et al., 2012; Atwater et al., 2014), tsunami deposits (e.g., Peters et al., 2007), microfossils (e.g., Engelhart et al., 2013), geochemical markers (e.g., O'Donnell et al., 2017), seafloor morphology (e.g., Watt et al., 2017), and soil liquefaction (Obermeier 1995), among other indicators of seismicity. Notably, the actual characteristics of ground-motions experienced in 1700 (or any other prehistoric earthquake) can be determined only through inverse-analysis of this evidence.

However, most artifacts are only loosely correlated to the intensity of shaking, if at all. That is, the evidence may suggest an earthquake occurred, but do little to quantify the shaking intensities experienced. In this regard, liquefaction is arguably the only artifact that is both: (i) presently documented at sites across the CSZ; and (ii) capable of "recording" motions. This latter distinction arises because well-validated, probabilistic, mechanics-based models are available to predict liquefaction as a function of seismic loading. In turn, these models can be inverted to back-calculate the ground motions that would, and would not, induce the observed evidence. In this way, liquefaction is more than a proxy of shaking, given that its presence and absence can be used to quantitatively constrain rupture characteristics and ground motion intensities across a region (e.g., as demonstrated by Rasanen and Maurer 2022ab; 2023). It should be noted that while earthquake parameters have also been derived elsewhere from felt-intensity reports (e.g., Rasanen and Maurer, 2023) and landslide evidence (e.g., Rasanen and Maurer, 2022), we are unaware of any such reports from 1700 CE, and relatively few, if any, landslides have been linked to CSZ interface earthquakes (Struble et al., 2017). Landslide evidence is also complicated by the fact that aseismic slides are common and can be difficult to distinguish from their coseismic counterparts, whereas aseismic phenomena resembling liquefaction are less common (e.g., Obermeier et al., 2011). Paleoliquefaction features have to date been identified at 185 locations in the Pacific Northwest, as compiled by Rasanen et al. (2021) from 24 publications documenting field investigations.

1.5 Prior Analyses of CSZ Paleoliquefaction

We are aware of one prior study in which estimates of shaking intensity were derived from CSZ paleoliquefaction. Obermeier and Dickenson (2000) conducted field investigations on Columbia River islands, building upon and benefiting from prior efforts (e.g., Atwater 1994, Obermeier 1995). This resulted in the documented presence, or judged absence, of paleoliquefaction on several islands, with features generally constrained to the year 1700 and decreasing in size and frequency moving inland. Obermeier and Dickenson (2000) used liquefaction models to back-calculate the motions that occurred on this transect in 1700, and notably, suggested that peak ground accelerations (*PGA*s) were as low as 20% of those predicted by recent forecasts (e.g., Frankel et al., 2018; Wirth et al., 2018). As stated by Obermeier and Dickenson (2000): "Our interpreted levels of shaking are considerably lower than current estimates." While this result is provocative, it comes with several important caveats.

First, Obermeier and Dickenson (2000) used the liquefaction triggering model of Youd and Nobel (1997), which relies on standard penetration test (SPT) data. However, the Youd and Nobel (1997) model has rarely been used since the 20th century; subsequent models have made important modifications and additions (e.g., to account for factors then unknown to be significant). Second, as new in-situ test methods have been developed, SPT-based models have fallen out of favor, with the more sophisticated cone penetration test (CPT) now recognized as preferable (NAE, 2021). Third, SPT measurements were not actually made. Instead, a small hand-held variant was used and then correlated to SPT measurements by unknown means. These test results were then input to the Youd and Nobel (1997) model, which predicts liquefaction at depth within a soil profile. Fourth, to predict whether liquefaction should be expected to manifest at the surface, Obermeier and Dickenson (2000) utilized the Ishihara (1985) model. However, newer manifestation models trained on significantly more data, and which also offer probabilistic predictions, are now available (e.g., van Ballegooy et al. 2014; Maurer et al. 2015). Fifth, liquefaction models like Youd and Nobel (1997) traditionally had embedded conservatism, such that the binomial threshold for triggering corresponded to a relatively low probability of liquefaction. This probability is now generally recognized to be ~15% (e.g., Boulanger and Idriss, 2014). The PGAs constrained by Obermeier and Dickenson (2000) may thus correspond to the 15th percentile of what possibly occurred, rather than to a median estimate. As stated by Obermeier and Dickenson (2000): "Our arguments are based largely on qualitative inferences of liquefaction susceptibility supported by preliminary geotechnical data." Considering major advances in liquefaction analytics made over the last two decades, CSZ paleoliquefaction features can and should be investigated using modern tools and methods. As concluded by Obermeier and Dickenson (2000): "More paleoliquefaction and geotechnical field studies are needed to bracket the strength of shaking."

1.6 Objectives

The relative lack of attention given to liquefaction in large CSZ interface earthquakes is notable when considering its potential impact looking forward (i.e., for damage prediction), and its potential value looking backward (i.e., to constrain the characteristics of past earthquakes). Accordingly, this study will first predict liquefaction-induced ground failure across the region for each of the 30 ground-motion simulations of Frankel et al. (2018) and Wirth et al. (2018). This effort will use the latest methods to predict site effects and liquefaction response, and the highest quality subsurface test data available. Nonlinear site-response analyses will propagate simulated motions to the surface and an ensemble of 18 CPT-based models will predict liquefaction manifestations at 400 locations in Oregon, Washington, and British Columbia. Second, this study will quantitatively assess which of the 30 ground simulations represent more and less plausible realizations of the motions experienced in 1700, as evidenced by the presence or absence of 1700 paleoliquefaction. This effort will compare predictions and observations at 10 distributed sites and will compute the probability that a given simulation would produce the set of field observations. Inherent to this process, the uncertainties of field observations (i.e., whether liquefaction did, or did not, occur in 1700) will be accounted for and modeled. This effort is made possible by decades of prior field research and will shed new light on the regional patterns of shaking intensity experienced in 1700. Collectively, this study provides the first rigorous investigation of the expected impacts of liquefaction throughout the Pacific Northwest in CSZ M9 ruptures, as well as the first rigorous inverse-analysis of 1700 CE paleoliquefaction evidence at regional scale. In this way, the study provides valuable insights looking both forward and backward.

2.0 DATA AND METHODOLOGY

To predict liquefaction response at 400 locations, CPT measurements and other surface data were compiled from numerous geotechnical investigations. Major sources of such data included the databases of Jeschke et al. (2019), Chin and Evans (2021), and the Washington State Dept. of Transportation (WSDOT, 2021). Groundwater depths were adopted from the geotechnical reports that documented CPTs while also considering nearby borings, well logs, and test pits. In general, multiple measurements of the groundwater depth were available in the vicinity of each CPT, given the high density of subsurface explorations in the region. However, because some sites of paleoliquefaction research significance lack existing, proximal CPTs, the authors also newly performed CPTs at seven sites using the U.S. Geological Survey CPT rig. The locations of the 400 study sites are mapped in Figure 2. These sites are spatially distributed in the CSZ but concentrated in major population centers (e.g., Portland, OR; Seattle, WA, and Vancouver, BC), transit corridors (e.g., Interstates 5, 405, and 84; U.S. Route 101), and coastal towns (e.g., Coos Bay, Newport, Cannon Beach, Astoria, and Aberdeen, among others). Also shown in Figure 2 are the PGAs predicted at study sites by simulations “csz-005” and “csz-025,” which are two of the thirty physics-based realizations produced by Frankel et al. (2018) and Wirth et al. (2018).

PGAs expected at study sites often differ considerably between the two realizations. Relative to simulation csz-025, csz-005 predicts somewhat stronger PGAs inland in the south (e.g., Portland) and weaker PGAs inland in the north (e.g., Seattle and Vancouver). In Seattle, PGAs are up to 260% larger at study sites in csz-025 relative to csz-005. On the Pacific Coast, both simulations predict strong shaking, but these predictions often differ greatly in the far north and far south. At Cape Blanco, OR, for example, the expected PGAs are 935% larger for csz-005. Although both realizations could potentially induce liquefaction at Cape Blanco in susceptible soils (with PGAs of 2.15 g and 0.23 g), the probabilities and expected severities of liquefaction manifestations would presumably be very different. While Cape Blanco provides an extreme example of the uncertainty of expected ground motions in an M9 CSZ event, the expected PGAs differ across the 30 simulations by at least 100% at all study sites. It is thus of interest to compute the distribution of expected liquefaction responses at each site, and to bound the conceivable “best case” and “worse case” scenarios in M9 interface ruptures.

Importantly, the physics-based simulations of Frankel et al. (2018) and Wirth et al. (2018) were underlain by the Stephenson et al. (2017) velocity model, which assigned a surficial shear-wave velocity (V_s) of 600 m/s. Although this surficial layer is adequate where soft rock is present at the surface, the velocity of soil is much lower. The PGAs shown in Figures 2a and 2b were thus predicted without considering the presence and influence of soils, which can dramatically change the ground motions. It would not be unexpected, as an example, for soft soil to increase or decrease intensity measures of engineering interest by 100% or more. To predict liquefaction at study sites, the simulated ground motions were therefore modified for surficial conditions using wave-propagation based site response analysis.

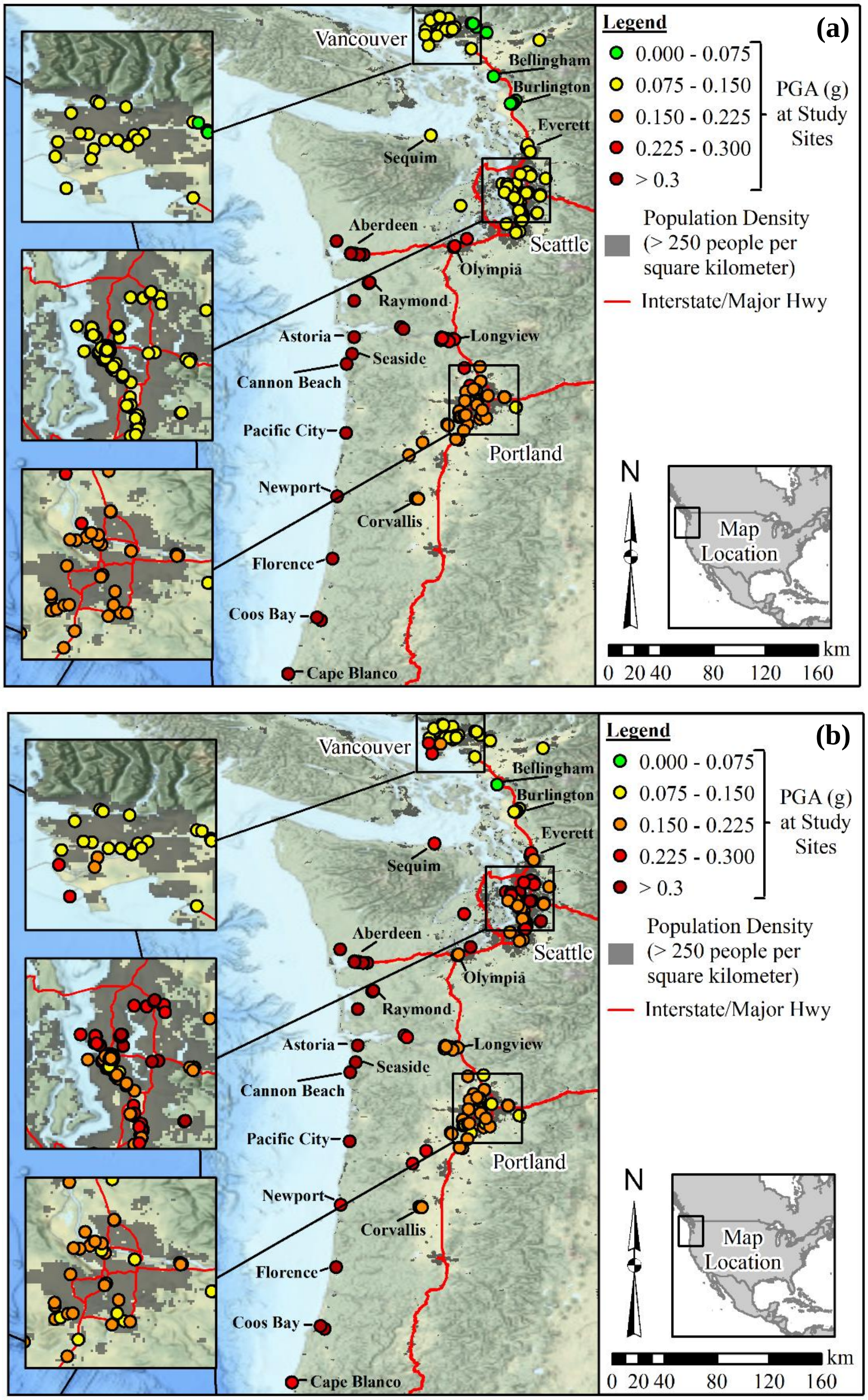


Figure 2. PGAs at 400 study sites predicted by ground-motion simulations: **(a)** csz-005 and **(b)** csz-025. PGA = peak ground acceleration; csz = Cascadia Subduction Zone.

