© Copyright 2021

Mertcan Geyin

Evaluation and Development of Liquefaction Occurrence and Consequence Analytics Driven by Emerging Data and Technologies

Mertcan Geyin

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Brett Maurer, Chair

Pedro Arduino

David Schmidt

Joseph Wartman

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Abstract

Evaluation and Development of Liquefaction Occurrence and Consequence Analytics Driven by Emerging Data and Technologies

Mertcan Geyin

Chair of the Supervisory Committee:
Brett Maurer, Assistant Professor
Civil and Environmental Engineering

For 50 years, significant investments have been made to develop models that predict soil liquefaction. Many of the results may be described as either: (i) wholly empirical models that require only geologic or geospatial data, and which are accessible to a broad userbase; or (ii) semi-mechanistic "simplified stress-based" models that are based on in-situ tests, and which are generally limited to use by geoengineers, with cone-penetration-test (CPT) based models currently considered best. While these models are widely used, all to-date were trained on modest datasets (e.g., 250 datapoints) and in general, have not been rigorously tested against unbiased data, or against one other. In addition, relative to predicting liquefaction occurrence, less effort has historically been given to predicting the downstream consequences of liquefaction, such as damage and loss. As a result, decisions about whether, and how, to mitigate liquefaction must generally be made in a heuristic manner. Meanwhile, recent earthquakes in the age of improved satellite remote

sensing and reconnaissance have significantly grown the data that could be used to model liquefaction and its effects, including ground settlement, foundation deformation, and insured losses. Moreover, the growth of remote sensing, in combination with emergent machine- and deep-learning algorithms, creates an opportunity to develop a new type of model that provides rapid, regional predictions of seismic impacts without the need to sample individual sites. Thus, the overarching goal of this work is to test and develop liquefaction occurrence and consequence analytics using emerging data and technologies. Namely, through ten objectives, this research aims to: (1) digitize and publicly compile all existing liquefaction case-histories where CPTs were performed, of which there are ~275, and which may be used to train or test CPT-based liquefaction models; and (2) grow the quantity of such cases to ~15,000 using data from earthquakes in New Zealand. These datasets will then be used to: (3) test existing liquefaction-prediction models based on either the CPT or on geospatial data, from which lessons for future improvement are identified; (4) test a new procedure for correcting "thin layer effects" on CPT data in the context of liquefaction model performance; (5) test existing CPT models that predict ground settlement induced by liquefaction; and (6) train new fragility functions that probabilistically predict liquefaction-induced ground failure. Using the data compiled in (2), in conjunction with additional foundation damage surveys and insurance loss assessments from New Zealand, a decision framework for mitigating liquefaction beneath buildings on shallow foundations is (7) proposed; and demonstrated. Using insights from the tests in (3), a next-generation geospatial modelling approach, wherein machine- and deep-learning are used to predict below-ground CPT data via above-ground geospatial information, is (8) proposed, trained, tested, and (9) coded into software. Lastly, existing and newly proposed analytics in (1-7) are (10) programmed into *Horizon*, a second software program for analyzing the occurrence, consequence, and mitigation of liquefaction.

TABLE OF CONTENTS

Chapter 1. Introduction 1

1.1 Background And Motivations 1

1.2 Dissertation Structure and Contents 6

1.3 A Succinct Summary of the Ensuing Chapters 8

Chapter 2. COMPILATION OF LIQUEFACTION CASE HISTORIES 10

2.1 Publication 1: Geyin et al. (2021). "CPT-Based Liquefaction Case Histories Compiled from Three Earthquakes in Canterbury, New Zealand." *Earthquake Spectra.* doi: 10.1177/8755293021996367. 10

2.2 Publication 2: Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI*. https://doi.org/10.17603/ 12

Chapter 3. TESTING AND DEVELOPING MODELS TO PREDICT LIQUEFACTION AND ITS SURFACE MANIFESTATION 14

3.1 Publication 3: 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. DOI: 10.1177/8755293019899951. 14

3.2 Publication 4: Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *ASCE Journal of Geotechnical and Geoenvironmental Engineering* 146 (12): 04020142. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002416. 17

3.3 Publication 5: Geyin M and Maurer BW (2019). “An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions.” *ASCE Geotechnical Special Publication No. 308*: 489–498..... 20

3.4 Publication 6: Geyin and Maurer (2020), “Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance.” *Engineering Geology*, 291, DOI: https://doi.org/10.1016/j.enggeo.2021.106221. 22

3.5 Publication 7: Geyin M, et al. (2021). “An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Scenario Planning.” *Soil Dynamics and Earthquake Engineering*, **In review**. 24

3.6 Publication 8: Geyin and Maurer. (2021). “RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction”, DesignSafe-CI. DOI: https://doi.org/10.17603/ds2-4bka-y039. 29

Chapter 4. A PRAGMATIC DECISION TOOL FOR LIFECYCLE COST-BENEFIT ANALYSIS OF LIQUEFACTION MITIGATION BENEATH LIGHTWEIGHT STRUCTURES. 31

4.1 Publication 9: Geyin M, Maurer, BW, and van Ballegooy S (2020). “Lifecycle Liquefaction Hazard Assessment and Mitigation.” *ASCE Geotechnical Special Publication* No. 320: 312-320 https://doi.org/10.1061/9780784482834.034. 32

Chapter 5. HORIZON: LIQUEFACTION OCCURRENCE AND CONSEQUENCE MODELLING SOFTWARE 40

5.1 Publication 10: Geyin M and Maurer BW (2021). “Horizon: CPT-Based Liquefaction Risk Assessment and Decision Software.” DesignSafe CI. doi: 10.17603/ds2-2fky-tm46 40

Bibliography 45

Appendix A 49

Appendix B 98

Appendix C 129

Appendix D 175

Appendix E 218

Appendix F 238

Appendix G 283

Appendix H 326

Appendix I 338

Appendix J 355

ACKNOWLEDGEMENTS

I would like to thank my supervisor Brett Maurer for all their help, patience and advice throughout my PhD. Thank you for your continuous support and tutelage. This work would not be possible without their encouragement and wisdom.

I also want to thank all my friends for their support, love and help during the course of my PhD.

Most of all, I want to thank my family for their love and support. I am forever indebted to their efforts at every stage of my life.

DEDICATION

Hasan dedeme

Chapter 1. INTRODUCTION

This dissertation follows the "manuscript format," consisting of ten publications that include datasets, software, conference papers, and journal articles. This introductory chapter first discusses the background and motivations behind the collective body of work presented herein. Next, the publications comprising this 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 two through five, each of the ten publication's principal products and findings are summarized in an expanded, yet still brief format. All publications are provided in separate appendices in-full.

1.1 BACKGROUND AND MOTIVATIONS

Earthquake-induced soil liquefaction is responsible for damaging a wide range of infrastructure assets, including dams and levees; buried lifelines; port facilities; residential, commercial, and industrial structures; and transportation systems, among others (Figure 1). These effects were vividly displayed, for example, during earthquakes that affected Canterbury, New Zealand, from 2010 to 2016 and resulted in at least ~$10B in infrastructure loss (Parker and Steenkamp, 2012). These and other recent earthquakes have demonstrated significant failings to predict and mitigate liquefaction and its consequences, while simultaneously generating truly unprecedented datasets that could be used to test or train liquefaction analytics (i.e., models for predicting liquefaction and its various effects). The Canterbury earthquakes, for example, affected a major urban area in an age of rapidly improved satellite remote sensing and reconnaissance. This facilitated and motivated the collection of vast amounts of data, including seismologic, hydrologic, geospatial, and geotechnical measurements, as well as LIDAR surveys of ground settlement, detailed inspections

of foundation damage, and pre- and post-earthquake insurance assessments (due to the government's status as a land insurer). Some of these data were uploaded to the Canterbury Geotechnical Database (CERA, 2013), now the New Zealand Geotechnical Database (NZGD, 2020). Others were obtained by the research team through international partners in academia and government. Collectively, these bulk, raw ingredients constitute the makings of an unprecedented quantity of case histories that could be used to better understand and mitigate liquefaction. Some of these ingredients, such as the foundation damage-surveys and insurance loss assessments, have not previously been available to researchers in any quantity, from any earthquakes before or since.

Figure 1. Infrastructure damage due to liquefaction-induced ground failure during recent earthquakes; photos courtesy of NSF Geotechnical Extreme Events and Reconnaissance (GEER).

While several "tiers" of liquefaction prediction model exist, most prevalent models in practice are based on in-situ geotechnical tests, among which the cone-penetration-test (CPT) has important advantages (NRC, 2016). Yet, while such models are widely used, all to-date were trained on modest datasets. For example, the CPT-based liquefaction "triggering" models of Boulanger and Idriss (2014) and Green et al. (2019), when developed, were trained on essentially all published case histories from all earthquakes combined, or 255 datapoints. In general, these and other models proposed over the last 20 years have not been tested against unbiased data, or against one another. The question of "which performs best?" thus remains popular and contentious. Also, while the CPT has advantages over other in-situ tests, it is still potentially obfuscated by the volume of soil

mobilized around the cone. This volume acts a low-pass filter on the true stratigraphy, potentially removing information that defines a thin soil stratum or the interface between two disparate soils. These smoothing effects have long been recognized and studied, but the first programmable model to correct them was proposed by Boulanger and DeJong (2018). While this model has been adopted into popular software that interprets CPT data and predicts liquefaction, it has yet to be tested in the context of liquefaction model performance, owing to a lack of case-history data.

In addition, novel "geospatial" liquefaction models, which aim to predict liquefaction using readily available geospatial variables, were recently proposed (Zhu et al., 2017). Examples of such variables include the ground slope and distance to a river, which may correlate to subsurface traits. In turn, combinations of these variables and seismic intensity data can be used to predict liquefaction regionally. These models are particularly suited for applications in which time- and cost-considerations outweigh the required, expected accuracy (e.g., regional earthquake simulations, post-event response). It is assumed that CPT models are more efficient than geospatial models, given that CPT models have a mechanistic underpinning and directly sample the subsurface, but are they? The efficacies of these models have not been directly compared using a consistent set of case histories and standard performance metrics. Such an assessment might elucidate strategies for improvement (e.g., what can each type of model teach the other?).

Moreover, the growth of remote sensing has given rise to hundreds of potential geospatial predictor variables (in contrast to the three or four variables used by current models). In combination with emergent algorithmic learning (i.e., machine- and deep-learning, or ML/DL) and the rise of community geotechnical datasets (e.g., CPTs) the opportunity arises to develop a next-generation style of geospatial model wherein below-ground geotechnical measurements are predicted using above-ground geospatial data. The predicted data could be input to a validated,

mechanistic model for predicting liquefaction. This would transfer the prediction target from liquefaction (with relatively sparse training data) to subsurface measurements (for which the training set is potentially vast). Moreover, liquefaction is a physical phenomenon best predicted by mechanics. Conversely, subsurface traits lack theoretical links to above-ground parameters, but surely correlate in complex, interconnected ways – a prime problem for ML/DL.

Ultimately, while it is critical to predict liquefaction occurrence (i.e., its "triggering" and/or expression at the ground surface), decisions about whether, and how, to mitigate liquefaction should be based on expected damage and loss, not on factors of safety or probabilities of liquefaction. In this regard, less effort has historically been given to predicting the downstream effects of liquefaction, such as ground settlement, foundation damage, and financial loss. CPT-based models for predicting ground settlement have been derived from laboratory test results (e.g., Zhang et al., 2002), but have rarely, if ever, been tested in the field against measurements. Models for predicting foundation deformation have similarly been proposed from physical models and computational simulations (e.g., Bullock et al., 2020), but empirical foundation-performance surveys are very rare. And to-date, the authors are unaware of any model for predicting losses due to liquefaction. Researchers are rarely afforded the opportunity to study damage and loss because these data are closely guarded by insurers (to protect intellectual property) and owners (to protect real-estate value). This has prevented popular methods for predicting liquefaction from being extended to also predict its downstream effects.

In addition, once the risk of such losses is assessed, a course of action must be devised. Should the risk be accepted as-is, or, should the risk be mitigated? The degree to which either choice is favorable is a function of its expected cost and benefit during the life of an infrastructure asset. These costs and benefits must be predicted for every earthquake that could impact an asset during

its lifespan, considering also the probabilities of these earthquakes occurring. However, because this requires an integrated seismological, geotechnical, and economic framework, a complete and pragmatic decision-framework for managing liquefaction risk has not been realized in contemporary practice. As a result, decisions are generally made in a heuristic manner – that is, one not guaranteed to be optimal or rational, but nevertheless relied upon for reaching a expeditious solution. Meanwhile, the potential for completing such a framework, and in a manner pragmatic for practitioners, was dramatically increased following the Canterbury earthquakes. The unique role of the New Zealand government led to their assessment of damage and loss for 80,000 buildings on shallow foundation systems.