2.1 Shear-Wave Profiles and Site Response Analysis

Two horizontal components of motion for 30 simulations were propagated to the surface at each of 400 study sites, for a total of 24,000 analyses. This required a measurement or estimate of the V_S versus depth profile at each site, which is a challenge given the quantity and regional distribution of study locations. For 394 sites, a common approach was used to develop V_S profiles. V_S was first estimated by averaging the predictions from six CPT-V_S models, which predict V_S from CPT measurements. These models included Andrus et al. (2007), Robertson (2009), McGann et al. (2015), and three different models trained by Assaf et al. (2023) using data from the Vancouver metropolitan region. However, CPTs oftentimes do not extend to a deep enough depth to reach V_S = 600 m/s, which is sometimes far more than 20 m (i.e., a common target depth of CPT testing).

Accordingly, the region-specific Wirth et al. (2021) soil velocity model, which predicts the near-surface V_S profile from V_S time-averaged over the top 30 m (V_{S30}), was also used to generate V_S profiles to 600 m/s. This model was selected over others (e.g., Marafi et al., 2021) because it was perceived to be slightly better at predicting profile characteristics that give rise to high-frequency ground motions, which liquefaction is most sensitive to, and best predicted by (e.g., Sideras, 2019; Wu et al. 2022). V_{S30} was predicted by the Geyin and Maurer (2023) model, which may be particularly advantageous in the Pacific Northwest given its anchorage to V_{S30} measurement sites via regression kriging, and considering that such measurements are plentiful in the region. Thus, where study sites are proximal to prior V_{S30} measurements, the predicted V_{S30} approaches the measured V_{S30}. To avoid phantom velocity contrasts that could result from concatenating the Vs profiles predicted from CPT data and by Wirth et al. (2021), the Wirth et al. (2021) functional form was refit to the dual Vs predictions. An example of this process is shown in Figure 3, where it can be seen how local measurements update the generic profile given by the regional model. For the remaining six study locations, which are located near sites of paleoliquefaction research significance, active and passive multi-channel analysis of surface wave (MASW) testing was performed to produce site-specific V_s profiles to a depth of V_s = 600 m/s. Additional details of the MASW testing, and the profiles derived from these measurements, are described by Rasanen et al. (2024).

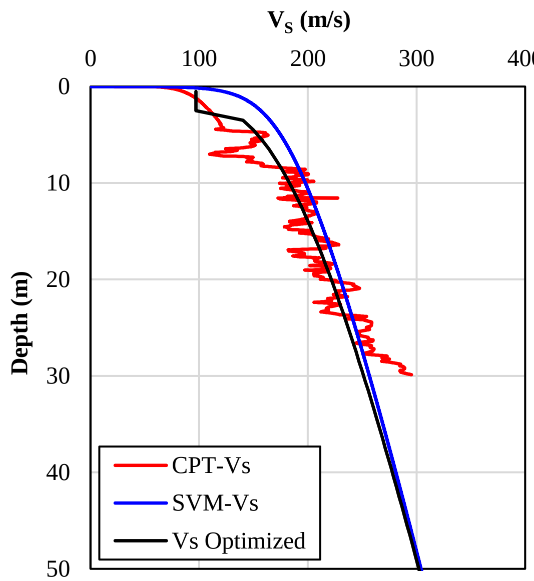


Figure 3. Example V_s profile developed from CPT data and the Wirth et al. (2021) soil velocity model (SVM) (note: only the first 50 m of the profile is shown).

To adjust the simulated motions from reference (V_S =600 m/s) to site-specific conditions, total stress site response analyses were performed using PySeismoSoil, a python implementation of the 1-D site response software platform SeismoSoil (Asimaki and Shi, 2017). Although effective-stress site response analyses could offer advantages and facilitate an entirely different approach to predicting liquefaction (i.e., the time-varying prediction of excess pore pressure), total-stress analyses are desired for consistency with the state-of-practice "stress-based" liquefaction models utilized in this study. Both equivalent linear analyses with frequency-dependent soil properties (EQL-FD) and fully non-linear analyses (NL) were performed. The latter is commonly presumed to be more accurate and is adopted in subsequent predictions of liquefaction, but the former was also performed to gain an understanding of where, and to what extent, the two methods make discrepant predictions. This comparison and its implications will be discussed later in the paper, as will the uncertainties of site-response analysis more broadly.

EQL-FD analyses were performed in the frequency domain following Assimaki and Kausel (2002), while the NL analyses were performed using the finite difference method in the time domain. Features of the NL approach include a memory-variable technique to model small-strain damping (Liu and Archuleta, 2006) and the hysteresis behavior model of Li and Assimaki (2010). For both the EQL-FD and NL analyses, V_S profiles were resampled in SeismoSoil such that a single wavelength in a given layer of constant V_s is represented by 10 discretization points, resulting in a spatial grid size (Δh) in a given sublayer of $\Delta h = V_s/(10f)$, where f is the maximum frequency of interest, with the default in SeismoSoil set to 30 Hz. The Hybrid Hyperbolic (HH) model, which has shown improved performance over the popular modified hyperbolic (MKZ) model, was used to represent the 1-D stress-strain behavior in each sublayer (Shi and

Assimaki, 2017). The HH model parameters and corresponding modulus reduction and damping curves were calibrated using the empirical correlations in Appendix D of Shi and Assimki (2017), estimated directly from the V_s in each layer. The simulated motions were modified prior to input using ground motion processing modules in PySeismoSoil; the time histories were initially truncated according to the Arias intensity, with timesteps before 0.001% and after 99.999% Arias intensity cut from the time series, then baseline corrected. Both EQL-FD and NL analyses were performed for each horizontal component of each simulated motion at each study site, from which the geometric mean horizontal PGA was computed.

2.2 Liquefaction Modeling

Six CPT-based models were used to predict the triggering of liquefaction over the depths of each CPT profile. These included Robertson and Wride (1998), Architectural Institute of Japan (2001), Moss et al. (2006), Idriss and Boulanger (2008), Boulanger and Idriss (2016), and Green et al. (2019). A CPT inversion filter was not applied to correct CPT data for thin-layer effects, based on the findings of Geyin and Maurer (2021b) and Yost et al. (2021). Prior to using any triggering model, liquefaction susceptibility was inferred from the soil behavior type index (I_c) (Robertson and Wride, 1998). Given the uncertainty in the relationship between I_c and susceptibility, the Maurer et al. (2019) model was used to predict the probability of susceptibility, as defined by Boulanger and Idriss (2006), from the measured I_c. Because liquefaction at depth may be inconsequential, or not express, at the ground surface, the results from triggering analysis were input to three other models to predict surficial manifestations of liquefaction (i.e., "ground failure"). While liquefaction manifestations do not necessarily indicate damage to infrastructure, they are a pragmatic and commonly adopted proxy of damage potential. The adopted manifestation models included Iwasaki et al. (1978), van Ballegooy et al. (2014); and Maurer et al. (2015). Using global case-history data, Geyin and Maurer (2021a) conditioned fragility functions on each of these 18 model combinations (i.e., six triggering models and three manifestation models), thereby permitting probabilistic predictions of surficial liquefaction manifestations. With this approach, the probability of manifestation was computed as:

$$P(\text{Manifestation}) = \int_{I_c} P(\text{Manifestation}|PGA_{Sim}, I_c\ threshold)\, f(I_c) \cdot \mathrm{d}I_c \qquad (1)$$

Where $f(I_c\, threshold)$ is the probability density function (PDF) of the I_c threshold for discriminating susceptibility (Maurer et al., 2019); PGA_{Sim} is the site adjusted PGA from NL site response analysis; and $P(\text{Manifestation}|PGA_{Sim}, I_c\ threshold)$ is the probability of surficial manifestation, which is conditioned on PGA_{Sim} and the threshold I_c for determining susceptibility, among numerous other inputs, and which is computed via the fragility functions of Geyin and Maurer (2021a). Fragility coefficients specific to each of the 18 models were obtained from Table 4 of Geyin and Maurer (2021a). Rasanen et al. (2023) previously tested the prediction efficiencies of the 18 models in the study area using 24 case histories from the 2001

Nisqually, WA, earthquake and found that while model efficiencies varied, the differences were rarely statistically significant considering finite-sample uncertainty. Consequently, in this study, the predictions from all 18 models were ensembled with equal weighting.

2.3 1700 CE Paleoliquefaction Study Sites

Ten well-studied sites of 1700 CE research significance were analyzed. These sites were developed utilizing the paleoliquefaction compilation of Rasanen et al. (2021) and considering both the presence of existing CPTs and the authors' ability to perform new CPTs. Undoubtedly, additional sites could be developed and analyzed in the future. To determine whether any ground-motion simulation best represents the motions experienced in 1700 CE, each site must be classified as "positive" (1700 CE evidence is observed) or "negative" (no 1700 CE evidence is observed). All "negative" sites have liquefaction evidence from earthquakes prior to and/or more recent than 1700 CE but lack 1700 CE evidence despite relatively extensive field studies, which sometimes included documentation of dozens of liquefaction features from other events and the searching of hundreds of square meters of exposed bank or tide flat. Obermeier (1995), for example, searched more than 100 km of riverbank in six different valleys in Western Washington and noted that at least 20% of the traversed length "had banks so freshly and cleanly eroded that dikes exceeding one cm in width would have been noticed." However, because "absence of evidence is not evidence of absence," it cannot be certain that "negative" sites are negative. Similarly, a liquefaction feature with a relatively well constrained age (e.g., 1700 ± 10) does not mean it was induced by the 1700 CE interface event, given the prevalence of crustal and intraslab earthquakes in the region, and the possibility that the feature might not be of seismic origin. Thus, it also cannot be certain that "positive" sites are positive. This is of course true for all forms of paleoseismic evidence.

To account for and model this uncertainty, each of the ten study sites was assigned a "field classification score," as defined in Table 1. These scores were assigned based on the rigor of prior field studies (e.g., the extent of searching or the quantity of features discovered), the age constraints given (considering both radiocarbon dating and the authors' holistic interpretations of all evidence), and conversations with the original authors, among other factors. Scores of 9 or 1 indicate 90% confidence in the field classification – the highest assigned – where 9 and 1 respectively indicate a high degree of confidence that liquefaction did, and did not, occur in 1700 CE. A score of 5 indicates a complete lack of confidence, with "positive" and "negative" being equally probable. This is akin to removing a study site from analysis altogether, since sites classified as 5 can do nothing to constrain the ground motions experienced. Scores less than 5 indicate a site is assigned as "negative" and scores greater than 5 indicate the opposite. Importantly, the sensitivity of the study's conclusions to these scores was also rigorously investigated, as will soon be discussed. The locations of the ten sites are shown in Figure 4, where the assigned, preferred field classification scores are also given. The locations of other paleoliquefaction study sites and of other CPTs, which could be used to develop additional analysis sites in the future, are also shown. In the following, the ten paleoliquefaction research sites are each succinctly described. The reader is referred to Rasanen et al. (2021), and most especially to the original publications cited below, for important notes and additional details (e.g., sand blow thicknesses, dike widths, dating information, interpretation of stratigraphic relationships).

Table 1. Assignment of observation uncertainties to 1700 CE study sites.

Field Classification Score	Probability that Observation is Positive (%)	Probability that Observation is Negative (%)
1	10	90
2	20	80
3	30	70
4	40	60
5	50	50
6	60	40
7	70	30
8	80	20
9	90	10

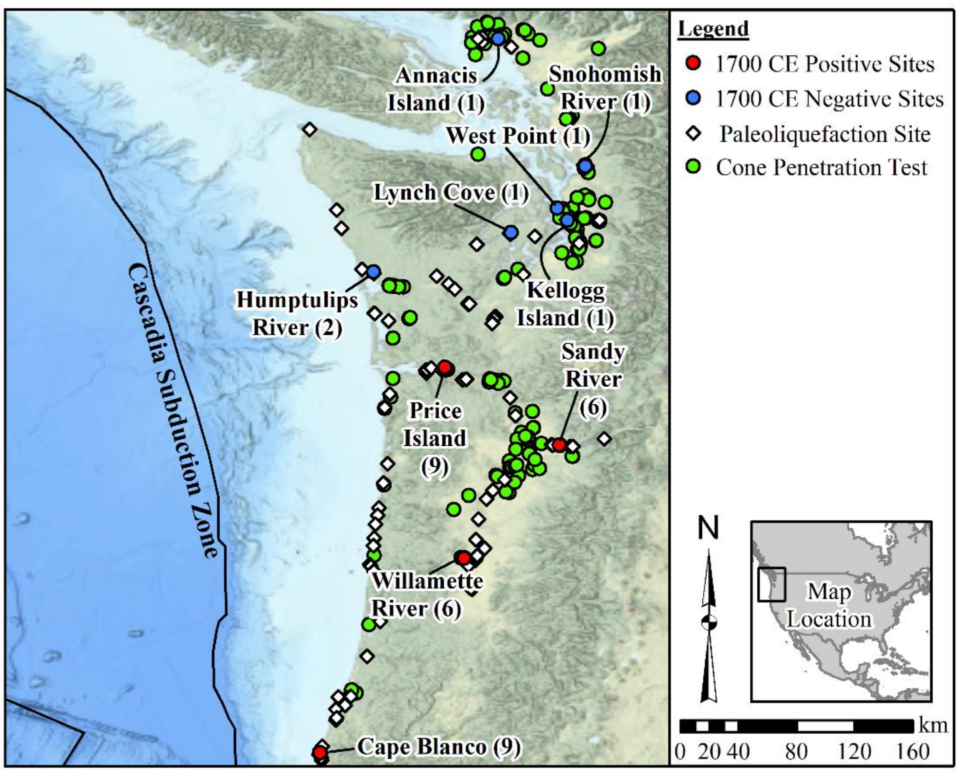


Figure 4. Sites of 1700 paleoliquefaction evidence analyzed herein.

West Point (Discovery Park), Seattle WA: Atwater (1999) obtained a radiocarbon date of 900-930 CE for a Douglas fir log that was embedded in a tsunami deposit discovered in 1992 during the expansion and upgrade of the West Point water treatment plant in Seattle. Atwater (1999) indicated the date of 900-930 CE applies to the subsidence and tsunami found at West Point, in addition to the uplift along the south side of the Seattle Fault, which produced the north-moving tsunami that left the tsunami sand deposits at West Point and Cultus Bay. Sand dikes were discovered extending up to the tsunami deposit at West Point (Atwater personal communication, 2020), indicating that the liquefaction that produced the sand dike was due to the same event that caused the tsunami. No liquefaction features younger in age (i.e., more recent) were discovered at WestPoint, despite >100-meter long excavations extending well below and above the feature discovered. Therefore, this CPT was assigned a score of 1 (10% probability that observation is positive, and 90% probability observation is negative).

Snohomish River, Everett, WA: Bourgeois and Johnson (2001) documented at least three episodes of paleoliquefaction in the Snohomish River delta that occurred after 800 CE. Several of these paleoliquefaction features (sand dikes) extend up to a tsunami deposit that was dated to 850-980 CE via radiocarbon dating of stumps of spruce trees that were submerged by earthquake induced subsidence. Bourgeois and Johnson (2001) indicate that the tsunami deposit at the Snohomish River sites, which the sand dikes extend up to, was consistent with the widespread tsunami deposit (caused by a north-moving tsunami) found at Cultus Bay and West Point dated to 900-930 CE (Atwater, 1999), indicating it may have been produced by the last Seattle fault earthquake dated to 900 – 930 CE. The youngest episode of liquefaction was inferred to have occurred ca. 1430 – 1640 CE based on the radiocarbon dating of a bark-free twig found within the sand volcano produce during the liquefaction event. Bourgeois and Johnson (2001) therefore considered it unlikely this liquefaction feature, nor any other feature discovered during their extensive search of the lower Snohomish delta, was caused by the 1700 CE Cascadia earthquake. Furthermore, they stated "there is no evidence that the A.D. 1700 Cascadia earthquake produced liquefaction in the Puget Lowland." Thus, this site was assigned a classification score of 1.

Kellogg Island, Seattle, WA: Located in the Duwamish Waterway, Kellogg Island offers a unique window into the pre-anthropogenic Duwamish River delta landscape (Feliks 2019). Despite fill placement in the late 1930s, paleoliquefaction is preserved on the island and has been documented during low tides. Davis (2018, 2019), for example, documented two distinct episodes of liquefaction, including sand dikes and/or sand volcanoes, at distributed sites on the island and on the east bank of the Duwamish River. Radiocarbon dating indicates these features were formed between 1000 CE and 1640 CE, making it highly unlikely they resulted from the 1700 CE Cascadia earthquake (Davis 2018, 2019). Although the minimum ages (i.e., latest dates of formation) have not been bound by dating for some features in the area, Davis (2018) concluded: “none of the dikes observed are likely to be as young as the 1700 Cascadia earthquake.” This conclusion is consistent with the lack of 1700 CE evidence elsewhere in the Puget Lowland (e.g., Bourgeois and Johnson 2001). Despite investigations at multiple locations with recurrent liquefaction episodes dated to other earthquakes, no evidence of the 1700 CE Cascadia earthquake has been found in Seattle, or in the broader Puget Sound. Therefore, this site was assigned a classification score of 1.

Humptulips River, Olympic Peninsula, WA:

Obermeier (1995) identified a liquefaction dike that cut through charcoal dated to approximately 2500 BC, suggesting that the liquefaction event was younger than the dated age. While Obermeier considers the possibility that the liquefaction feature could have been caused by the 1700 CE Cascadia earthquake, he also notes that this is unlikely. The uncertainty surrounding the age of the liquefaction event results in a score of 2 (20% probability that observation is positive, and 80% probability that observation is negative).

Price Island, Columbia River, WA

Liquefaction feature likely caused by January 26th, 1700 CE Cascadia Subduction Zone earthquake based on geological evidence (Atwater, 1994), tree-ring dating (Atwater et al., 1991; Jacoby et al., 1997; and Yamaguchi et al., 1997) and tsunami records in Japan (Satake et al., 1996). More specifically, the liquefaction features dikes extruded through and vented on top of a buried soil (Atwater, 1994). The buried soil is correlated to coseismic subsidence at (i) Willapa Bay and the Copalis River, which has been dated to be between 1680-1720 CE through dendrochronology (Atwater et al., 1991); and (ii) at Price Island, Willapa River, and Copalis River, which has been dated to be between 1699-1700 CE through dendrochronology (Jacoby et al., 1997; Yamaguchi et al., 1997). Because the coseismic subsidence was likely caused by the 1700 CE CSZ earthquake (Atwater et al., 1991), and because Satake et al. (1996) dated this CSZ earthquake from tsunami evidence to January 26th, 1700 CE, the buried soil is likely from 1700 CE. The stratigraphic position of the liquefied sand venting on top of the buried soil (sand blow), indicate the liquefaction is likely from the 1700 CE CSZ earthquake. A maximim limiting age can found from the mud cap above the sandy source layer that has two marker beds with the most recent being the ash and pumice likely from Mt. St. Helens erupting between 1479-1482 CE. In addition, the maximum calibrated 2 sigma (from tubers in soil older than the liquefaction) was 710-1021 CE (at the oldest 710 CE). Thus, based on the maximum limiting age of the tubers (at the oldest 710 CE), the maximum limiting age of the Mt. St. Helen's ash layer (1479-1482 CE), and the age of the buried soil likely being from 1700 CE, the liquefaction feature at this site is highly likely to be from the 1700 CE CSZ earthquake and thus is assigned a score of 9 (90% probability that the observation is positive and 10% probability it is negative). Regarding the buried soil, radiocarbon ages, tree-ring counts, inferred stratigraphic correlations, and tsunami records together suggest that the submergence resulted from a plate-boundary earthquake in 1700 CE. Furthermore, just like

Price Island, Hunting, Brush, and Marsh islands nearby in the Columbia River also have vented sand on top of the buried 1700 CE soil (Atwater 1994). In addition, Price, Hunting, and Marsh islands have sand dikes intruding through the 1700 buried soil (Atwater 1994). Considering all the evidence at Price Island, as well as its presence at other nearby islands, the liquefaction is highly likely to have been produced by the 1700 CSZ earthquake.

Annacis Island, Vancouver, BC

Clague et al. (1997) discovered two large sand blows and several sand dikes in a 6-meter-deep large excavation for the expansion and upgrade of a sewerage treatment facility at Annacis Island in Vancouver, BC, part of the Fraser delta. Clague et al. (1997) assigned a maximum calibrated 2 sigma age to the liquefaction features of 135-408 CE based on three radiocarbon age dates obtained from a log buried by the vented sand, as well as a branch and western redcedar leaves buried just below the sand blow. Clague et al. (1997) indicate this is a closely limiting maximum age for liquefaction based on the location of the dated features (i.e., either buried by the vented sand or in the paleosol immediately below). This date is similar to that of subsidence at estuaries in southern Washington believed to be from a plate boundary earthquake (e.g., Atwater and Hemphill-Haley, 1996). While Clague et al. (1997) indicate a plate boundary earthquake could have caused this liquefaction, they favor that a crustal fault earthquake caused it based on subsidence in the Fraser delta. More specifically, Clague et al. (1997) state: "We favor this alternative hypothesis [the liquefaction features at Annacis Island were from a crustal fault] because the subsidence that is inferred to have accompanied the earthquake is more consistent with displacement on a crustal fault than with plate-boundary rupture (Mathewes and Clague 1994)." While other liquefaction features have been found in the Fraser delta (Clague et al. 1992), no conclusive evidence of more than one liquefaction event at Annacis Island or any other site in the Fraser delta has been established. Consequently, the recorded instances of liquefaction in the Fraser delta are scarce, with no direct association to the 1700 CE plate boundary earthquake nor any event after approximately 300 CE. Therefore, the Annacis Island site is assigned a classification score of 1.

Willamette River, Corvallis, OR

During an archaeological survey for a municipal wastewater reclamation project on the east bank of the Willamette River, Peterson et al. (2014) documented several sand dikes and sills in trench exposures. These features were located approximately 150 meters east of the modern Willamette River channel and intruded into overbank mud deposits from a depth of 3.4 meters to 0.4 meters. Radiocarbon dating conducted by Peterson et al. (2014) on peat, wood, and charcoal provided a maximum calibrated 2 sigma age of 1400-1440 CE for the liquefaction features, indicating that the liquefaction event occurred after this period. Furthermore, based on historic earthquake records and the absence of reported large earthquakes in northwest Oregon over the last 175 years (as noted by Berg and Baker, 1963), Peterson et al. (2014) argue that the liquefaction features at this site could potentially be attributed to the 1700 CE earthquake. It is important to acknowledge that the Corvallis site exhibits greater uncertainty compared to other sites in terms of dating and stratigraphic relationships. Consequently, a classification score of 6 is assigned to this site, indicating a 60% probability that liquefaction occurred during the 1700 CE Cascadia Subduction Zone earthquake and a 40% probability that the liquefaction was unrelated to the event. It should be noted that while this uncertainty slightly favors liquefaction from the 1700 CE earthquake, its impact on the final results is negligible, as discussed later on.