Accordingly, the research presented herein is motivated by the confluence of: (i) unprecedented datasets resulting from recent earthquakes, which include a range of scientific, engineering, and economic data; and (ii) emergent technologies, including satellite remote sensing (which also significantly aided (i)) and machine- and deep-learning algorithms, which can provide new insights into longstanding earthquake engineering problems. Using these data and technologies, the overarching goal is to test and develop liquefaction occurrence and consequence analytics, with a focus on models based on, or benefiting from, the CPT. Namely, this dissertation will: (1) digitize and publicly compile all existing liquefaction case-histories where CPTs were performed, of which there are ~275; and (2) grow the quantity of such cases to ~15,000 using data from earthquakes in New Zealand. These datasets will then be used to: (3) test existing liquefaction-prediction models based on either the CPT or on geospatial data, from which lessons for future improvement are identified; (4) test a new procedure for correcting "thin layer effects" on CPT data in the context of liquefaction model performance; (5) test existing CPT models that predict ground settlement induced by liquefaction; and (6) train new fragility functions that probabilistically predict

liquefaction-induced ground failure. Using the data compiled in (2), in conjunction with additional foundation damage surveys and insurance loss assessments from New Zealand, a decision framework for mitigating liquefaction beneath buildings on shallow foundations is (7) proposed; and demonstrated. Using insights from the tests in (3), a next-generation geospatial modelling approach, wherein machine- and deep-learning are used to predict below-ground CPT data via above-ground geospatial information, is (8) proposed, trained, tested, and (9) coded into *RapidLiq*, a software program for geospatial modelling of soil liquefaction with near-real-time capabilities. Lastly, existing and newly proposed analytics in (1-7) are (10) coded in *Horizon*, a second software program for analyzing the occurrence, consequence, and mitigation of liquefaction.

1.2 Dissertation Structure and Contents

Table 1 lists all publications resulting from this dissertation, parsed into four themes/chapters. The principal products and findings of these publications, which range from "In Review" to "Published," 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 ten publications in Table 1, and the connections between them, is shown in Figure 2.

Table 1. Publications resulting from this dissertation.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology,* In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

[1]This paper and the data it compiles earned the 2021 *Outstanding Dataset Award* from the NSF DesignSafe C-I.
[2]A presentation by M. Geyin based on this paper was awarded *2nd Place* in the 2019 ASCE GeoCongress Poster Competition.
[3]A presentation by M. Geyin based on this paper was awarded *1st Place* in the 2020 ASCE GeoCongress Poster Competition.

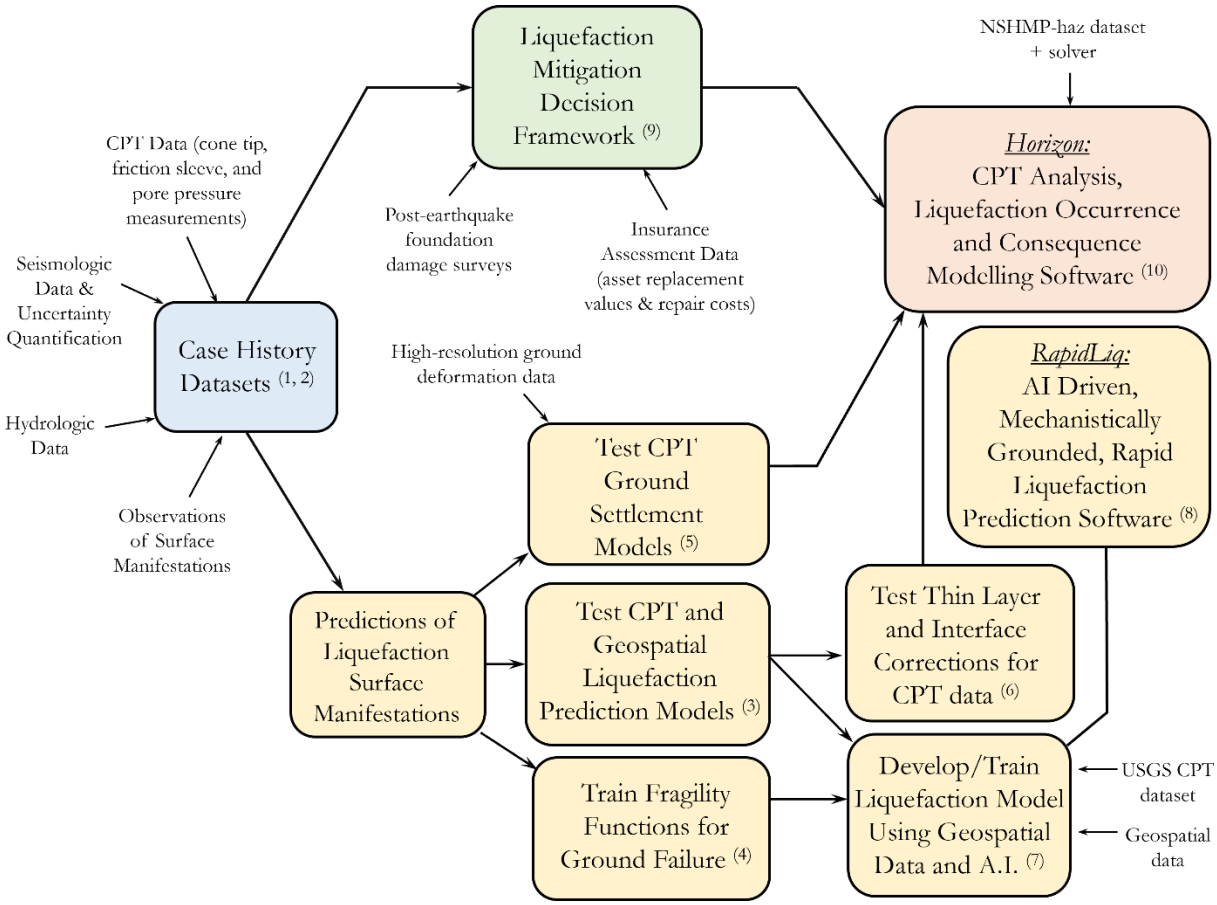


Figure 2. Schematic showing the flow and interconnection of research subjects comprising this dissertation. The four colors correspond to the ensuing four chapters, wherein research is grouped by theme. Numbers in each box correspond to the publications appearing in the first column of Table 1.

1.3 A Succinct Summary of the Ensuing Chapters

Chapter two summarizes efforts to: (i) compile 15,000 CPT-based liquefaction case-histories from three earthquakes in the Canterbury region of New Zealand (i.e., the "Canterbury dataset"); and (ii) digitize and compile from the literature all existing CPT-based case histories resulting from all other earthquakes combined, of which there are ~275 (i.e., the "global dataset"). Chapter three summarizes efforts to test and develop models for predicting liquefaction and its manifestation at

the ground surface. These efforts utilize approximately 50 times more case history data than any prior study. *First*, 23 existing liquefaction-prediction models based on either CPT or geospatial data are tested using the Canterbury and global datasets, from which various lessons and avenues for potential improvement are highlighted. *Second*, fragility functions for probabilistically predicting the occurrence and severity of liquefaction surface manifestations (i.e., "ground failure") are trained using both datasets. *Third*, existing models for predicting the magnitude of ground settlement are tested using data from New Zealand. While such models are widely used, they are derived from laboratory test data and have rarely, if ever, been tested in the field. *Fourth*, a newly proposed procedure for correcting CPT data for "thin layer effects" is tested in the context of liquefaction model performance. *Fifth*, building off the lessons learned from testing geospatial models, and the development of fragility functions developed for CPT models, a next-generation approach to geospatial modelling that utilizes machine- and deep-learning is developed.

Chapter four summarizes efforts to develop and demonstrate a pragmatic framework by which analysts can decide whether, and how, to mitigate liquefaction beneath lightweight structures on shallow foundation systems. Chapter five summarizes the development of *Horizon*, a windows-executable software program wherein all previously mentioned analytics and models are coded. *Horizon* improves upon commercially available software, but will be freely available and open source, while also facilitating user adoption of new models resulting from this research.

Chapter 2. COMPILATION OF LIQUEFACTION CASE HISTORIES

CPT-based liquefaction case histories may be used to test or train liquefaction models based on CPT data. Two separate datasets (i.e., the "Canterbury" and "Global" datasets), were compiled by assembling: (i) identifying information (e.g., geographic coordinates); (ii) processed CPT data; (iii) peak ground accelerations (*PGA*) and their uncertainties; (iv) earthquake magnitudes (M_w); (v) groundwater table (GWT) depths; and (vi) the classified occurrence and severity of liquefaction manifestation at the ground surface, with focus on free-field level ground sites. The scheme of the assembly, to include methodology, CPT post-processing, classification of liquefaction manifestations, and event-specific measurements, are described in the following two publications:

2.1 PUBLICATION 1: GEYIN ET AL. (2021). "CPT-BASED LIQUEFACTION CASE HISTORIES COMPILED FROM THREE EARTHQUAKES IN CANTERBURY, NEW ZEALAND." *EARTHQUAKE SPECTRA.* DOI: 10.1177/8755293021996367.

- This study compiles a curated digital dataset of approximately 15,000 CPT-based case-histories from three earthquakes in Canterbury, New Zealand: 4 Sept. 2010 M_w7.1, 22 Feb. 2011 M_w6.2, and 14 Feb. 2016 M_w5.7 events.
- As shown in Figure 3, this work considerably augments the data available for CPT-based model training and testing (*by at least 50x*), presenting the profession with a unique opportunity to advance the science of liquefaction prediction.
- The compiled, post-processed data is presented in a dense structural array (i.e., a single file), allowing researchers to easily access and analyze a wealth of information pertinent to free-

field liquefaction response (i.e., triggering and surface manifestation). The structure of the file is shown in Figure 4 and is provided in both Matlab and Python formats.

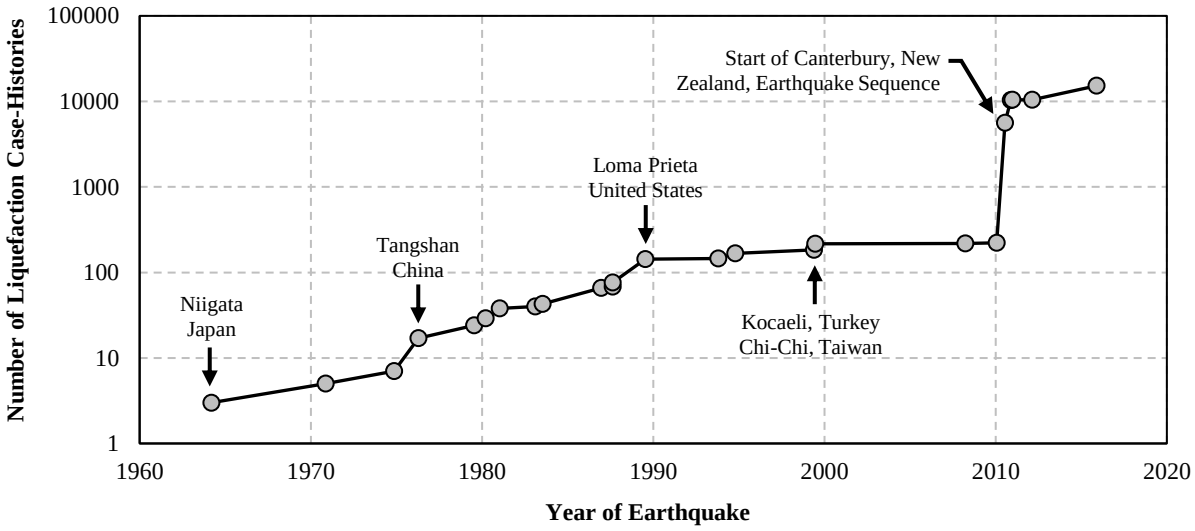


Figure 3. Chronology of CPT-based liquefaction case histories.