Sandy River, Troutdale, OR

At the confluence of the Sandy River and Columbia River, small sand dikes and an abundance of sills were discovered by Obermeier (1995) and Peterson and Madin (1997). Peterson and Madin (1997) obtained a maximum calibrated 2 sigma age of 1413-1642 CE and a minimum calibrated 2 sigma age of 1461-1878 CE, effectively constraining the liquefaction features to have occurred between 1413-1878 CE. Local historic seismicity near Portland, as recorded by Siskowic et al. (1994), indicates events of M_w5-6 have occurred. However, these local events have not left apparent evidence of liquefaction along the Sandy River in the most recent flood deposits (post ~1850 CE), as noted by Obermeier (1995). Furthermore, Siskowic et al. (1994) report that sediments after approximately 1850 CE appear to be undeformed. Noteworthy, Obermeier (1995) and Peterson and Madin (1997) in their searches along several islands of the Columbia River, observed a general trend of decreasing dike widths as moving further upstream the Columbia River, which aligns with expectations if the earthquake source was the CSZ interface rupture. However, localized zones of dike width variability exist. Obermeier (1995) also computed that a PGA of only 0.1 g would be required to induce liquefaction at the ground surface at Sandy River, although this result is subject to the same limitations discussed in Section 2.5. Considering the available evidence, both Obermeier (1995) and Peterson and Madin (1997) conclude that the 1700 CE Cascadia Subduction Zone earthquake could have been responsible for the liquefaction at the Sandy River site. As a result, a classification score of 6 is assigned to this site. Importantly, the sand dikes at the Sandy River site terminate at least 0.5 meters below the current ground surface (Peterson and Madin 1997). While this is quite close to the ground surface, this adds uncertainty to the analysis given that liquefaction manifestation models are used. Thus, the models used in the likelihood analysis would require a higher loading than would be required to produce features that terminate 0.5 meters below the ground surface.

Cape Blanco, Port Orford, OR

Kelsey et al. (2002) discovered and dated sand dikes from two distinct episodes of liquefaction at a cutbank site along the Sixes River at Cape Blanco State Park. Through radiocarbon dating and stratigraphic relationships, Kelsey et al. (1998, 2002) established that the first episode of liquefaction occurred around 500 BC, while the second episode occurred in 1700 CE, both induced by CSZ interface earthquakes. Kelsey et al. (1998) observed that the second episode of liquefaction intrudes through the same buried soil as the one found at the mouth of the Sixes River, which experienced coseismic subsidence inferred to be from the 1700 CE Cascadia earthquake. The roots of spruce trees within the youngest buried soil at the site of liquefaction, provided a maximum limiting calibrated age of 1707 CE, indicating that the spruce trees died after the 1700 CE event while the subsidence causing the buried soil occurred during the 1700 CSZ earthquake (Kelsey et al. 1998). The second episode of liquefaction intrudes through this buried soil and extends up to approximately 0.15 meters from its top (i.e., the 1700 ground surface). The latter supports the use of a manifestation model for this site, given the dikes proximity to the 1700 ground surface. Based on the evidence in Kelsey et al. (1998, 2002), a classification score of 9 was assigned to the site.

Lynch Cove, Belfair, WA

Martin and Bourgeois (2012) interpreted a sand layer found at several sites across Lynch cove as vented sediment from earthquake induced liquefaction. This sand layer, which is up to 100 cm thick, is found between tidal-flat mud abruptly overlain by forest peat. The abrupt mud to forest peat contact is interpreted as evidence for coseismic uplift with an age of 680 to 1020 CE (Bucknam et al. 1992). Martin and Bourgeois (2012) searched 182 cores and excavations at Lynch Cove, 53% of which contained the sandy deposit between the mud and peat. The sand deposit ranges in thickness, but on average (where present) was 24 cm. While this sand layer was previously interpreted to be a tsunami deposit (i.e., Hemphill-Haley 1996 and Jovanelly and Moore 2009), Martin and Bourgeois (2012) interpret this sand layer to be vented sediment rather than a tsunami deposit based on: (i) the range of sand layer thickness, at both small and large scales; (ii) the absence of sand at half the sites; (iii) the lack of the sand layer thinning landward; (iv) the high average thickness of the sand layer; and (v) the presence of intraclasts from a lower stratigraphic level. Furthermore, at 21 of the sites sand dikes were found leading to the vented sand layer. Based on the sand layer residing between the mud and peat, along with the liquefaction sand dikes extending to this vented sand layer, this indicates that the liquefaction that caused the dikes and vented sand occurred at the time of coseismic uplift (i.e., 680 to 1020 CE). There are several possible earthquake sources during this date range (e.g., Seattle fault, Tacoma fault, and Cascadia Subduction Zone). However, because the focus of this study is on the 1700 event, this site is assigned a classification score of 1.

2.4 Constraining 1700 CE Ground Motions

The simulations most and least likely to match the classified observations at the 10 sites were identified through likelihood analysis. Specifically, the likelihood that a simulated ground motion represents a past earthquake, given a set of field observations, may be computed as the product of the probabilities of those observations, conditioned on the simulation. Thus, the likelihood of any simulation, given a set (x) of field observations at N different sites, is computed as:

$$Likelihood\ (Sim|x) = P(X = x|Sim) = \prod_{i=1}^{N} P(X_i = x_i|Sim) \tag{2}$$

where $P(X_i = x_i|Sim)$ is the probability of the outcome observed at site i (liquefaction or no liquefaction), given a CSZ simulation. Incorporating the liquefaction modelling results and the confidence scores assigned to each study site, this probability expands to:

$$P(X_i = x_i|Sim) = P(\text{Manifestation}|PGA_{Sim}, I_c\ threshold)_i \times P(Positive\ Site)_i + [1 - P(\text{Manifestation}|PGA_{Sim}, I_c\ threshold)_i] \times P(Negative\ Site)_i \tag{3}$$

where $P(\text{Manifestation}|PGA_{Sim}, I_c\ threshold)$ is defined by Equation 1 and $P(Positive\ Site)_i$ and $P(Negative\ Site)_i$ are the probabilities assigned to each study site by the authors using Table 1. By repeating this process for all 30 ground-motion simulations, those that are more and less like the 1700 CE event, as evidenced by paleoliquefaction data, are distinguished.

To account for the presence of artificial fill (typically ~1-2 m) at several study sites (i.e., to correct for its absence in 1700 CE), the soil profile and groundwater depth at the time of testing were used to stress-normalize CPT measurements, whereas the trimmed CPT profile with fill removed was used to infer saturation and compute the cyclic stress ratio at the time of shaking. In this regard, the groundwater table in 1700 CE was assumed to be at the same elevation as present day (i.e., it was closer to the ground surface in 1700 CE wherever fill was subsequently placed). While this introduces uncertainty, the overall conclusions are ultimately insensitive to this assumption, as will be discussed, and would only change if the groundwater was considerably deeper in 1700 CE than at present, which is highly unlikely.

3. RESULTS AND DISCUSSION

3.1 Liquefaction Predictions: Looking Forward

Using the data and methods described above, the probability of liquefaction manifesting at the ground surface was computed at 400 study sites for each of 30 ground-motion realizations. Figure 5 shows the median probability computed from all simulations, which represents a best estimate of what would occur in an M9 event. This median probability exceeds 40% at a majority of the sample sites in Seattle and Vancouver, and at a slight minority of sites in the vicinity of Portland. These statements also apply to many other towns and cities with test data. Given the sites available for study, median probabilities exceeding 60% are predicted as far south as Cape Blanco, OR, as far north as North Vancouver, BC, and as far east as Issaquah, WA. Collectively, the predictions suggest that liquefaction could be both pervasive in the region and affect significant hubs of population, transportation, and commerce.

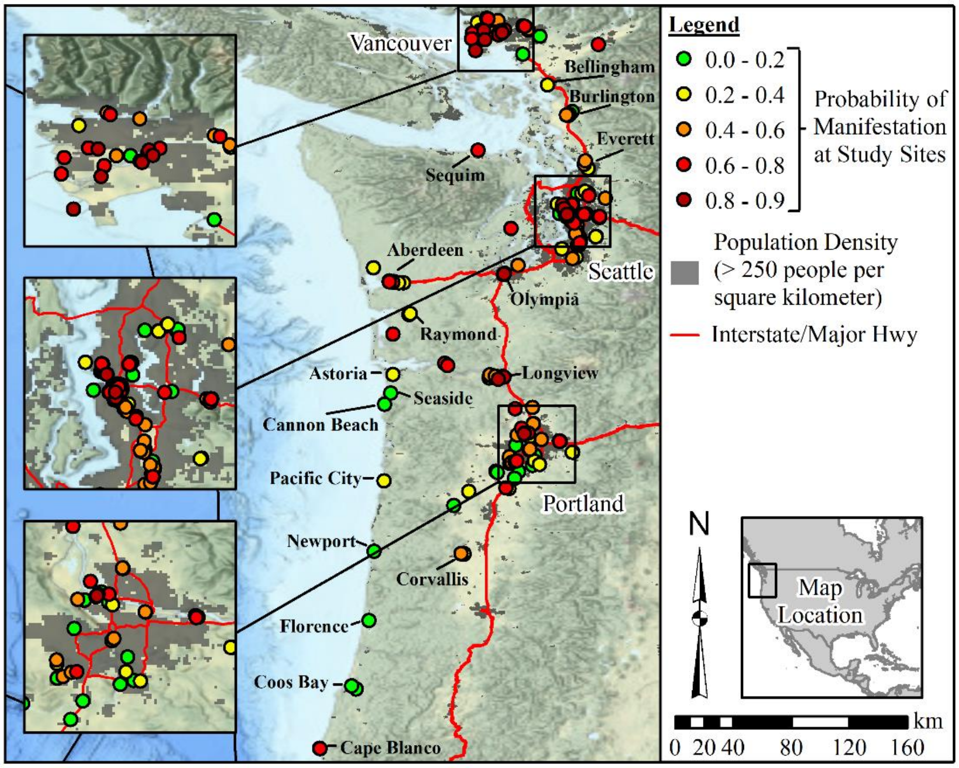


Figure 5. Median probability of liquefaction surface manifestation at 400 CPT sites considering 30 ground motion simulations.

Noticeably, the probabilities of liquefaction manifesting at study sites in many of Oregon's coastal towns are relatively low (0-20%), even though expected shaking intensities are typically strongest there. This is because the sampled soils are interpreted to not be sufficiently loose, saturated, or cohesionless to liquefy, and/or because the strata predicted to liquefy are interpreted to not be sufficiently thick or shallow to produce manifestations at the surface. Most often, study sites in estuarine environments on the immediate coast are interpreted to have a prevalence of clayey soils in the subsurface, with I_c values far exceeding the thresholds commonly adopted to denote susceptibility. Additionally, Obermeier (1995) observed that heading inland from the coast, small river valleys in Washington were mainly filled with coarse gravel and concluded that "highly susceptible sediment appears rare," having found only one small liquefaction feature within 50 km of the coast. However, it must be emphasized – here and again – that while the predictions in Figure 5 and elsewhere may generally extrapolate to areas surrounding each test site, liquefaction is a localized phenomenon and therefore nothing can be confidently concluded at unsampled locations. In this regard, liquefaction manifestations are judged to be unlikely at many coastal test sites, but liquefiable soils do exist on the coast and could exist within the coastal towns where test data is available, but additional data would be needed to assess this. Nonetheless, the liquefaction hazard does appear to be greater elsewhere (e.g., in Seattle, Vancouver, and elsewhere in the Puget Lowlands).

To illustrate the expected liquefaction response at higher spatial resolution, Figure 6a shows the computed median probability of liquefaction manifestation for a portion of Seattle. Median probabilities frequently exceed 50%, and occasionally 80%, in each of the three areas where an abundance of test data is available: the Duwamish, Interbay, and University District neighborhoods. Notably, these are also areas where artificial fill is prevalent and where liquefaction was previously observed in earthquakes in 1949, 1965, and/or 2001 (e.g., Chleborad and Schuster, 1990). By contrasting the predictions of liquefaction in M9 CSZ events against those previously computed by the authors for the 2001 Nisqually earthquake – in which ground motions were densely recorded in Seattle – it may be appropriate to draw preliminary comparisons, at least where test data is available. To that end, the median probabilities in Figure 6a are very similar in Duwamish, but considerably higher in Interbay and the University District (up to 10 times greater), as compared to those computed at the study sites during the Nisqually earthquake using the same liquefaction modeling approach (Rasanen et al., 2023). Therefore, in the "best guess" scenario (i.e., the median prediction), the frequency and severity of liquefaction manifestations in the Duwamish neighborhood may be similar to what occurred there in 2001. It is difficult to draw meaningful comparisons in Interbay and the University District, other than to say that manifestations are likely in M9 events, and considerably more likely relative to the Nisqually earthquake. This change in the expected response is due to the Nisqually earthquake's northward attenuating ground motions, which were sufficient to induce liquefaction in Duwamish, but of only marginal threat to neighborhoods north of downtown. Conversely, an M9 rupture would produce ground motions sufficient for liquefaction across a vast area.

It should again be noted, in the context of Figure 6a, that predictions cannot not be extrapolated well beyond the study sites. In the University District, for example, all study sites are in the vicinity of the University of Washington's athletic facilities, which are generally on land that was formerly beneath the surface of Lake Washington before its lowering ca. 1915, and where artificial fill is common. Yet just several hundred meters to the west, most of the University campus sits at higher elevation and atop glacial till, where the liquefaction hazard is generally expected to be very low. Similar statements could be made for the Interbay and Duwamish neighborhoods, which are likewise surrounded by terrain comprised of less susceptible soils. This should not be interpreted to mean that liquefaction should not be expected in Seattle outside the three neighborhoods discussed above. There are numerous other areas of artificial fill in the city, and as demonstrated by historic and paleo-liquefaction observations in the region, and by earthquakes occurring globally each year, alluvium and other natural deposits can be equally susceptible. Still the predictions mapped in Figure 6a are also not representative of the city of Seattle as a whole. Analogous statements could be made elsewhere, given that CPTs are necessarily performed in ground that is relatively penetrable. The study sites thus tend to be located where liquefiable soils have more potential to be found.

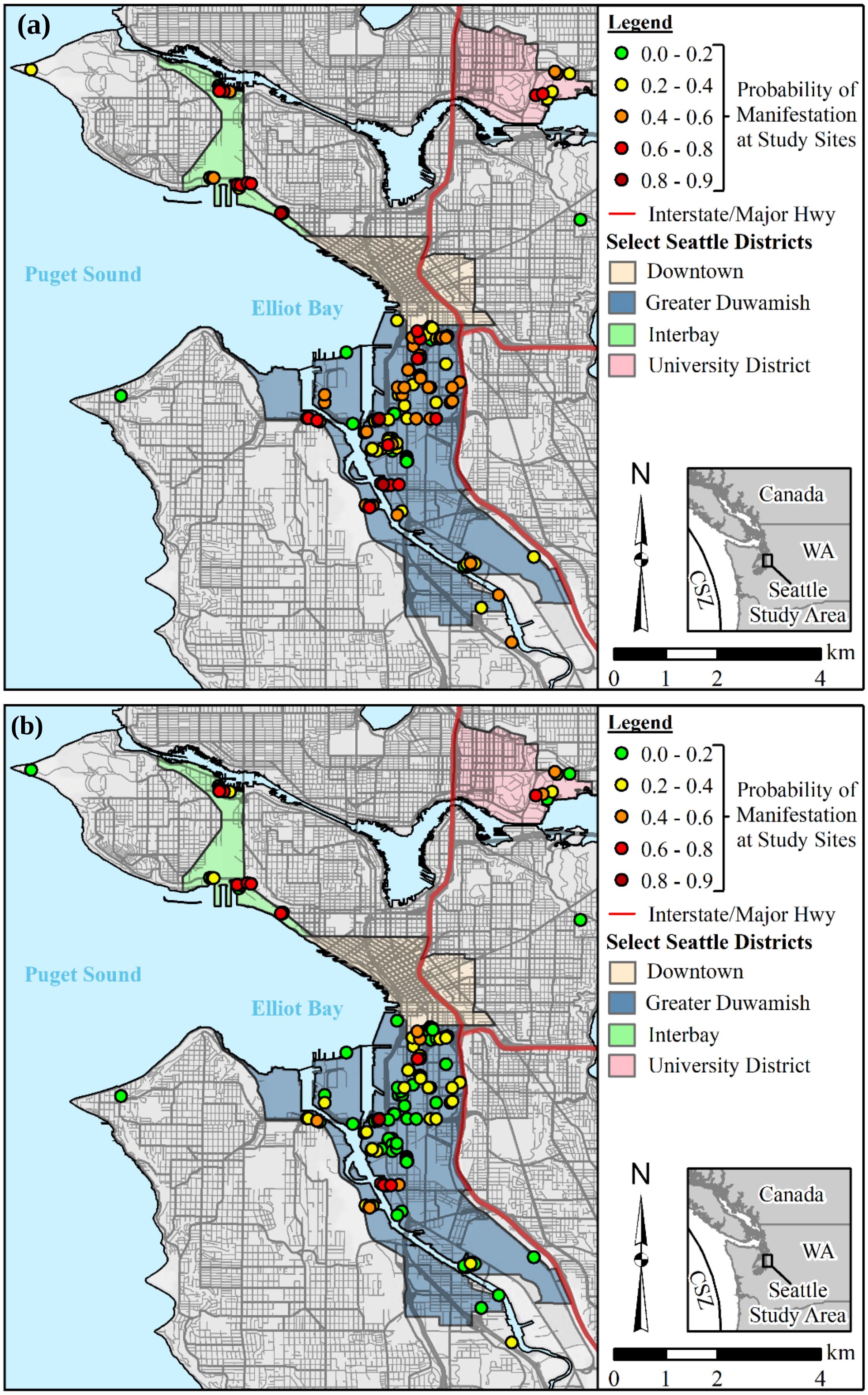


Figure 6. Probability of liquefaction surface manifestation at the **(a)** 50th and **(b)** 2.5th percentiles at study sites in Seattle, WA, considering 30 ground motion simulations.

To evaluate the distribution of expected outcomes at each study site, 2.5th and 97.5th percentile manifestation probabilities were also extracted from the analyses. The 2.5th percentile probabilities, for example, are mapped in Figure 6b for a portion of Seattle. Together, the values provide a 95% confidence interval on the expected liquefaction response and are provided for all study sites in a supplemental GIS map file, where the median predictions are also given. These maps allow readers to explore the predictions at full resolution. It should be emphasized, however, that these confidence intervals primarily reflect the uncertainty inherent to the 30 ground-motion simulations and do not include all conceivable sources of uncertainty, which could potentially broaden the probability distributions.

Although about 80% of study sites in Figure 6b have manifestation probabilities of 0-40%, the remainder have higher probabilities that sporadically exceed 70 or 80%. Comparing again against the 2001 Nisqually earthquake, the 2.5th percentile probabilities in Figure 6b are now lower in Duwamish, but still much higher in Interbay and the University District, as compared to those computed by the authors at these sites in the Nisqually event (Rasanen et al., 2023). Therefore, in the "best case" scenario (i.e., for the weakest predicted ground motions), liquefaction is still expected to occur in Seattle, although the frequency and severity of manifestations in the Duwamish neighborhood may be less than what occurred there in 2001. Of course, what may be overlooked in making such a comparison, and what is particularly problematic about M9 CSZ ruptures, is the immensity of the area likely to be affected. The 2.5th percentile probabilities mapped in Figure 6b are associated with a rupture scenario that produces relatively weak shaking in Seattle, but much stronger shaking elsewhere in the region. Thus, while the 2.5th percentile outcome is inherently less likely, it is further unlikely that this outcome would occur at all study sites in any one M9 earthquake. The sparing of one area from high seismic loading tends to mean a different area receives significantly more. In the "worst case" scenario, for example, the 97.5th percentile predictions in Seattle (provided in the appendix) suggest that liquefaction manifestations would be widespread, and often severe, at the study sites shown in Figure 6b.

The preceding analyses show that like the differences in PGA predicted across the 30 ground-motion simulations, the differences in predicted liquefaction response also often vary considerably. Therefore, it would be useful to determine which, if any, simulation or set of simulations best represent the ground motion intensities experienced in the 1700 CE rupture.

3.2 Liquefaction Predictions: Looking Backward

Using all ten well studied 1700 sites and their assigned classification uncertainties, the likelihood of each simulation to have produced the observed paleoliquefaction evidence was computed. This analysis showed csz-005 to be the most likely simulation to have produced the 1700 paleoliquefaction evidence. Simulation csz-005, shown previously in Figure 2a, predicts weaker shaking inland in the north (e.g., in Seattle, WA and Vancouver, BC) and stronger shaking inland in the south (e.g., Portland, OR). It is important to note that Figure 2a shows csz-005 prior to site-response analysis (i.e., all sites are considered to be on soft rock). The absence of site effects allows for a clearer picture of the regional prediction trends. The likelihood analysis was repeated numerous times incorporating the following changes: (i) several plausible classification uncertainties were varied to assess sensitivity; (ii) classification uncertainties were removed altogether; (iii) CPTs were (and were not) trimmed to remove fill; (iv) the ground water table was raised and lowered (e.g., by 2 meters); and (v) study sites were added and removed altogether. Regardless of ground water table, regardless of trimming, regardless of classification uncertainty (for a wide range of plausible uncertainties), and for several combinations of these, the outcome was always the same, csz-005 was the most likely to have produced the current 1700 paleoliquefaction evidence.

While simulation csz-005 fits the paleoliquefaction observations the best among the 30 simulations, it still fits the observations poorly. That is, relatively high probabilities of liquefaction manifestation are computed at sites where none was observed from the 1700 event. A simulation that would match the current paleoliquefaction evidence better would be one that does not predict liquefaction in the Puget Sound and northward. Such a simulation might be similar to csz-005, but the shaking in Washington would have to be less. To illustrate this point, Figure 7 shows the probability of ground failure for csz-005 across the region along with the ten 1700 study sites labeled. This simulation predicts widespread liquefaction in the Puget Lowland, yet no liquefaction has been tied to the 1700 event, nor to any prior CSZ rupture, in the Puget Lowland. Regarding prior CSZ interface earthquakes that predate 1700 CE, the likelihood analysis performed uses liquefaction observations tied to the 1700 earthquake, and thus the strength of shaking trends deemed most likely from the likelihood analysis are only for the 1700 earthquake and may not reflect past interface ruptures. The CSZ has ruptured many times in the past (e.g., Atwater and Hemphill-Haley 1997) and ruptures prior to 1700 may have had different rupture properties and therefore different ground motions and liquefaction fields. To elucidate the shaking in these pre-1700 ruptures, paleoliquefaction evidence tied to each specific rupture would need to be analyzed. In this regard, it is crucial for researchers to conduct ongoing field explorations to document and date new evidence for all CSZ ruptures, as well as other earthquake sources that have previously caused significant ruptures, like the Seattle fault (Atwater 1999). It is peculiar, however, that liquefaction has not been tied to any CSZ interface earthquake in the Puget Lowland or Fraser delta thus far.