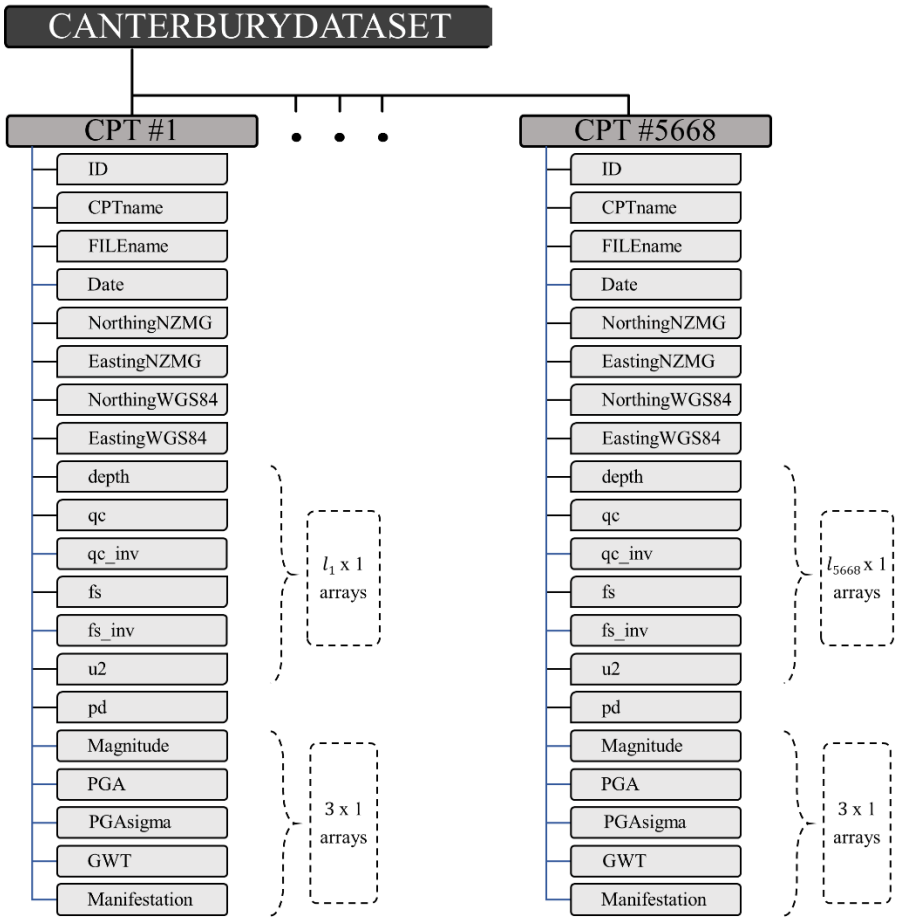


Figure 4. Depiction of the Canterbury case-history dataset structure array.

- Nuances and caveats that analyst should be aware of are discussed in detail. These are presented in four categories as follows: (i) additional data exclusion criteria; (ii) alternative sources of data; (iii) correlations and decisions for analysis; and (iv) lingering uncertainties.

2.2 PUBLICATION 2: GEYIN M AND MAURER BW (2021). "CPT-BASED LIQUEFACTION CASE HISTORIES FROM GLOBAL EARTHQUAKES: A DIGITAL DATASET (VERSION 1)." *DESIGNSAFE CI*. HTTPS://DOI.ORG/10.17603/

- To compare various findings in Canterbury with regions elsewhere, this study compiled 275 CPT-based liquefaction case histories from 21 globally distributed earthquakes.

- This is effectively the sum of all CPT-based case histories from all earthquakes combined outside of Canterbury. An electronic compilation of these cases was not previously available.
- These cases were sourced from the literature, to include observations of manifestation severity, CPT soundings, and estimation of GWT depth and ground-motions, as generally reported by original investigators. When available, recent refinements were adopted from the literature.
- Approximately half of the CPTs required digitization from paper documents.
- The compiled, post-processed data is presented in the same structure array format as the Canterbury data (a single file), allowing researchers to easily analyze both datasets.
- Information describing the contents and structure of the case-history data are provided in a "ReadMe" document; the complete database is available from the citation above.

Chapter 3. TESTING AND DEVELOPING MODELS TO PREDICT LIQUEFACTION AND ITS SURFACE MANIFESTATION

Using the two datasets compiled in the previous chapter, efforts were initiated to test and develop models for predicting liquefaction and its manifestations at the ground surface. These efforts utilized approximately 50 times more case history data than any prior study (and in some cases, are the first such studies of their kind). These efforts are summarized in five publications:

3.1 PUBLICATION 3: 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. DOI: 10.1177/8755293019899951.

- This study tested 23 existing liquefaction-prediction models based on either CPT or geospatial data using both the Canterbury and global datasets, from which various lessons and avenues for potential improvement were highlighted.
- CPT models have never been tested against geospatial models, while the popular question of "which CPT model performs best?" remains contentious, owing to the prior lack of test data.
- Namely, six liquefaction triggering models (Robertson and Wride, 1998 [RW98]; Architectural Institute of Japan, 2001 [AIJ01], Moss et al., 2006 [Mea06]; Idriss and Boulanger, 2008 [IB08]; Boulanger and Idriss, 2014 [BI14]; and Green et al., 2019 [Gea19]), used in combination with three different liquefaction manifestation models (Iwasaki et al., 1978 [LPI], van Ballegooy et al., 2014 [LSN], and Maurer et al., 2015 [LPI_{ISH}]), were tested against five "geospatial" liquefaction models that use readily available surface parameters (Zhu

et al., 2015; 2017) of which three are region-specific, developed for New Zealand [RGM 1, 2, and 3] and two are globally applicable [GGM 1 and 2].

- Model performance was evaluated via the Area Under a Receiver Operating Characteristic Curve. This area, or *AUC*, is statistically equivalent to the probability that a positive observation has a higher prediction-model index than a negative observation. In other words, the higher the *AUC*, the more efficient a model is at separating positive and negative cases.
- Results in the form of median *AUC* are summarized in Figure 5 and ordered by the year in which each model was proposed. In this regard, CPT models are dated in accordance with the triggering model but are grouped/symbolized by manifestation model. Thus, Figure 5 allows for assessments to be made regarding the evolution of performance (i.e., prediction efficiency) through twenty years of model research and development.

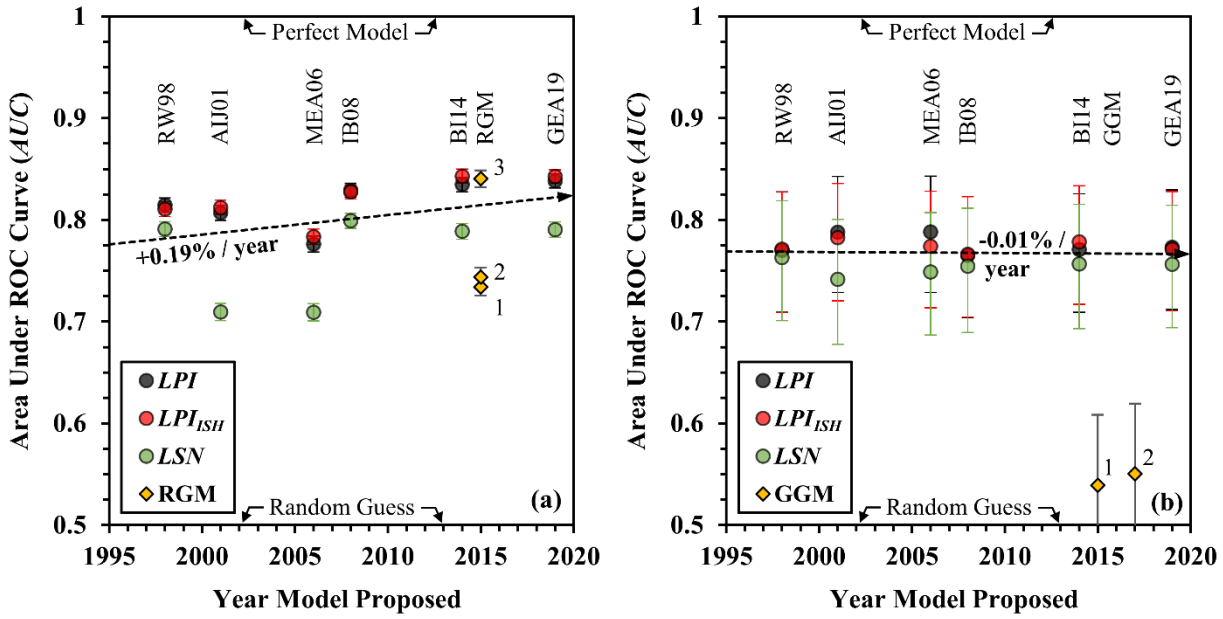


Figure 5. Summary of liquefaction-model performance – quantified by *AUC* – for 23 popular models, ordered by year proposed: **(a)** Canterbury dataset; and **(b)** Global dataset. Markers denote median *AUC*; bars are 95% confidence intervals on *AUC*; all model acronyms are identified in the text. Trendlines are developed from linear regression and do not include RGM or GGM data points.

- Finite-sample uncertainty was also accounted for and used to establish statistical significance in the form of p-values (i.e., probabilities that two samples of performance could have come from the same parent distribution). In some cases, for example, a model with better *AUC* could be found to be not significantly better than another model once this uncertainty is considered.
- The analyses showed that geotechnical predictions are significantly more efficient than geospatial models on a global scale. This finding was not altogether surprising, given the difficulty of inferring below-ground conditions from above-ground parameters (i.e., the task of geospatial models), especially across diverse geomorphic, topographic, and climatic environs. However, the analyses also showed that successive CPT models proposed over the last twenty years show little or no demonstrable improvement.
- In addition, geospatial models performed equally well for large subsets of the data – a provocative result given the relative time- and cost-requirements underlying these predictions. In all, the study demonstrates both the provocative potential of geospatial modelling as well as the significant shortcomings of existing models. Building off this study, a new geospatial model informed by machine- and deep-learning is proposed in Publication 7.
- As more test data becomes available, and as better geospatial predictors are identified, improved global performance should result. In contrast, considering the 20-year trendlines in Figure 5, it appears less likely that geotechnical models will soon improve greatly, should the status quo continue. Neither fine-tuning of model parameters, nor incremental grown in training data, are likely to inflect these trends upwards.
- Several lessons for improving geotechnical and geospatial models were identified and discussed in detail. This are succinctly summarized as: (a) identifying and addressing sources of inter-event instability; (b) identifying better geospatial proxies of soil type, and by corollary,

liquefaction susceptibility; (c) developing geotechnical triggering and manifestation models within a more rational and internally consistent framework; and (d) more rigorous investigation and modelling of the multi-scale interactions which drive a soil profile's cumulative response.

3.2 PUBLICATION 4: GEYIN M AND MAURER BW (2020). "FRAGILITY FUNCTIONS FOR LIQUEFACTION INDUCED GROUND FAILURE." *ASCE JOURNAL OF GEOTECHNICAL AND GEOENVIRONMENTAL ENGINEERING* 146 (12): 04020142. HTTPS://DOI.ORG/10.1061/(ASCE)GT.1943-5606.0002416.

- This study essentially extended the deterministic CPT models in Publication 3 to probabilistically predict the severity of liquefaction manifested at the ground surface, which is a popular and pragmatic proxy of damage potential for infrastructure.
- Towards this end, the Liquefaction Potential Index (*LPI*) and similar models are commonly used, and often codified, to predict surface manifestations on level ground. These predictions typically use deterministic thresholds from the literature - obtained via calibration on case-history data - to classify the expected manifestation. While widely adopted, such thresholds obscure the uncertainty of expected outcomes and are incompatible with probabilistic frameworks. Proposed thresholds are also intimately tied to the liquefaction analytics used to compute them and to the methodology used to select them, each of which can conflict with forward applications, leading to erroneous predictions.
- Accordingly, using the compiled Canterbury and Global datasets, this study developed fragility functions conditioned on CPT-based triggering and manifestation models popular in practice (i.e., the same 18 CPT models studied in Publication 3). When used in series, these models attempt to characterize the cumulative behavior (a.k.a. "system response") of soil profiles.

- The proposed fragility functions all have the same functional form, such that users can easily extend their deterministic workflows by adopted fragility coefficients specific to the models of their choosing. As an example, functions conditioned on *LPI* using the BI14 triggering model are shown in Figure 6. A means to incorporate uncertainty in the functions was also provided.

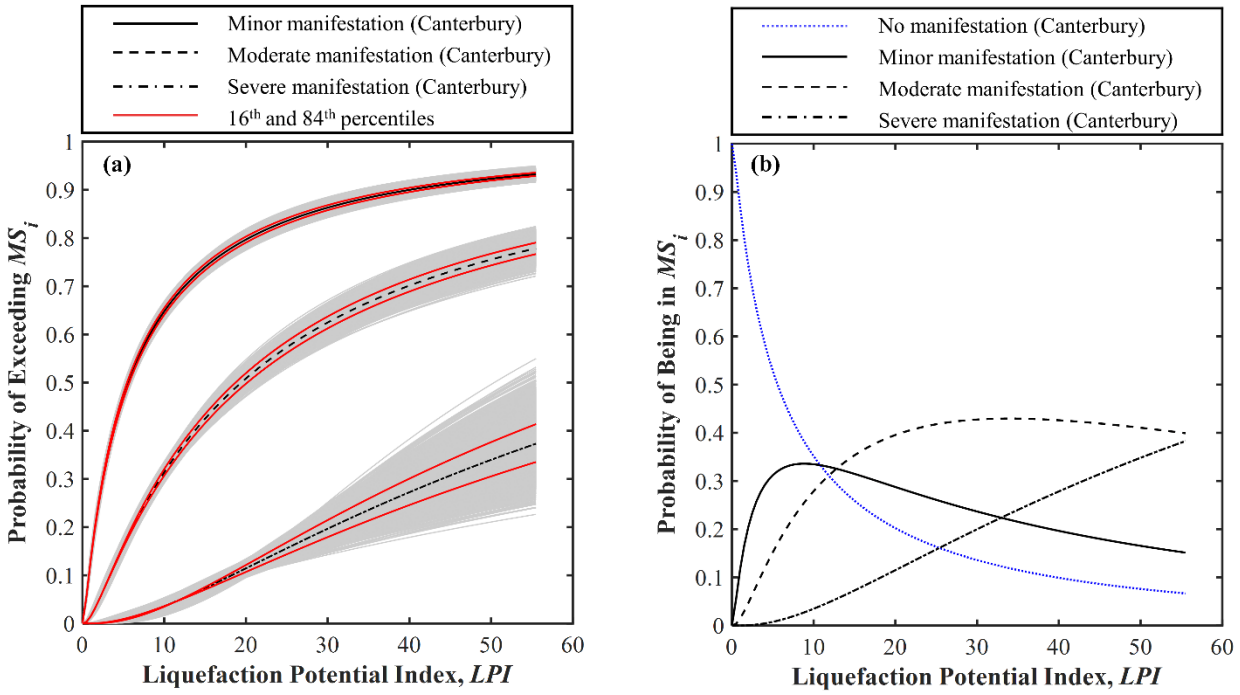


Figure 6. Probability of manifestations **(a)** exceeding, and **(b)** being in severity class, MS_i, conditioned on *LPI* computed using BI14. In **(a)**: grey lines are bootstrap simulations; black lines are median functions.