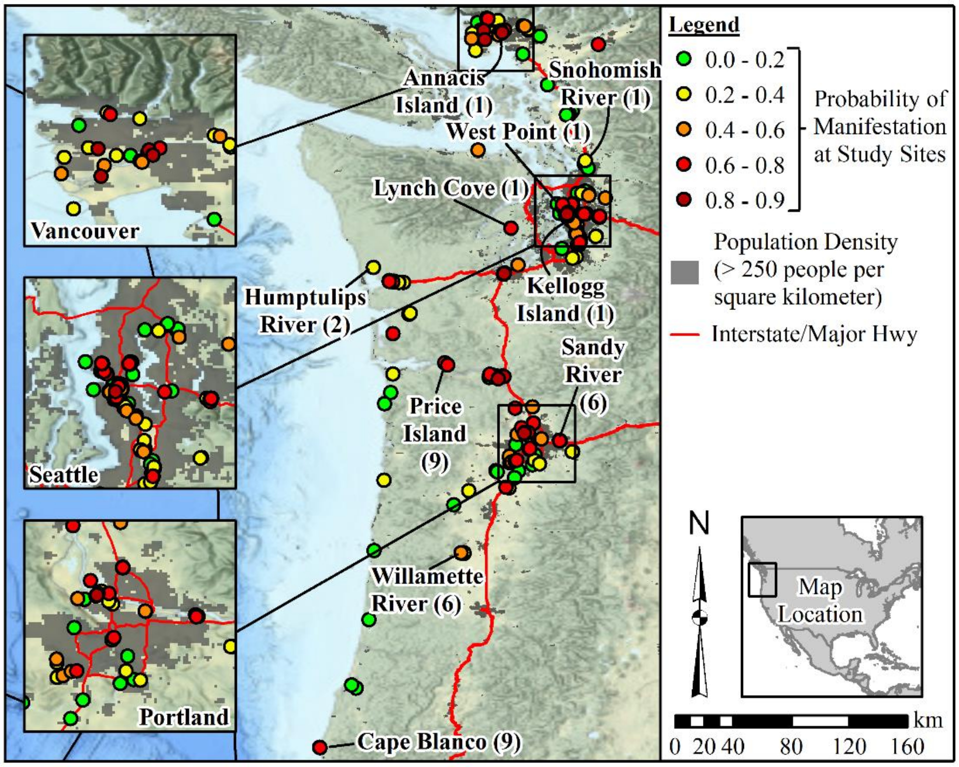


Figure 7. Probability of liquefaction surface manifestation for csz-005 at all 400 study sites with the 10 well-studied 1700 sites labeled.

There are several possible explanations for the discrepancy between predictions and observations of liquefaction at regional scale. These include: (i) there is an abundance of 1700 paleoliquefaction evidence that has not been discovered (e.g., in the Puget Lowland); (ii) component models used to predict site response and liquefaction are erroneous; (iii) the simulations overpredict PGAs in inland hubs such as Seattle, WA; and/or (iv) the 1700 CE rupture was not a full-margin rupture. The second is relatively unlikely given that the adopted techniques are trusted worldwide in engineering practice and relatively well validated (e.g., Geyin et al., 2020). It is true, however, that liquefaction models have been trained on relatively few case histories from earthquakes of M_w8 and above. These models have not been shown to be less accurate or more uncertain in such events, but the possibility persists. The first is plausible but also perplexing, given that there are many locations in the Puget Lowland and Fraser River delta where paleoliquefaction has been discovered from earthquakes that are younger and/or older than 1700 CE. This indicates the strength of ground shaking in the Puget Lowland and Fraser River delta may have been lower than the simulations predict, given the lack of paleoliquefaction evidence despite relatively rigorous efforts to find it. In this regard, the fourth possibility could explain the lack of paleoliquefaction evidence in the Puget Lowland and Fraser River delta, given that smaller M_w ruptures have shorter duration of shaking and would likely produce lesser shaking intensities in certain regions of the CSZ. Furthermore, recent studies (Melgar, 2021) have shown the 1700 CSZ rupture need not be a full-margin rupture to produce the other types of paleoseismic evidence (e.g., subsidence and tsunami) that have been tied to the 1700 rupture.

Prior to any conclusion being made, more paleoliquefaction study sites are needed. That is, sites with CSZ interface evidence (both positive and negative) and with geotechnical data (e.g., CPT data). There are many other well-studied positive and negative 1700 sites available for study, particularly along the Columbia River (e.g., Marsh, Brush, and Deer Islands amongst several others), but they require the expense of in-situ geotechnical testing such as a CPT truck. In that regard, searching for new paleoliquefaction evidence near CPT data would alleviate the costs if paleoliquefaction were to be found at those locations. Finding new evidence could significantly alter our understanding of the 1700 CE event and others (e.g., the Seattle fault rupture of 900-930 CE (Atwater 1999)). Ultimately, as more well-studied 1700 study sites with geotechnical data are developed, better constraint of regional shaking intensities will result.

3.3 Limitations, Uncertainties, and Further Discussion

While this is the most advanced study to date, which incorporated many of the best models available, there are many uncertainties that warrant further discussion than has been detailed thus far. First, 1D EQL-FD and NL site response analyses tend to become less accurate when large shear strains occur in the subsurface. Large shear strains are typically associated with strong ground shaking. The level of strain at which predicted PGAs start to differ from the “true PGA” has ranged in the literature from 0.1% (Zalachoris and Rathje 2015) to 0.4% (Kaklamonos et al. 2013). At high frequencies (e.g., PGA), the latter study states NL site response analyses are necessary when shear strains are greater than 0.4%, while Zalachoris and Rathje (2015) indicate that after 0.1% shear strain EQL-FD overpredicts and NL underpredicts ground motion intensity at high frequencies. The median peak shear strains at 400 CPT sites considering 30 ground motion simulations are shown in Figure 8. All median peak shear strains are less than 0.4% in Seattle and Portland, while only three are greater than 0.4% percent in Vancouver. Furthermore, of the sites with median peak shear strains >0.4%, none of them exceed 1% strain. Thus, while almost all sites inland have median shear strains less than 0.4%, there are some that exceed 0.4% and thus have more uncertainty in what the true PGA is, which impacts the liquefaction predictions. Importantly, if readers view the 97.5th percentile liquefaction predictions in the appendix or Map Package, the reader should be aware that there is more uncertainty with these predictions due to larger shear strains being imparted, which tend to decreases the NL PGA and thus the true 97.5th percentile liquefaction predictions could be larger than shown in the 97.5th percentile figure. To directly compare NL and EQL-FD, and to illustrate their differences, the computed geometric mean horizontal PGAs for all CPT-simulation combinations (12,000 total) are shown in Figure 9. The two methods generally agree at peak shear strains less than 0.4%, but in general as the peak shear strain increases, typically when PGAs are large, equivalent linear tends to predict higher PGAs. This discrepancy becomes especially apparent when peak shear strains are greater than 1%. Therefore, using EQL site response analysis for sites along the coast where shaking is strong (and thus where higher peak shear strains are more likely to occur) could lead to erroneous predictions. However, in inland locations such as Seattle, Portland, and Vancouver, use of EQL analysis may be adequate for determining site adjusted PGA if the peak shear strain is less than 0.4%. In addition to large shear strains, the 1D site response analyses may have more uncertainty for sites located in basins, which may have important 2D and 3D effects, potentially causing seismic waves to not approach the ground surface vertically, causing more uncertainty with 1D site response analysis. However, the Cascadia simulations account for basin effects, so to some degree this is accounted for already.

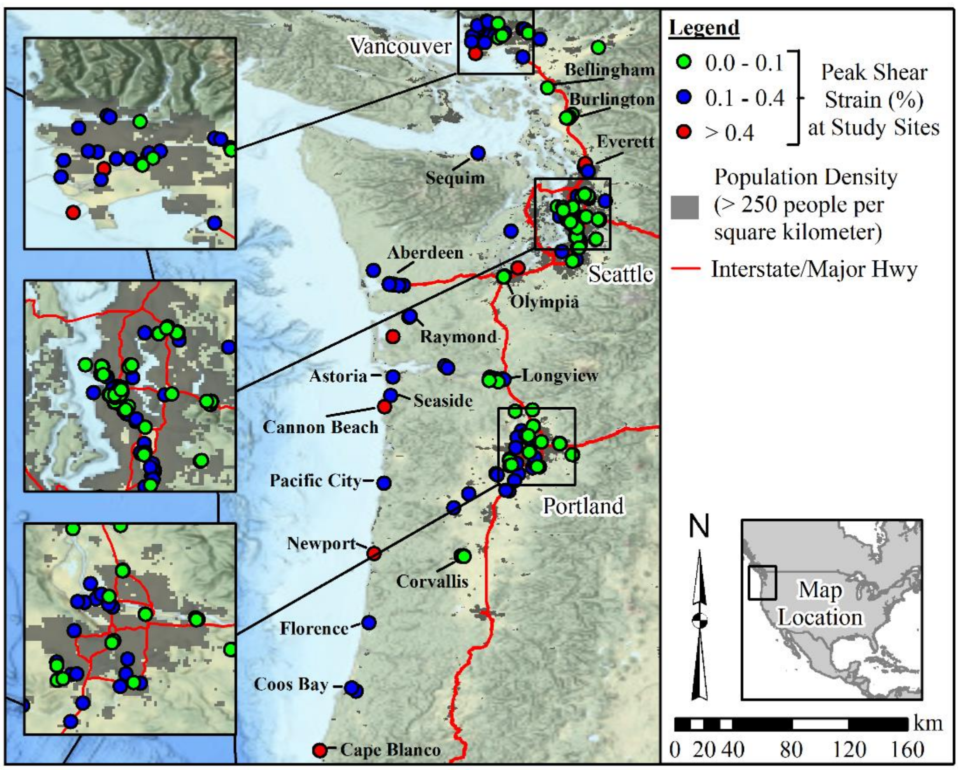


Figure 8. Median peak shear strain (%) at 400 CPT sites considering 30 ground motion simulations.

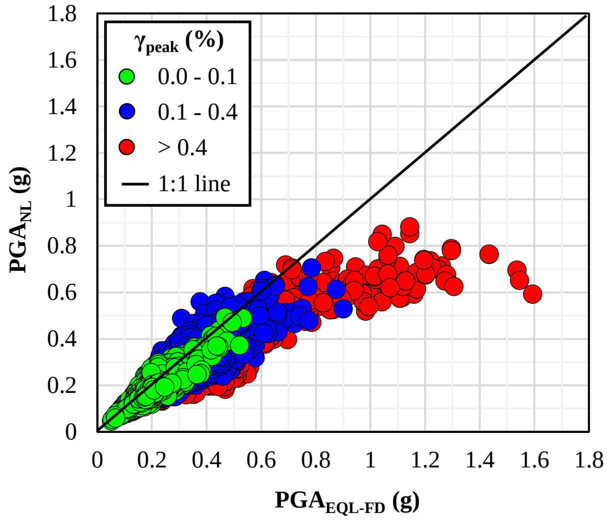


Figure 9. Comparison of PGA predicted by NL and EQL-FD site response analysis for all CPT-Simulation combinations in addition to the corresponding peak shear strain (γ_{peak}).

Second, liquefaction models have been trained on relatively few case histories from earthquakes of M_w8 and above, where in this study they were used for M9 earthquakes, which typically have longer duration and greater amplitude than those which liquefaction models were trained upon. These state-of-art liquefaction models, which are mechanistic in nature, and scale with Mw (e.g., duration), have not been shown to be less accurate or more uncertain in M9 events, but the possibility persists.

Third, given the use of CPT data, liquefaction susceptibility was determined by *Ic*. Because the *Ic* threshold at which a soil is susceptible to liquefaction is uncertain, the *Ic* – susceptibility model of Maurer et al. (2019) was used to probabilistically predict susceptibility, as defined by Boulanger and Idriss (2006), from the measured *Ic*. While this model was trained on data from New Zealand, it is the only known model of its type and its applicability elsewhere has not been disproven.

Fourth, and has been mentioned previously, liquefaction manifestation predictions are made at the locations of CPT data and while these predictions may generally extrapolate to areas surrounding each site, liquefaction is a localized phenomenon and therefore nothing can be confidently concluded at unsampled locations. In this regard, CPT data along coastal towns in the current dataset are sparse, and the locations judged unlikely to have liquefaction manifestations may not be pervasive. Liquefiable soils do exist on the coast and could exist within the coastal towns where test data is available, but additional data would be needed to assess this.

Other forms of uncertainty include: (i) cyclic softening was not considered (i.e., ground deformation could occur in plastic soils); (ii) seismic compression in unsaturated soil was not considered but should be very minor; (iii) manifestations do not consider potential for lateral spreading; (iv) and paleo comparison issues may exist. Regarding the latter, paleo comparison issues may exist between 1700 CE and the time of testing of the CPT (e.g., change in GWT, change from fill, change to penetration resistance over time, evidence is hard to find (liquefaction may not manifest directly above its source), etc.). The uncertainties in the latter were accounted for in the likelihood analysis by iterating through numerous scenarios where these parameters were changed to other plausible values as discussed previously.