- Model application was demonstrated by predicting: (i) liquefaction manifestations in Christchurch, NZ, resulting from an Alpine Fault earthquake, wherein a logic-tree was used to ensemble predictions from 18 models, as shown in Figure 7; and (ii) the return period of liquefaction manifestations in the SODO district of Seattle, USA, wherein predictions were compared to historical observations. As an example, the predicted return period of minor or greater manifestations is shown in Figure 8. Historically, manifestations were observed in the

study area – particularly that with lowest computed return period of 60 years – following earthquakes in 1949, 1965, and 2001 (Chleborad and Schuster 1990; Bray et al. 2001). Assuming a Poisson process, and adopting (i) the return period of 60 years; and (ii) a 170-year exposure window beginning 1850 (ca. Seattle's founding), the probability of three observations is 22% (the only quantity more likely is two, having 23% probability). While discrepancies between observed and predicted return periods would not necessarily discredit the latter, their close agreement nonetheless gives credence to the developed functions and results.

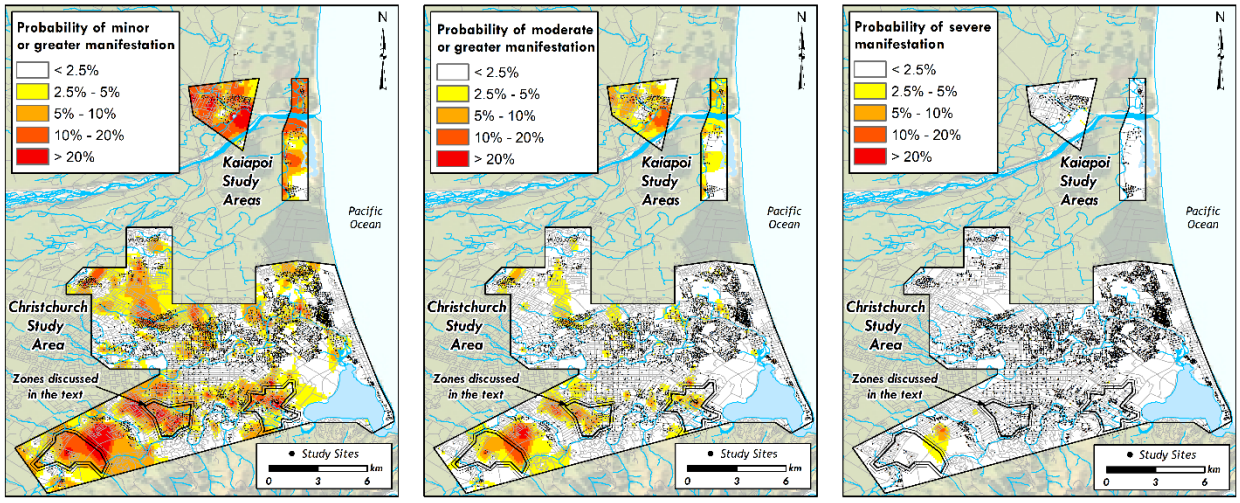


Figure 7. Probabilities of minor, moderate, and severe liquefaction manifestations in Christchurch, NZ, during an Alpine Fault earthquake. Results are obtained by assembling 18 models.

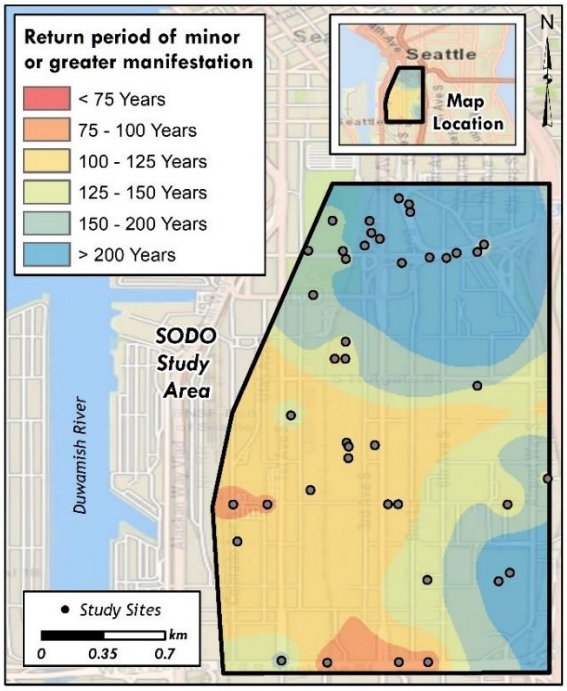


Figure 8. Computed return period of liquefaction manifestations across SODO study area.

3.3 PUBLICATION 5: GEYIN M AND MAURER BW (2019). "AN ANALYSIS OF LIQUEFACTION-INDUCED FREE-FIELD GROUND SETTLEMENT USING 1,000 + CASE-HISTORIES: OBSERVATIONS VS. STATE-OF-PRACTICE PREDICTIONS." *ASCE GEOTECHNICAL SPECIAL PUBLICATION NO. 308*: 489–498.

- Liquefaction-induced free-field ground settlement can be attributed to several mechanisms, including volumetric reconsolidation strain, ground loss due to ejecta, and vertical components of lateral spreads and flow slides. In practice, the volumetric-strain settlement is typically predicted using one of several semi-empirical methods, but there is no consensus of which, if any, is most accurate. Moreover, no method has been proposed to predict the magnitude of settlement owed to ejecta, even though this component could be larger than any other.

- The lack of a complete and trusted model for predicting free-field settlement is due, in part, to the scarcity of case-history data needed to apportion and model its individual components. For starters, there are very few measurements of settlement in the field.
- However, the Canterbury earthquakes resulted in an unprecedented LIDAR survey of settlements induced by each event, corrected for tectonic movements, as well as the geotechnical and seismological data studied in the previous publications.
- Accordingly, this study compiled and analyzed 1,013 high-quality ground settlement case-histories from Canterbury, such that the contributing components of settlement could be reasonably accounted for and quantified.
- These analyses found that popular CPT-based models for predicting ground settlement exhibit heteroskedastic bias, such that smaller total settlements (< ~6 cm) are underpredicted while larger total settlements (> ~6 cm) are overpredicted, as shown in Figure 9a.

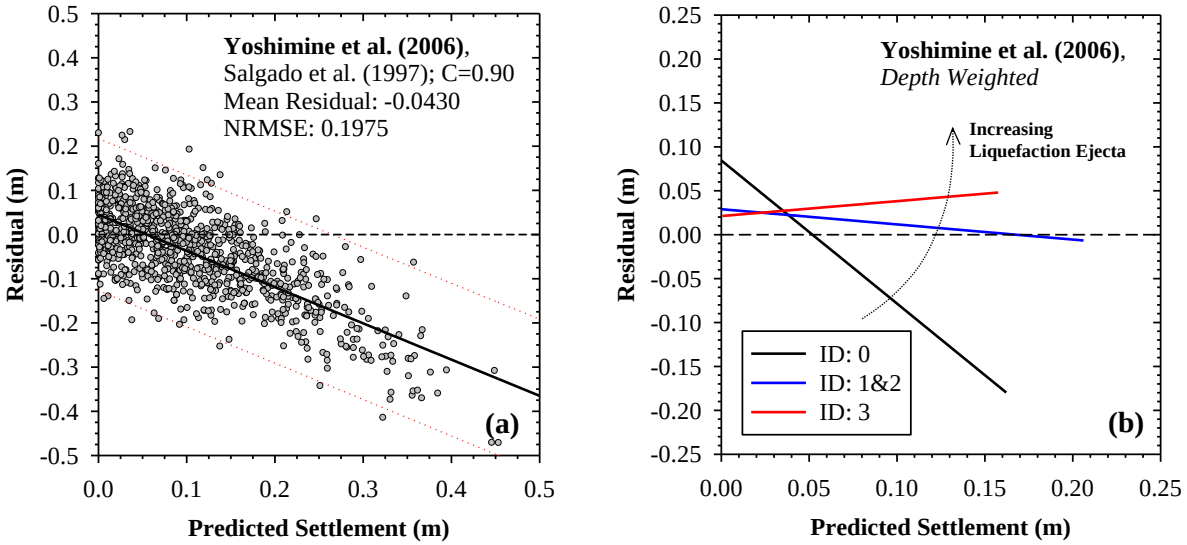


Figure 9. (a) Settlement predicted by a popular CPT model vs. residual errors, exhibiting strong heteroskedastic bias; **(b)** residual trendlines grouped by observed severity of liquefaction ejecta, showing the need to account for and model the contributions of settlement owed to ejecta.

- Introduction of a depth-weighting function, which most current CPT-based settlement models do not utilize, lessened the heteroskedastic prediction bias, showing promise for future study.
- Moreover, it was found that prediction-performance was related to the observed volume of liquefaction ejecta, as shown in Figure 9b, illustrating the need to account for and model this mechanism of settlement. Specifically, it was seen that as the observed volume of ejecta increases, residuals tend to decrease, especially for larger predicted settlements. In other words, overpredictions of observed settlement diminish as the volume of ejecta increases. Thus, the general tendency for existing prediction methods to overpredict total settlement (e.g., due to the lack of depth-weighting schemes) is to some degree counteracted by the fact these methods do not account for ejecta, the omission of which should result in underprediction of settlement.
- Collectively, the findings suggest that popular methods for predicting free-field settlement have significant room for improvement.

3.4 Publication 6: Geyin and Maurer (2020), "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology*, 291, DOI: https://doi.org/10.1016/j.enggeo.2021.106221.

- The volume of soil mobilized by the cone penetration test (CPT) acts as a low-pass filter on the true stratigraphy, potentially removing information such as the data defining a thin layer of soil or the interface between two dissimilar soils, as shown in Figure 10. These spatial smoothing effects, commonly called "thin layer effects" could reduce the efficacy of CPT-based liquefaction models. Following Publication 3 (Geyin et al., *Earthquake Spectra*) several researchers noted that the performance of liquefaction models computed therein was strongly correlated to the

"interbeddedness" of profiles. That is, liquefaction predictions were most inaccurate for profiles with many thin layers of dissimilar soil.

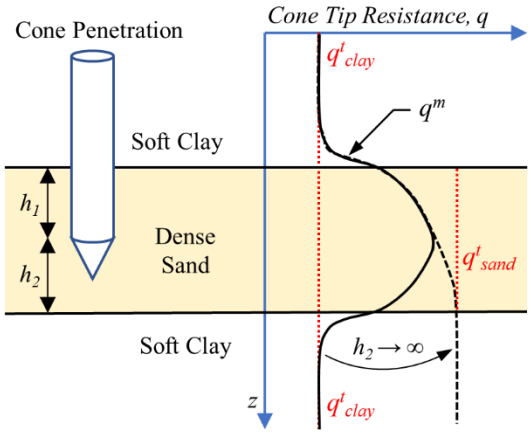


Figure 10. Conceptual schematic showing cone penetration in layered soil. The measured cone-tip resistance (q^m) deviates from the "true" resistance (q^t) that would be measured in each material if measurements reflected conditions at discrete points.

- While thin-layer effects have long been recognized and studied, Boulanger and DeJong (2018) proposed the first automated correction procedure, the efficacy of which has never been tested.
- Accordingly, this study evaluated the Boulanger and DeJong (2018) procedure in the context of CPT-based liquefaction model performance. Using the Canterbury and Global datasets, 18 different liquefaction models were studied, resulting in 36 performance trials. The results, summarized in terms of the change in model *AUC* when the procedure, are shown in Figure 11.
- In 1 of these trials, the CPT inversion procedure increased model efficiency to a statistically significant degree, but in 23 others it significantly decreased efficiency. This decline in performance tended to grow as profiles became more stratified, opposite of what should ideally occur. To explore remedies, a liquefaction triggering curve was rederived from inverted CPT data, such that its training and forward implementation were made consistent. Nonetheless, this exacerbated the decline in prediction efficiency.

- Ultimately, the results of this study are not a direct assessment of the Boulanger and DeJong (2018) procedure. However, the results do provide evidence that this procedure – when applied to existing CPT-based liquefaction models – provides no demonstrable performance benefit.

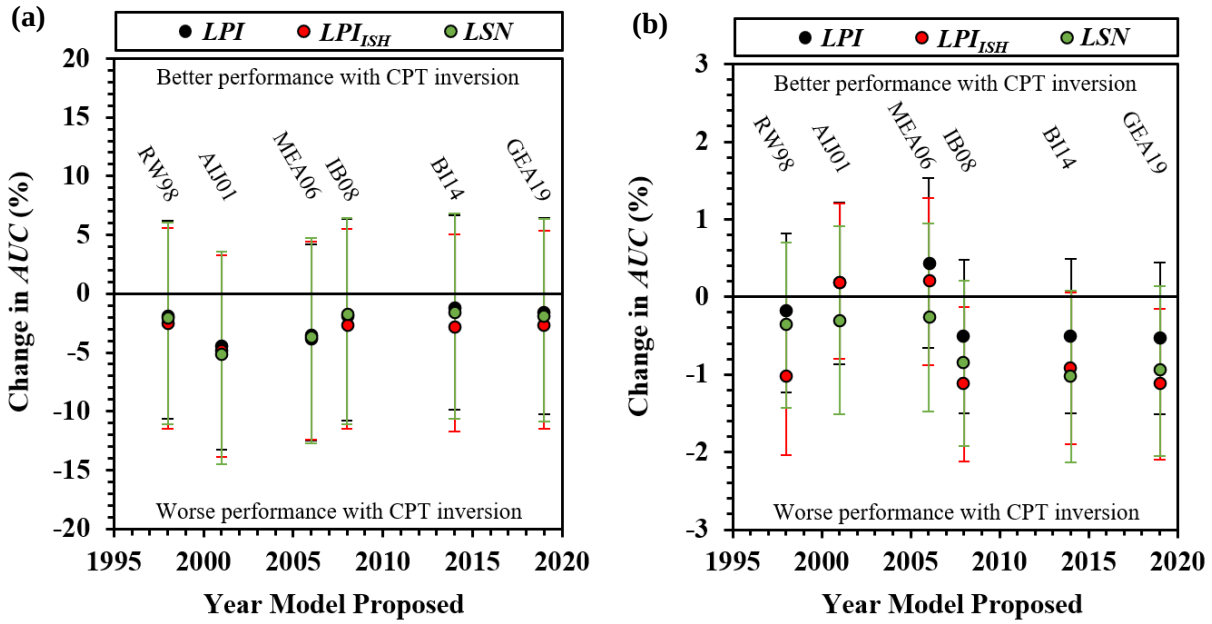


Figure 11. Summary of change in liquefaction-model performance – quantified by *AUC* – for 18 models to which the Boulanger and DeJong (2018) procedure was applied, ordered by year proposed: (a) Global dataset; and (b) Canterbury dataset.

3.5 Publication 7: Geyin M, et al. (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Scenario Planning." *Soil Dynamics and Earthquake Engineering*, **In review**.

- Using insights from the tests of geospatial liquefaction models in Publication 3 (summarized in Figure 5), this study proposed, trained, and tested a next-generation geospatial liquefaction model. In combination with satellite remote-sensing measurements, machine- and deep-learning algorithms (or ML/DL) were used to predict below-ground CPT data using above-

ground information. These predictions were then input to the CPT-based fragility functions developed in Publication 4, an example of which is shown in Figure 6.

- Soil liquefaction is inherently a regional phenomenon. Predictions of liquefaction would thus ideally be made: (i) rapidly (e.g., in real-time after an event); (ii) at high spatial resolution (e.g., commensurate with the scale of single assets); and (iii) over the regional scope of transportation networks. Problematically, state-of-practice prediction models rely on costly subsurface data, such as that from the CPT. Given the growth of community geotechnical data, remote sensing, and machine learning, a more accurate regional scale solution may be possible.
- Towards that end, interest in models based on free and readily available geospatial data has recently grown. In Publication 3, the geospatial model of Zhu et al. (2017) was compared against 18 CPT-based models. This elucidated both the provocative potential of a geospatial model (it outperformed nearly all CPT models in New Zealand - a startling result) as well as significant room for improvement (globally it performed only slightly better than random guessing). Informed by these analyses, this study proposes a geospatial model driven by algorithmic learning (benefiting from machine- and deep-learning insights) but pinned to a physical framework (benefiting from mechanics and a tradition of regression modelers).
- Several limitations of the Zhu et al. (2017) model may be overcome using the proposed approach. First, the Zhu et al. model was trained directly on outcomes (i.e., observations of ground failure) rather than on the mechanistic causes of those outcomes (i.e., subsurface engineering properties). While this lack of a mechanistic underpinning can be overcome with vast training data (e.g., how voice transcription apps predict words without understanding language), current ground failure inventories are arguably too sparse. Given the adopted approach, inadequate training data can result in a divergence from mechanistic principles (e.g.,

prediction of liquefaction given shaking too weak, from a mechanistic perspective, to induce liquefaction). Second, the Zhu et al. model uses just 5-variables (four terms represent capacity (distances to river and coast, precipitation, and mapped V_{S30}) and one represents demand (peak ground velocity)). Notably, none of these variables are likely to correlate to the type of soil, or by corollary, to the susceptibility of the soil to liquefaction. This was a common cause of mispredictions identified in Publication 3, with the Zhu et al. model expecting susceptibility if the ground is flat, saturated, and near water. However, such profiles can consist mostly of soils less- or un-susceptible to liquefaction (e.g., clays, peats, or gravels). An improved model might thus use mapped geologic data, when available. Third, the Zhu et al. model is a regression equation. This method of modelling inherently requires hypotheses of what is believed to matter and how (beliefs that are unnecessary with ML/DL). The efficient prediction of subsurface traits likely requires more than four geospatial variables, yet regression limits the number easily modeled. Algorithmic learning would allow more geospatial predictor variables to be used, with greater potential for those variables/data to be exploited fully.

- Accordingly, this study proposed a novel geospatial model that is driven by algorithmic learning but pinned to a mechanistic framework. The benefits of this approach include: (i) a mechanistic underpinning; (ii) a significantly larger training set; and (iii) ML/DL algorithms.
- Using the proposed approach, three ML/DL models were trained, using ten geospatial variables, to predict *LPI* values computed from in-situ geotechnical tests subjected to a range of hypothetical ground motions. The goal of the ten geospatial variables is to predict the traits which give rise to low or high *LPI*. The USGS national CPT database, which consists of 1,712 CPTs, was used to create training and test sets. In effect, the models predict below-ground geotechnical data using above-ground geospatial information. To use these models in the

forward direction, the predicted *LPI* values are input to the fragility functions from Publication 4, which predict the probability of liquefaction manifestation (i.e., "ground failure") as a function of *LPI*. An illustrative synopsis of the proposed approach is shown in Figure 12.

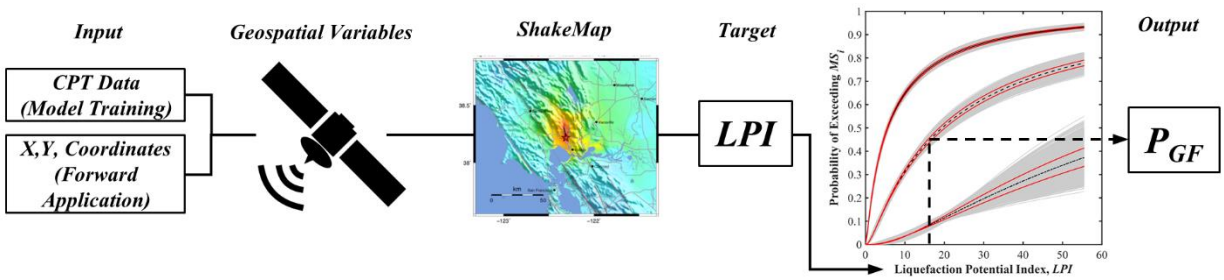


Figure 12. Synopsis of the modelling approach to predicting probability of ground failure (P_{GF}).

- Various ML/DL techniques were explored, to include Gaussian process models, support vector machines, tree ensembles, bagging and boosting, random forests, and neural networks, both shallow and deep. Each comes with many options (e.g., net training algorithm, back-propagation method, hyperparameter optimization), which were explored to identify optimal models. K-fold validation was used to control overfitting. Ultimately, three models were developed. The first – referred to as the ML model – is a boosted ensemble of decision trees, wherein numerous relatively weak models are coalesced to form one high-quality model. The second – referred to as the DL model – is a deep (7-layer) artificial neural network model. A third "Ensemble model" was also created by averaging the outputs of the ML and DL models.
- Using unbiased liquefaction observations, the three models were tested against Rashidian and Baise (2020) [RB20], a lightly modified update of the Zhu et al. (2017) model. An example comparison is shown in Figure 13, using RB20 and the Ensemble model, for the 2011 Mineral, Virginia and 2016 Ridgecrest, California earthquakes.

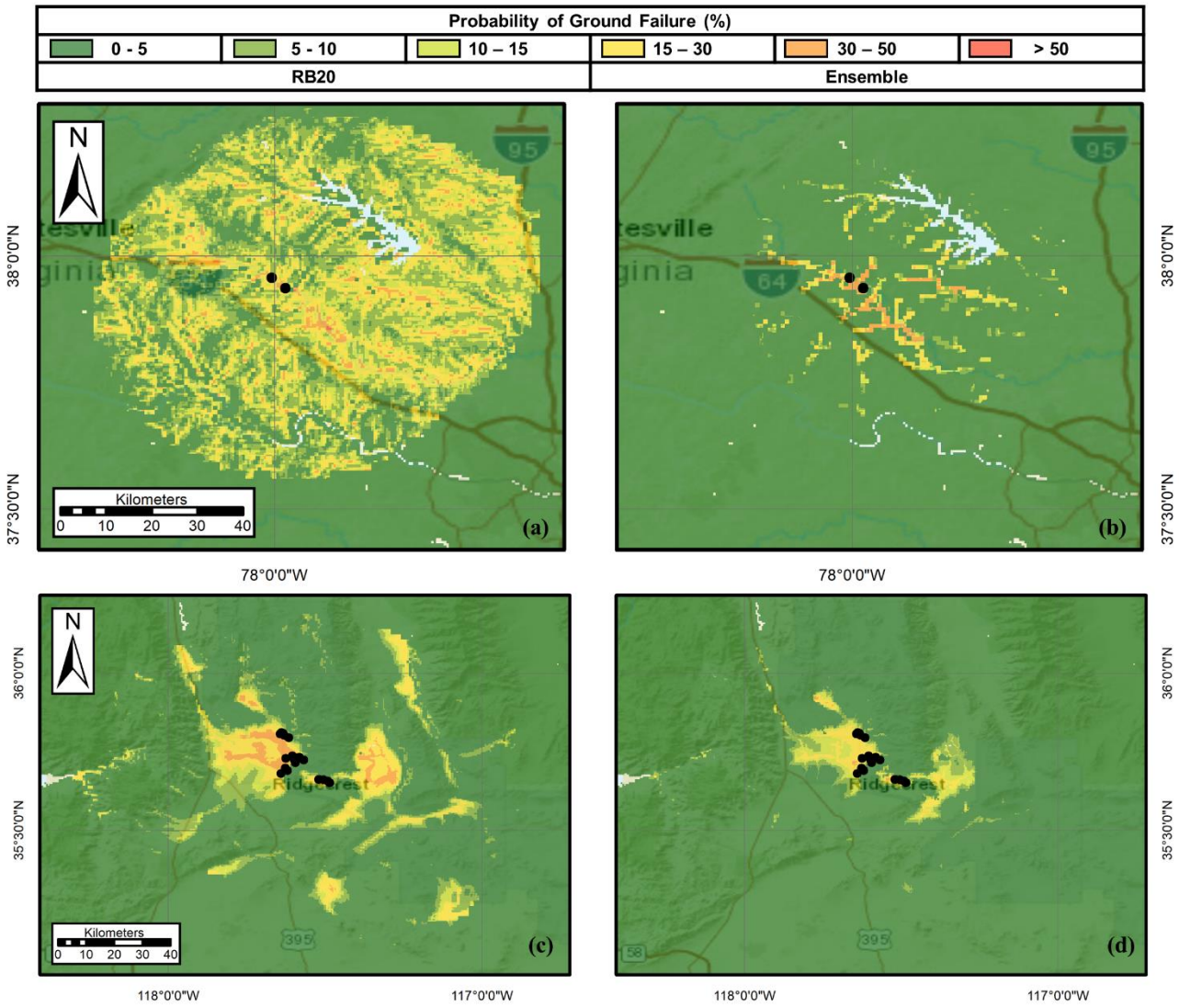


Figure 13. Probabilities of ground failure in the 2011 Mineral earthquake, as computed by (a) RB20 and (b) the Ensemble model; and in the 2016 Ridgecrest earthquake, as computed by (c) RB20; and (d) the Ensemble model. Black dots are observed ground failures.