CONCLUSIONS

This study led to: (i) the first regional scale M9 CSZ liquefaction assessment using geotechnical data; and (ii) constraint of the intensity of ground shaking from the 1700 CE earthquake, to the degree permitted by current paleoliquefaction evidence. Looking forward, the predictions at 400 study sites across the PNW suggest that liquefaction could be both pervasive in the region and affect numerous significant hubs of population, transportation, and commerce, with the potential for damage across hundreds of square kilometers. Looking backward, the analyses suggest that 1700 ground motion intensities may have been less in northern portions of the CSZ (e.g., Seattle and Vancouver) than prevailing forecasts indicate, given the simulations predict far more liquefaction than has been found in the Puget Lowland and Fraser delta and the peculiar paucity of 1700 evidence throughout these areas. This finding has important nuances and uncertainties that were discussed in detail, and which warrant further investigation. Ultimately, further discovery and analysis of CSZ paleoliquefaction, or lack thereof, will confirm or modify the conclusions drawn herein.

ACKNOWLEDGEMENTS

The presented study is based on work supported by the National Science Foundation (NSF) under Grant No. CMMI-1751216 and by the NSF Graduate Research Fellowship Program under Grant No. DGE-1762114. However, any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

REFERENCES

Andrus RD, Mohanan NP, Piratheepan P, Ellis BS, Holzer TL (2007). Predicting shear-wave velocity from cone penetration resistance. In: *Proceedings of the 4th international conference on earthquake geotechnical engineering*, vol. 2528. Thessaloniki, Greece.

Architectural Institute of Japan (2001). *Recommendations for design of building foundations*, 486 p.

Asimaki, D. and Shi, J. (2017). "SeismoSoil User Manual, v1.3" doi.org/10.5281/zenodo.7813062.

Assaf, J., Molnar, S., & El Naggar, M. H. (2023). CPT-Vs correlations for post-glacial sediments in Metropolitan Vancouver. *Soil Dynamics and Earthquake Engineering*, *165*, 107693.

Assimaki, D., and E. Kausel (2002). "An equivalent linear algorithm with frequency- and pressure-dependent moduli and damping for the seismic analysis of deep sites." *Soil Dynamics and Earthquake Engineering*, 22(9–12), 959–965.

Atwater, B.F. (compiler) (1994). “Geology of Holocene liquefaction features along the lower Columbia River at Marsh, Brush, Price, Hunting, and Wallace Islands, Oregon and Washington.” *U.S. Geol. Surv. OpenFile Rept.* 94-209, 30 pp.

Atwater, B.F. (1999). “Radiocarbon dating of a Seattle earthquake to A.D. 900–930.” *SRL* 70: 232.

Atwater, B.F., Carson, B., Griggs, G.B., Johnson, H.P., and Salmi, M.S. (2014). “Rethinking turbidite paleoseismology along the Cascadia subduction zone.” *Geology*, 42(9): 827-830.

Atwater, B.F., and Hemphill-Haley, E. (1997). “Recurrence Intervals for Great Earthquakes of the Past 3,500 Years at Northeastern Willapa Bay, Washington.” *U.S. Geological Survey Paper 1576.*

Atwater, B.F., Stuiver, M., and Yamaguchi, D.K. (1991). “Radiocarbon test of earthquake magnitude at the Cascadia subduction zone.” *Nature*, 353: 156-158.

Boulanger, R.W. and Idriss, I.M. (2006). “Liquefaction susceptibility criteria for silts and clays.” *Journal of Geotechnical and Geoenvironmental Engineering* 132(11): 1413-1426.

Boulanger, R. W. and Idriss, I. M. (2014). “CPT and SPT based liquefaction triggering procedures.” Report No. UCD/CGM-14/01, Center for Geotechnical Modeling, Department of Civil and Environmental Engineering, University of California, Davis, CA, 134 pp.

Bourgeois, J., and Johnson, S.Y. (2001). “Geologic evidence of earthquakes at the Snohomish delta, Washington, in the past 1200 yr.” *Geological Society of America Bulletin* 113(4): 482-494.

Bucknam, R.C., Hemphill-Haley, E. and Leopold, E.B. (1992) Abrupt uplift within the past 1700 years at southern Puget Sound, Washington. *Science*, 258, 1611–1614.

Chin, T., and Evans, T.M. (2021). “GOSEP: Georeferenced Oregon Soil Engineering Properties.” Ver. 2.0, Accessed from http://ohelp.oregonstate.edu/GOSEP/user_manual.pdf on 6 June 2023.

Chleborad AF and Schuster RL (1990) Ground failure associated with the Puget Sound region earthquakes of April 13, 1949 and April 29, 1965. United States Geological Survey Open-File Rept.

Clague, J.J., Naesgaard, E., and Sy, A. (1992). “Liquefaction features on the Fraser delta: evidence for prehistoric earthquakes?” Canadian Journal of Earth Sciences 29: 1734.1745.

Clague, J.J., Naesgaard, E., and Nelson, A.R. (1997). “Age and significance of earthquake-induced liquefaction near Vancouver, British Columbia, Canada.” Can. Geotech. J. 34: 53-62.

Davis E. (2018). “Evidence for liquefaction and flooding in the past 1,000 years along the Duwamish River, Seattle, Washington.” M.S. thesis, University of Washington.

Davis E. (2019). Seattle liquefaction features along the Duwamish Waterway, Washington. In: Seismological Society of America Annual Meeting, 23–26 April, Seattle, USA.

DHS (2019). Department of Homeland Security. “Resiliency Assessment: Washington State Transportation Systems,” Regional Resiliency Assessment Program (RRAP), March 2019.

Engelhart, S.E., Horton, B.P., Nelson, A.R., Hawkes, A.D., Witter, R.C., Wang, K., Wang, P.-L., and Vane, C.H. (2013). “Testing the use of microfossils to reconstruct great earthquakes at Cascadia.” *Geology* 41 (10): 1067–1070.

Feliks, B. (2019). "All Over the Map: Kellogg Island is a 19th-century time capsule."

FEMA (2019). “Hazus Earthquake Model Technical Manual.” Federal Emergency Management Agency, Washington D.C.

Frankel, A., Wirth, E., Marafi, N., Vidale, J., and Stephenson, W. (2018). “Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust Based On 3D Simulations and Stochastic Synthetics (Part 1): Methodology and Overall Results.” *BSSA* 108 (5A): 2347-2369.

Geyin, M. and Maurer, B.W. (2021a) “Fragility functions for liquefaction induced ground failure.” *Journal of Geotechnical and Geoenvironmental Engineering*, 146(12): 04020142.

Geyin, M., and Maurer, B. W. (2021b). Evaluation of a cone penetration test thin-layer correction procedure in the context of global liquefaction model performance. *Engineering Geology*, *291*, 106221.

Geyin, M., and Maurer, B. W. (2023). US National VS30 Models and Maps Informed by Remote Sensing and Machine Learning. *Seismological Research Letters*, 94(3): 1467-1477.

Geyin, M., Baird, A. J., & Maurer, B. W. (2020). “Field assessment of liquefaction prediction models based on geotechnical versus geospatial data, with lessons for each.” *Earthquake Spectra*, *36*(3), 1386-1411.

Goldfinger, C., Nelson, C.H., Morey, A.E., Johnson, J.E., Patton, J.R., Karabanov, E., Gutiérrez-Pastor, J., Eriksson, A.T., Gràcia, E., Dunhill, G., Enkin, R.J., Dallimore, A., and Vallier, T., (2012). “Turbidite event history - methods and implications for Holocene paleoseismicity of the Cascadia subduction zone.” *U.S. Geological Survey Professional Paper 1661–F*, 170 pg.

Green, R. A., J. J. Bommer, A. Rodriguez-Marek, B. W. Maurer, P. J. Stafford, B. Edwards, P. P. Kruiver, G. De Lange, and J. Van Elk (2019). "Addressing limitations in existing ‘simplified’ liquefaction triggering evaluation procedures: application to induced seismicity in the Groningen gas field." *Bulletin of Earthquake Engineering* 17(8), 4539-4557.

Hemphill-Haley, E. (1996) Diatoms as an aid in identifying late-Holocene tsunami deposits. *The Holocene*, 6, 439–448.

Idriss, I.M., and Boulanger, R.W. (2008). “Soil liquefaction during earthquakes.” *Monograph MNO-12* 2008; Earthquake Engineering Research Institute, Oakland, CA, 261 pp.

Ishihara, K. (1985). “Stability of natural deposits during earthquakes.” *Proceedings of the 11th international conference on soil mechanics and foundation engineering* 1: 321–76.

Iwasaki, T, Tatsuoka, F, Tokida, K, and Yasuda, S.A. (1978). “Practical method for assessing soil liquefaction potential based on case studies at various sites in Japan.” *2nd Int. Conf. on Microzonation*, San Francisco, USA.

Jacoby, G. C., D. E. Bunker, and B. E. Benson (1997). Tree-ring evidence for an A.D. 1700 Cascadia earthquake in Washington and northern Oregon, *Geology* 25, 999–1002.

Jeschke, D.A., Eungard, D.W., Troost, K.G., Wisher, A.P, (2019). “Subsurface database of Washington State - GIS data.” Washington Geological Survey Digital Data Series 11, version 2.1, previously released August 2017, Accessed from https://geologyportal.dnr.wa.gov/ on 6 June 2023.

Joseph W. Berg, Charles D. Baker (1963). Oregon earthquakes, 1841 through 1958. *Bulletin of the Seismological Society of America*; 53 (1): 95–108. doi: https://doi.org/10.1785/BSSA0530010095

Jovanelly, T.J. and Moore, A.L. (2009) Sedimentological analysis of an ancient sand sheet of multiple origins at Lynch Cove, Puget Sound, Washington. *J. Coast. Res.*, 25, 294–304.

Kaklamanos, J., Bradley, B. A., Thompson, E. M., & Baise, L. G. (2013). Critical parameters affecting bias and variability in site-response analyses using KiK-net downhole array data. Bulletin of the Seismological Society of America, 103(3), 1733-1749.

Kelsey, H.M., Witter, R.C., and Hemphill-Haley, E. (1998). Response of a small Oregon estuary to coseismic subsidence and postseismic uplift in the past 300 yr: *Geology*, v. 26, p. 231–234.

Kelsey, H.M., Witter, R.C., Hemphill-Haley, E. (2002). “Plate-boundary earthquakes and tsunamis of the past 5500 yr, Sixes River estuary, southern Oregon. *GSA Bulletin* 114(3): 298-314.

Li, W., and D. Assimaki (2010). “Site- and motion-dependent parametric uncertainty of site-response analyses in earthquake simulations.” *Bulletin of the Seismological Society of America*, 100(3), 954–968.

Liu, P.-C., and R. J. Archuleta (2006). “Efficient modeling of Q for 3D numerical simulation of wave propagation.” *Bulletin of the Seismological Society of America*, 96(4A), 1352–1358.

Marafi, N. (2018). “M9 Simulations.” Personal Communication; available from University of Washington News https://www.flickr.com/photos/uwnews/sets/72157688124103084/.

Marafi, N. A., Grant, A., Maurer, B. W., Rateria, G., Eberhard, M. O., & Berman, J. W. (2021). “A generic soil velocity model that accounts for near-surface conditions and deeper geologic structure.” *Soil Dynamics and Earthquake Engineering*, *140*, 106461.

Martin, M.A., and Bourgeois, J. (2012). “Vented sediments and tsunami deposits in the Puget Lowland, Washington – differentiating sedimentary processes.” *Sedimentology* 59: 419-444.

Mathewes, R.W., and Clague, J.J. 1994. Detection of prehistoric large earthquakes in the Pacific Northwest by microfossil analysis. Science (Washington, D.C.), 264: 688.691.