- It was shown that ML, DL and Ensemble models provide efficient predictions in unbiased, forward prediction of liquefaction induced ground failure.
- All developed models were statistically tested against the RB20 model using ROC analyses as described in Publication 3. Collectively, these tests indicated that the proposed models predict

ground failure with similar or better efficiency as RB20 and thus warrant application and further evaluation.

- Immediate applications include: (i) regional loss estimation and disaster simulation; ii) city planning and policy development; (iii) emergency response; (iv) post-event reconnaissance (e.g., to remotely identify sites of interest, thus maximizing the efficiency of field reconnaissance); and (v) areas that lack geotechnical testing (e.g., humanitarian/aid studies).
- Ultimately, the presented models were developed using a modest training set and similarly limited set of predictor variables. Expanding upon the proposed approach, improved geospatial liquefaction models could thus be developed for regional, national, or global application.

3.6 PUBLICATION 8: GEYIN AND MAURER. (2021). "RAPIDLIQ: SOFTWARE FOR NEAR-REAL-TIME PREDICTION OF SOIL LIQUEFACTION", DESIGNSAFE-CI. DOI: HTTPS://DOI.ORG/10.17603/DS2-4BKA-Y039.

- One of the main limitations of any Artificial Intelligence model is the lack of a defined analytical expression easily ported and executed via hard copy. It is thus critical that trained ML/DL models be provided as code, ideally in a format accessible to a broad userbase. Despite this necessity, enumerable ML/DL models in the field of geotechnics and geohazards have been published without code, meaning – whether accidentally or intentionally – that the respective authors describe the development of a model that will never be used. In addition to the challenge above, geospatial models require that multiple predictor variables are compiled at regional scale before predictions can be made.

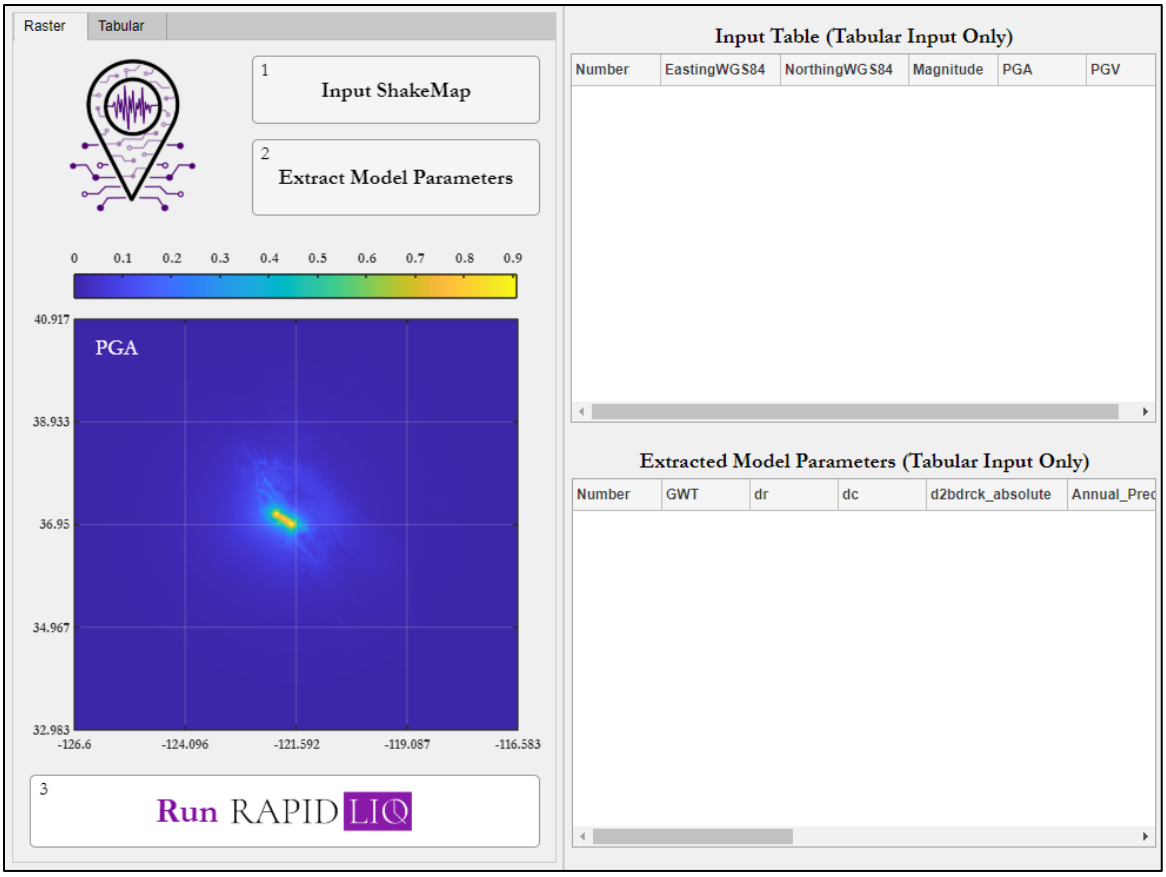


Figure 14. RapidLiq Interface: ShakeMap Input.

- To facilitate user adoption and evaluation, the aforementioned models in Publication 7 were coded into *RapidLiq*, a new Windows software program with a simple-to-use interface (Figure 14). All predictor variables were compiled within *RapidLiq*, making user implementation trivial. The only required input is a ShakeMap of ground-motion intensity measures, either in Extensible Markup Language (.xml) or Geotagged Image File (.tiff) format. The first is easily downloaded from the USGS earthquake catalog minutes after an earthquake, or for numerous future scenario events. Moreover, *RapidLiq* also accepts tabular inputs, should a user wish to enter specific sites of interest and ground-motion parameters at those sites, rather than study the regional effects of an earthquake.

Chapter 4. A PRAGMATIC DECISION TOOL FOR LIFECYCLE COST-BENEFIT ANALYSIS OF LIQUEFACTION MITIGATION BENEATH LIGHTWEIGHT STRUCTURES.

Decisions about risk intervention, such as ground improvement to mitigate soil liquefaction, should be made based on the expected cost and lifecycle benefit (i.e., reduction in loss over the life of an infrastructure asset). This requires accurate predictions of loss, with and without intervention, for every earthquake that could impact an asset during its lifespan, considering also the probabilities of these earthquakes occurring. Yet, because field-assessments of damage and loss due to liquefaction are historically sparsely available, decisions are generally made in a heuristic manner – that is, one not guaranteed to be optimal or rational, but nevertheless relied upon for reaching a expeditious solution. Meanwhile, the potential for completing a practical decision framework was dramatically increased following the Canterbury earthquakes. The unique role of the New Zealand government led to their assessment of damage and loss for 80,000 lightweight buildings on shallow foundation systems. Thus, motivated by the unprecedented engineering and economic data compiled during the Canterbury earthquakes, Chapter 4 proposes a pragmatic decision tool for lifecycle cost-benefit analysis of liquefaction mitigation. This tool will be: (i) applicable to lightweight structures on shallow foundation systems; (ii) developed as an extension to CPT-based liquefaction analytics already in popular use; and (iii) allow for lifecycle benefits to be quantitatively weighed for a range of intervention strategies. Ultimately, this tool could be used to decide whether, and how, liquefaction should be mitigated, or to demonstrate the economic feasibility of a new type of ground improvement (e.g., bio-inspired and bio-mediated technologies). Towards this end, the proposed functions are demonstrated via applications to a variety of sites and locations in the United States.

4.1 PUBLICATION 9: GEYIN M, MAURER, BW, AND VAN BALLEGOOY S (2020). "LIFECYCLE LIQUEFACTION HAZARD ASSESSMENT AND MITIGATION." *ASCE GEOTECHNICAL SPECIAL PUBLICATION* NO. 320: 312-320 HTTPS://DOI.ORG/10.1061/9780784482834.034.

- This study, presented as a conference paper, introduced the planned concept and developed fragility functions for predicting damage, and vulnerability functions for predicting loss due to liquefaction. This was done considering one liquefaction model and all shallow foundation systems. Later, separate functions are conditioned on different liquefaction models and developed for several distinct shallow-foundation systems that populate the dataset as implemented in Chapter 5.
- In a performance-based earthquake engineering framework, shown schematically in Figure 15, a series of functions are used to predict the seismic hazard, system response, and damage and loss. Knowledge of the expected loss then informs whether, and how, the hazard is mitigated. In the jargon of performance-based earthquake engineering, fragility functions predict damage and vulnerability functions predict loss. Both are central to risk management strategy.

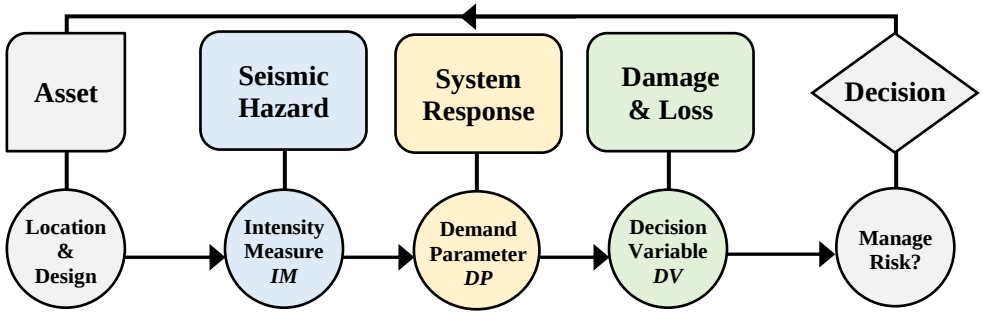


Figure 15. Schematic of the progression that leads to losses within a performance-based earthquake engineering framework, the provenance of which is summarized by Kramer (2008).

- This study used data obtained from the New Zealand government that consisted of: (i) surveys of foundation damage due to liquefaction, wherein the mode and magnitude of damage was classified; and (2) pre- and post-event insurance assessments, from which monetary loss may be determined. All previously mentioned seismic and geotechnical data was also used.
- Three types of shallow foundation systems were inspected: (1) concrete slab-on-grade with a thickened edge or "ring beam" footing; (2) suspended timber floors with perimeter concrete footings; and (3) suspended timber floors supported by shallow, small diameter piles (~20 cm). Between 1999 and 2015, 50.9% of residential construction in the U.S. utilized slab-on-grade foundations, consistent with type 1; an additional 16.9% utilized crawl-space designs consistent with types 2 and 3 (Bradtmueler et al., 2015). In addition, many lightweight commercial and industrial structures have similar designs. The results of this research are thus applicable to design practices in the U.S. and elsewhere.
- This study combined the three foundation types together. Seven modes of foundation deformation were documented: stretching, hogging, dishing, racking, tilting, discontinuity, and global settlement. The severity of each mode, for each foundation in each earthquake, was classified based on its magnitude as either: none-to-neglible, moderate, or severe. Many foundations showed multiple modes of damage with different severities. To develop fragility and vulnerability function herein, the "worst" foundation-damage severity was recorded for each foundation in each event (i.e., the greatest severity, independent of deformation mode, if multiple modes were observed).
- In parallel with the foundation inspection program, analysts assessed the repair cost and replacement value of each residential building following each earthquake. The results of these

assessments are represented via the Building Damage Ratio (*BDR*), which allows for losses to be compiled for various buildings having different values. *BDR* is defined as:

$$BDR = \frac{Repair\ Cost}{Replacement\ Value} \quad (1)$$

- Using this data, fragility functions for predicting damage were developed, an example of which is in Figure 16.

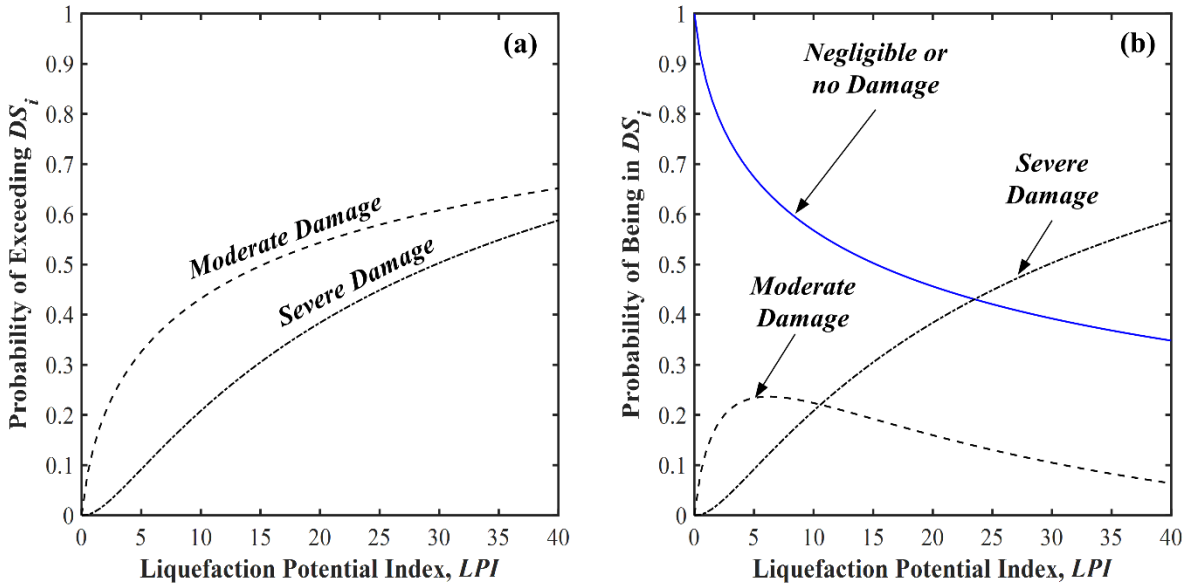


Figure 16. Fragility functions conditioned on *LPI*, as computed using Boulanger and Idriss (2014) liquefaction triggering model, for predicting the probability of **(a)** exceeding a foundation damage-state; and **(b)** being in a foundation damage-state.

- The framework is further expanded to model risk-aversion behaviors or to include environmental costs and losses. Vulnerability functions were developed using *BDR* and the form of Ramirez et al. (2012):

$$E[BDR|LPI] = \sum_{i=1}^{3} E[BDR|DS_i, LPI] \cdot P[DS_i|LPI] \quad (2)$$

Where *BDR* is as previously defined; E denotes the expected value of *BDR*; subscripts refer to the three foundation-damage states (i.e., 1 = negligible or no damage; 2 = moderate damage; and 3 = severe damage); and $P[DS_i|LPI]$ describes the suite of fragility functions shown in Figure 15b. The expression for $E[BDR|LPI]$ is shown conceptually in Figure 17a.

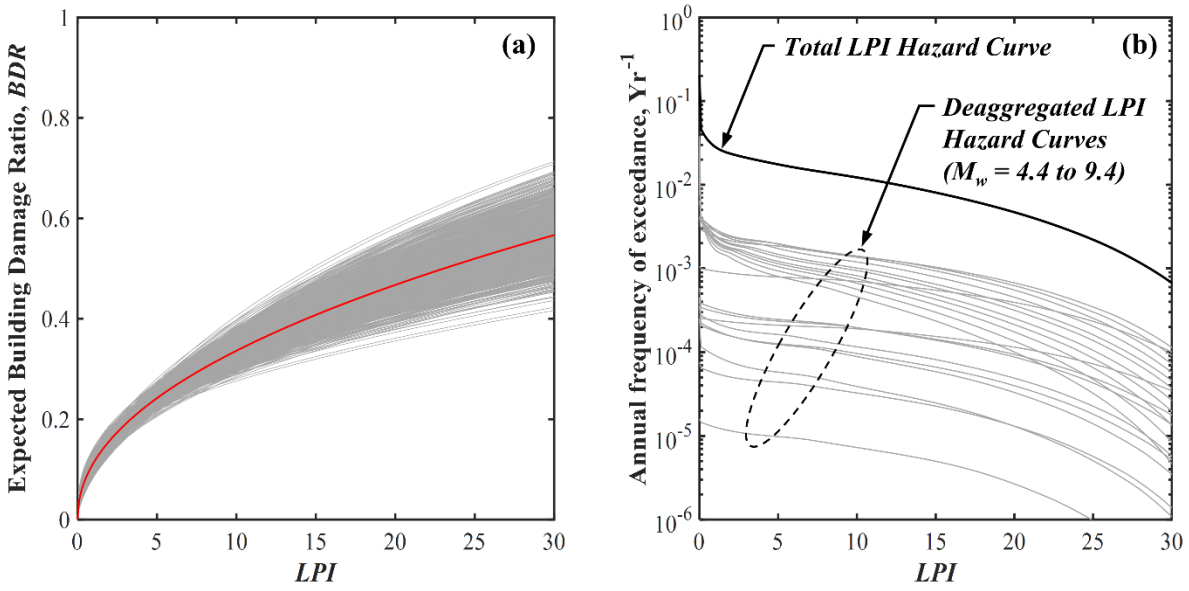


Figure 17. (a) Bootstrapped vulnerability curves (grey) for expected *BDR* as a function of *LPI*; the median relation is shown in red. **(b)** *LPI* hazard curves – for an arbitrary location in Seattle – deaggreaged on earthquake magnitude, such that the total *LPI* hazard curve is equal to the sum of the *LPI* hazard curves for all earthquake magnitudes.

- To account for and visualize finite-sample uncertainty, the underlying dataset will be bootstrapped 1,000 times (i.e., resampled with replacement), with each simulation resulting in a vulnerability function. A building's expected annual loss (*EAL*) due to liquefaction will then be computed using: (i) the developed fragility and vulnerability functions; (ii) CPT data at the building's location; and (iii) the *LPI* hazard curve, which describes the mean annual rate of

LPI exceeding a given value (λ_{LPI}) at the building's location. With this approach, *EAL* is computed by integrating over all *LPI* values:

$$EAL = V \cdot \int_{LPI=0}^{\infty} E[BDR|LPI] \cdot \left|\frac{d\lambda_{LPI}}{dLPI}\right| \cdot dLPI \quad (3)$$

where V is the replacement value of the building, in dollars; $E[BDR|LPI]$ is defined in Equation 5; and $|d\lambda_{LPI}/dLPI|$ is the absolute value of the derivative of the *LPI* hazard curve, which is computed from CPT data and integration with the *PGA* hazard curve, which describes the mean annual rate of *PGA* exceeding a given value (λ_{PGA}). Specifically, to compute a building site's λ_{LPI}, the *PGA* hazard-curve must first be deaggregated by earthquake magnitude. This requirement results from use of both *PGA* and magnitude as necessary inputs for predicting liquefaction triggering (Boulanger and Idriss, 2014). λ_{LPI} is then computed as:

$$\lambda_{LPI} = \sum_{j=1}^{N_{Mw}} \sum_{i=1}^{N_{PGA}} P\left(LPI > lpi \middle| pga_i, m_j\right) \Delta\lambda_{pga_i,m_j} \quad (4)$$

where N_{Mw} and N_{PGA} are respectively the number of magnitude and *PGA* increments into which the computed seismic hazard is subdivided, and $\Delta\lambda_{pgai,mj}$ is the incremental annual-exceedance rate for intensity measure, pga_i, and magnitude, m_j, which follows an established approach in performance-based liquefaction modelling (Kramer and Mayfield 2007).

- To demonstrate this process conceptually, a CPT from the South-of-Downtown (SODO) district of Seattle, USA, is analyzed. Deaggregated *PGA* hazard-curve data was obtained for the CPT's location (47.587130, -122.331487) using the USGS Unified Hazard Tool. Using the above equations, the *LPI* hazard curve was computed and is shown in Figure 17b. Using that

hazard curve within Equation 3, the *EAL* due to liquefaction can be computed for a lightweight structure that: (i) is located at the CPT site coordinates given above; (ii) is built atop a shallow foundation

- To conceptually demonstrate the proposed decision tool, a life-cycle cost-benefit analysis was performed for the site in SODO considering two hypothetical ground improvement strategies and a building with replacement cost of $500,000. For simplicity, these hypothetical strategies either double or quadruple the normalized, fines-corrected CPT tip resistance ($q_{c1N\text{-}cs1}$) at the ground surface, with both improvements decreasing linearly to a depth of 20 m, where no improvement is achieved. These strategies are henceforth referred to as Mitigation Options 1 and 2, respectively, where the former has an expected upfront cost of $30,000 and the latter (which achieves more improvement) has an expected upfront cost of $50,000. The relative efficiencies of these intervention strategies can be assessed and compared using various economic decision-modeling tools. As an example, the cost-efficiency of a strategy may be defined by its cost-benefit difference (*CBD*), such that a *CBD* less than zero indicates an auspicious decision:

$$CBD = Cost - Benefit = Cost - (EAL_o - EAL_m)\left(\frac{1-e^{-rt}}{r}\right) \quad (5)$$

where *Cost* is the upfront expense of the strategy; *Benefit* is the reduction in loss obtained from the strategy; *EAL* is the expected annual loss from liquefaction; subscript *o* refers to the original state; subscript *m* refers to the mitigated state (e.g., with ground improvement); *t* is the planning horizon, or service life, of the infrastructure asset, in years; and *r* is an interest rate for determining the present value of future cash flows. An investment in remediation thus makes

sense if the service life of the asset exceeds the time required for *CBD* to reach zero. Moreover, negative *CBD*'s convey the time-dependent net benefit resulting from an up-front investment in remediation. To compute *CBD* values for each intervention strategy, *LPI* hazard curves are first computed for each CPT following the approach taken previously, the results of which are shown in Figure 18a.

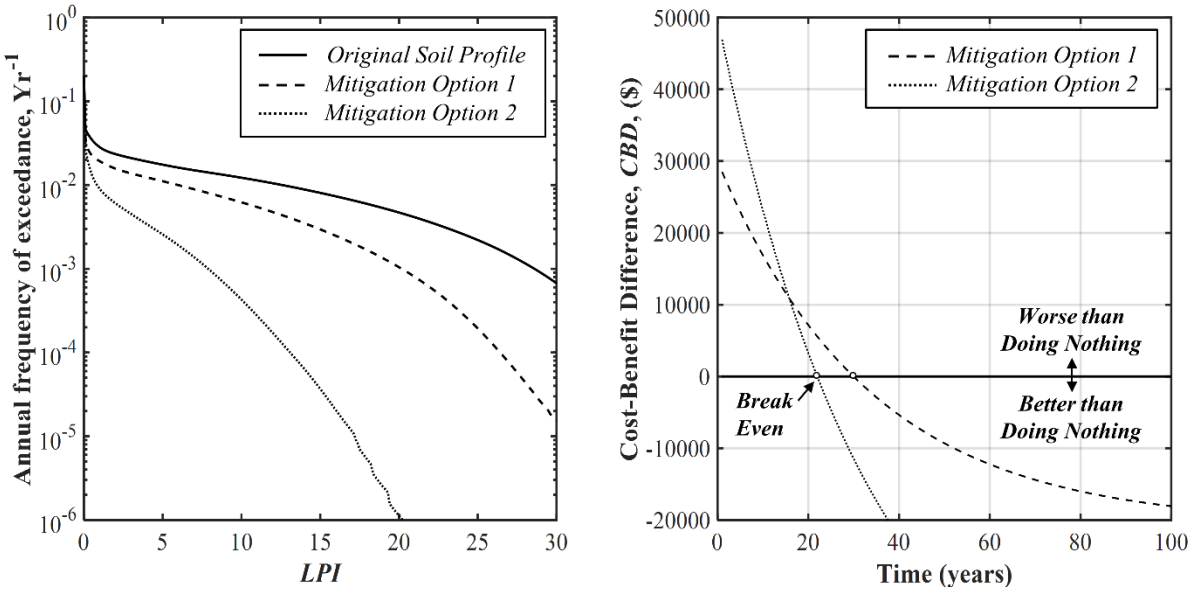


Figure 18. (a) *LPI* hazard curves for the SODO study site, with and without two different ground-improvement methods for liquefaction mitigation; **(b)** Effect of planning-horizon on cost-benefit-difference analysis of ground improvement.

- Thus, as illustrated in Figure 18a, mitigation strategies are detected and incorporated into the analysis through changes in λ_{LPI}, such that the annual exceedance rate of any *LPI* value is reduced due to changes to the soil profile (in this case, increased penetration resistance). By corollary, the *EAL* due to liquefaction will also diminish, per Eq. (3). In this case, the modified *EAL*s due to liquefaction are reduced from \$6,528 to \$5,010 for Option 1, and to \$3,422 for Option 2. Using these results, and assuming an interest rate of

3%, *CBD* values are computed and plotted in Figure 18b for Mitigation Options 1 and 2. It can be seen that the more costly of the two methods – Option 2 – reaches its break even point sooner than Option 1. In this case, Option 2 pays for itself through reductions in *EAL* after approximately 21 years, whereas Option 1 breaks even after approximately 30 years. This modelling approach can be used to decide whether, and how, liquefaction should be mitigated, or to demonstrate the economic feasibility of a new ground improvement technology. For example, given expected (or actual) ground-improvement costs and outcomes, the economic feasibility of any bio-inspired or bio-mediated technology can be readily demonstrated. Ultimately, the economic feasibility of any liquefaction mitigation strategy is dependent on the interactions of numerous location-, profile- and asset-specific factors. However, this study will present a tool for studying these factors in a consistent, practical, and rational manner.