Maurer, B.W. and Sanger, M.D. (2023). “Why ‘AI’ Models for Predicting Liquefaction have been Ignored, plus some that Shouldn’t Be.” *Earthquake Spectra*, doi: 10.1177/87552930231173711.

Maurer, B.W., Green, R.A. and Taylor, O.D.S. (2015). “Moving towards an improved index for assessing liquefaction hazard: lessons from historical data.” *Soils and Foundations* 55(4): 778-787.

Maurer, B.W., Green, R.A., van Ballegooy, S. and Wotherspoon, L. (2019). “Development of region-specific soil behavior type index correlations for evaluating liquefaction hazard in Christchurch, New Zealand.” *Soil Dynamics and Earthquake Engineering* 117: 96-105.

McGann, C. R., Bradley, B. A., Taylor, M. L., Wotherspoon, L. M., & Cubrinovski, M. (2015). Development of an empirical correlation for predicting shear wave velocity of Christchurch soils from cone penetration test data. *Soil Dynamics and Earthquake Engineering*, 75, 66-75.

Moss R.E.S., Seed R.B., Kayen R.E., Stewart J.P., Der Kiureghian A., Cetin K.O. (2006). “CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential.” *Journal of Geotechnical and Geoenvironmental Engineering*, 132(8), 1032-1051.

NAE (2021). State of the art and practice in the assessment of earthquake-induced soil liquefaction and its consequences. National Academy of Engineering, Washington, D.C.

O’Donnell III, R.J., Hawkes, A.D., Lane, C.S., Engelhard, S.E., Horton, B.P., Bobrowsky, P., Sawai, Y., Witter, R.C., Nelson, A.R., and Tanigawa, K. (2017). “Assessing the utility of δ13C and bulk geochemistry in estuaries along the Cascadia subduction zone for coastal paleoseismology.” *Geological Society of America Abstracts with Programs* 49(6): 58-2.

Obermeier, S. F. (1995). “Preliminary estimates of the strength of prehistoric shaking in the Columbia River Valley and the southern half of coastal Washington, with emphasis for a Cascadia Subduction Zone earthquake about 300 years ago.” *U.S. Geol. Surv.* Open-File Rept. 94-589, 46 pp.

Obermeier, S.F. and Dickenson, S.E. (2000). “Liquefaction evidence for the strength of ground motions resulting from late Holocene Cascadia subduction earthquakes, with emphasis on the event of 1700 A.D.” *Bulletin of the Seismological Society of America*, 90(4): 876-896.

Parker, M and Steenkamp, D. (2012). “The economic impact of the Canterbury earthquakes.” Bulletin of the *Reserve Bank of New Zealand:* 75(3): 13–25.

Peters, R., Jaffe, B., and Gelfenbaum, G. (2007). “Distribution and sedimentary characteristics of tsunami deposits along the Cascadia margin of western North America.” *Sed Geo* 200: 372-386.

Peterson, C.D., Kristensen, K., and Minor, R. (2014). “Large-Scale Fluidization Features from Late Holocene Coseismic Paleoliquefaction in the Willamette River Forearc Valley, Central Cascadia Subduction Zone, Oregon, USA.” *Open Journal of Earthquake Research*, 3: 82-99

Peterson C.D. and Madin, I.P. (1997). “Coseismic Paleoliquefaction Evidence in the Central Cascadia Margin, USA.” Oregon Geology 59, 51-74.

Rasanen, R. A., & Maurer, B. W. (2022). “Probabilistic seismic source location and magnitude via inverse analysis of paleoliquefaction evidence.” *Earthquake Spectra*, *38*(2), 1499-1528.

Rasanen, R. A., & Maurer, B. W. (2022a). “Seismic Source Parameters from Regional Paleoseismic Evidence”. In *Geo-Congress 2022* (pp. 411-420).

Rasanen, R. A., & Maurer, B. W. (2023). Probabilistic seismic source inversion of the 1886 Charleston, South Carolina, earthquake from macroseismic evidence: A major updating. Engineering Geology, 312, 106958.

Rasanen, R. A., Marafi, N. A., & Maurer, B. W. (2021a). “Compilation and forecasting of paleoliquefaction evidence for the strength of ground motions in the US Pacific Northwest.” *Engineering Geology*, *292*, 106253.

Rasanen, R.A., Geyin, M., and Maurer, B.W. (2023). “Select Liquefaction Case Histories from the 2001 Nisqually, Washington Earthquake: A Digital Dataset and Assessment of Model Performance.” *Earthquake Spectra*, doi: 10.1177/87552930231174244.

Rasanen, R.A., Wood, C.M., and Maurer, B.W. (2024). "Constraining Cascadia Subduction Zone Ground Motions via Paleoliquefaction Evidence: A Case Study from Kellogg Island, Washington, with Regional Implications." In: *Proceedings of GeoCongress 2024*, Submitted.

Rashidian, V., & Baise, L. G. (2020). “Regional efficacy of a global geospatial liquefaction model.” *Engineering geology, 272*, 105644.

Robertson, P. K. (2009). Interpretation of cone penetration tests—a unified approach. *Canadian geotechnical journal*, 46(11), 1337-1355.

Robertson, P.K., and Wride, C.E. (1998). "Evaluating cyclic liquefaction potential using cone penetration test." *Canadian Geotechnical Journal*, 35(3), 442-459.

Satake, K., K. Shimazaki, Y. Tsuji, and K. Ueda (1996). Time and size of a giant earthquake in Cascadia inferred from Japanese tsunami record of January 1700, *Nature* 379, 246–249.

Seed, H.B. & Idriss, I.M., 1971. Simplified procedure for evaluating soil liquefaction potential, *ASCE Journal of Soil Mechanics and Foundations Division* 97(9), 1249–1273.

Shi, J., & Asimaki, D. (2017). "From Stiffness to Strength: Formulation and Validation of a Hybrid Hyperbolic Nonlinear Soil Model for Site-Response Analyses." *Bulletin of the Seismological Society of America*, *107*(3), 1336–1355.

Sideras, S. (2019). *Evolutionary intensity measures for more accurate and informative evaluation of liquefaction triggering.* PhD Dissertation, University of Washington, Seattle, WA

Siskowic, J.D., Anderson, D.A., Peterson, B., Peterson, C.D., Soar, M., Travis, P, and Vockler, K., 1994, Possible coseismic liquefaction evidence at the Sandy River delta, Portland, Oregon: Tentative correlation with the last great Cascadia rupture [abs.]: Geological Society of America Abstracts with Programs, v. 26, no. 2, p. 92.

Stephenson, W. J., Reitman, N. G., & Angster, S. J. (2017). P- and S-wave velocity models incorporating the Cascadia subduction zone for 3D earthquake ground motion simulations—Update for Open-File Report 2007–1348. In Open-File Report (Version 1.). https://doi.org/10.3133/ofr20171152.

Thrush, C., and Ludwin, R.S. (2007). "Finding fault: Indigenous seismology, colonial science, and the rediscovery of earthquakes and tsunamis in Cascadia." *American Indian Culture and Res J* 31(4): 1-24.

Van Ballegooy, S., Malan, P., Lacrosse, V., Jacka, M.E., Cubrinovski, M., Bray, J.D., O'Rourke, T.D., Crawford, S.A., and Cowan, H. (2014). "Assessment of liquefaction-induced land damage for residential Christchurch." *Earthquake Spectra*, 30(1): 31-55.

Watt, J.T., Brothers, D.A., Bennett, S.E.K., Kluesner, J.W., Roland, E., Conrad, J.E., Sliter, R.W., Dartnell, P., Goldfinger, C., Patton, J., Michalak, M.J., Sherrod, B.L., Gomberg, J., and Wells, R.E. (2017). "Toward a systematic characterization of Cascadia upper plate morphology, structure, and Quaternary deformation history: an integrated onshore-offshore approach." *GSA Abstracts* 49(6): 58.

Wirth, E. A., Frankel, A. D., Marafi, N., Vidale, J. E., and Stephenson, W. J. (2018). "Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust based on 3-D Simulations and Stochastic Synthetics (Part 2): Rupture Parameters and Variability." *BSSA*, 108 (5A): 2370-2388.

Wirth, E. A., Grant, A., Marafi, N. A., & Frankel, A. D. (2021). Ensemble shakemaps for magnitude 9 earthquakes on the Cascadia subduction zone. *Seismological Research Letters*, *92*(1), 199-211.

WSDOT (2021). "Geotech Data Historic Borings." Washington State Department of Transportation (WSDOT), WSDOT Online Map Center. Accessed from https://geo.wa.gov/ on 6 June 2023.

Wu, Q., Li, D.Q., and Du, W. (2022). "Identification of optimal ground-motion intensity measures for assessing liquefaction triggering and lateral displacement of liquefiable sloping grounds." *Earthquake Spectra*: 87552930221094344.

Yamaguchi, D. K., B. F. Atwater, D. E. Bunker, B. E. Benson, and M. S. Reid (1997). Tree-ring dating the 1700 Cascadia earthquake, *Nature* 389, 922–923 (correction 390, 352).

Yost, K. M., Green, R. A., Upadhyaya, S., Maurer, B. W., Yerro-Colom, A., Martin, E. R., & Cooper, J. (2021). Assessment of the efficacies of correction procedures for multiple thin layer effects on Cone Penetration Tests. *Soil Dynamics and Earthquake Engineering*, *144*, 106677.

Youd, T.L., and S.K. Noble (1997). "Liquefaction criteria based on probability analysis." *Proc. of the NCEER Workshop on Evaluation of Liquefaction Resistance of Soils*, T. L. Youd and I. M. Idriss (Editors), Technical Report NCEER 97-0022, 201–216.

VITA

The author was born on April 23, 1992, in Seattle, Washington. His family moved for work to beautiful, although at times frozen, Minot, North Dakota where he grew up and then attended college at the University of North Dakota in Grand Forks, North Dakota – the coldest city in contiguous United States. It was during his time as a garbage collector that he had a firsthand experience with geologic hazards. In June 2011, when the levees broke due to the Souris River flood in Minot, only the garbage collectors and National Guard were in the flood zone. This experience motivated him to pursue studies in Geological Engineering, and he graduated summa cum laude with a 4.0 GPA in December 2014.

After graduation, he found his way back to the Pacific Northwest, gaining consulting experience for two years in slope stability analyses and geotechnical field investigations. Building upon this practical knowledge of geomaterials, he enrolled at the University of Washington (UW) to pursue an M.S. degree in Civil (Geotechnical) Engineering. While at UW, he sought out research with Dr. Steven Kramer on probabilistic liquefaction modelling. Through this experience, he has studied the body of knowledge on liquefaction susceptibility, triggering, and effects. He also has conducted sensitivity analyses on the leading probabilistic liquefaction models to analyze their performance for differing site conditions, with the ultimate goal of formulating an improved model. The author graduated with a Masters degree in Civil (Geotechnical) Engineering from the University of Washington in June 2019. The knowledge of probabilistic liquefaction triggering models was essential for formulating the mathematical tools to perform inverse-analysis of ancient liquefaction evidence, which was the main focus of the author's PhD as detailed in the dissertation.

The author graduated with his PhD in Civil (Geotechnical) Engineering in June 2023 under the wing of Dr. Brett Maurer, who was instrumental in all aspects throughout the author's PhD.

While at UW, he held a leadership role as the President of the Geo-Institute Graduate Student Society (GIGSS). In this position, he organized multiple regional graduate student conferences, facilitating collaboration and knowledge sharing among geotechnical engineering students from various universities in the Pacific Northwest. Additionally, he coordinated numerous geotechnical lectures, and developed strong connections with the Seattle ASCE Geotechnical Group. To further engage the community, he initiated a monthly lecture series where practicing professionals from local geotechnical firms presented their work and connected with graduate students. Through these efforts, he raised over $15,000 in funds to support these events. Furthermore, he mentored GIGSS members, ensuring a smooth transition for future leaders and sustaining the impact of the organization in the community and the geotechnical profession.

As the author contemplates "what's next" - the question all PhD students face – he is considering academic, government, and consulting positions. However, above all, the author's primary aspiration is to make a strong positive impact on the community in which he lives.