Chapter 5. HORIZON: LIQUEFACTION OCCURRENCE AND CONSEQUENCE MODELLING SOFTWARE

Thus far, the dissertation has produced several chapters in which each chapter is a prerequisite for the one that follows. Chapter two compiled an unprecedented quantity of liquefaction case histories. Chapter three tested and developed various models for predicting liquefaction and its manifestation at the ground surface. Chapter four developed and demonstrated a pragmatic framework by which analysts can decide whether, and how, to mitigate liquefaction beneath lightweight structures on shallow foundation systems. To synthesize these various efforts, and to facilitate user adoption, Chapter 5 develops *Horizon*, a windows-executable software program wherein all previously mentioned analytics and models are coded.

5.1 PUBLICATION 10: GEYIN M AND MAURER BW (2021). "HORIZON: CPT-BASED LIQUEFACTION RISK ASSESSMENT AND DECISION SOFTWARE." DESIGNSAFE CI. DOI: 10.17603/DS2-2FKY-TM46

Horizon is a windows-executable software for analyzing the occurrence, consequence, mitigation, and economics of seismically induced soil liquefaction. *Horizon* improves upon commercially available software, but is freely available and open source, while also facilitating user adoption of new models resulting from this research. Salient features include:

- CPT processing: Inverse filtering and interface correction via the Boulanger and DeJong (2018) procedure; CPT sensor alignment via statistical cross-correlation. User interface of this feature is shown in Figure 19.

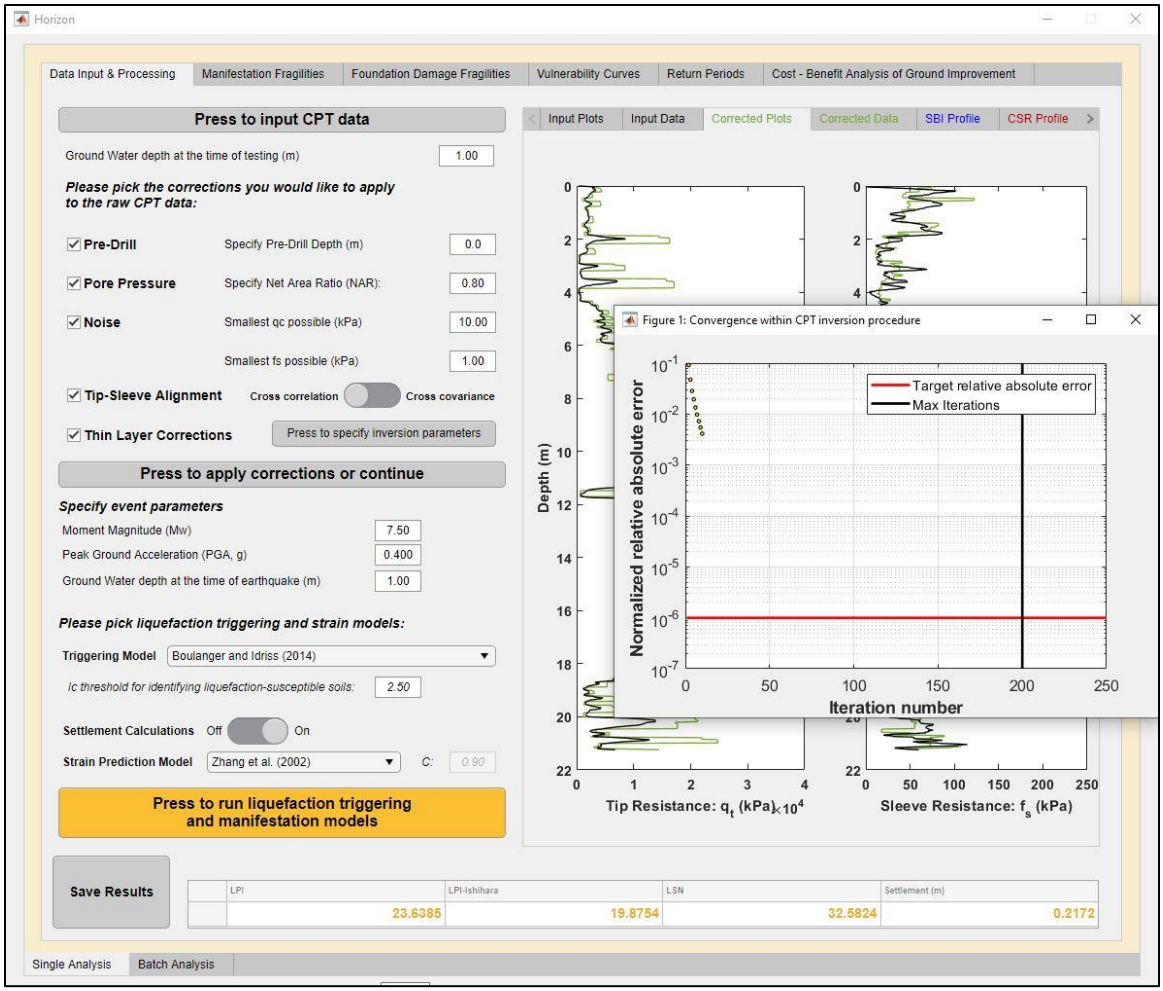


Figure 19. Horizon interface: data input and processing tab.

- Six CPT-based liquefaction triggering models; four free-field manifestation models; fragility and vulnerability models for lightweight structures on shallow foundation systems (Figure 20).

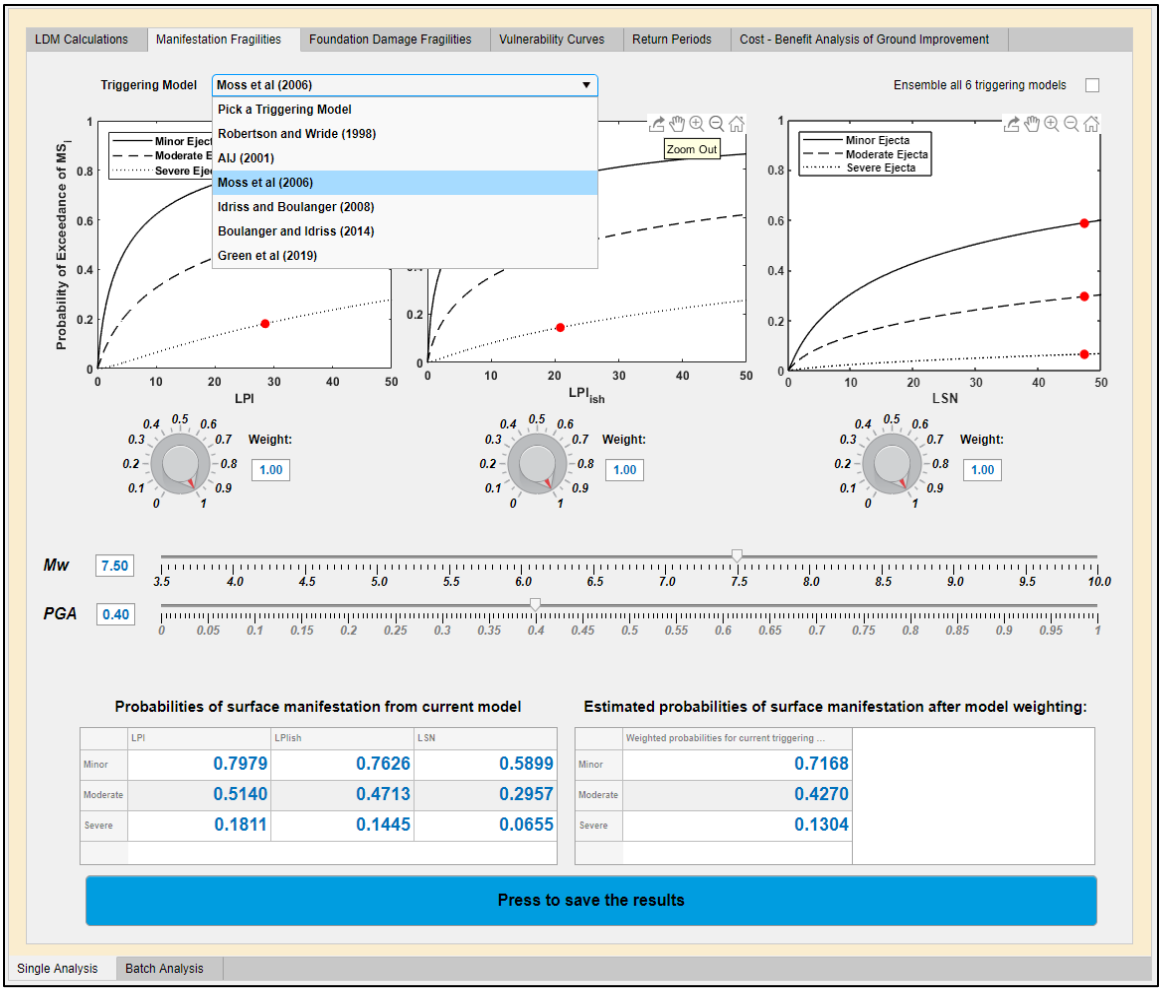


Figure 20. Horizon interface: manifestation fragilities with 6 different liquefaction manifestation models.

- Seamless integration with USGS probabilistic seismic hazard analysis (PSHA) framework; Compute return periods of free-field manifestation, foundation damage, and economic loss (Figure 21).

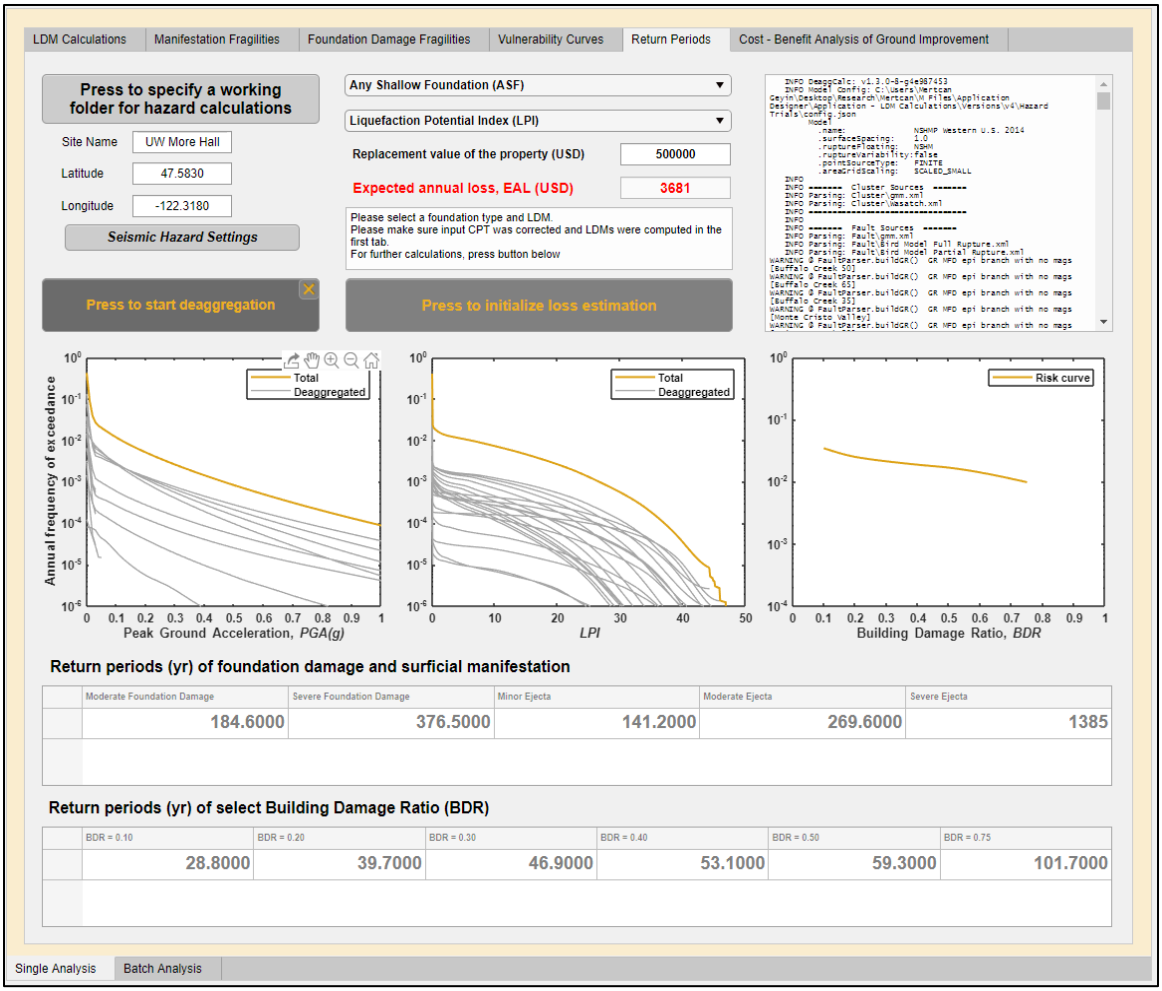


Figure 21. Horizon interface: encoded NSHMP-HAZ for PSHA framework.

- Cost-benefit analytics for deciding whether/how to mitigate liquefaction risks; Quantify value of planned or constructed mitigation activities, wherein improvement is detected and incorporated via changes to CPT sensing data (Figure 22).

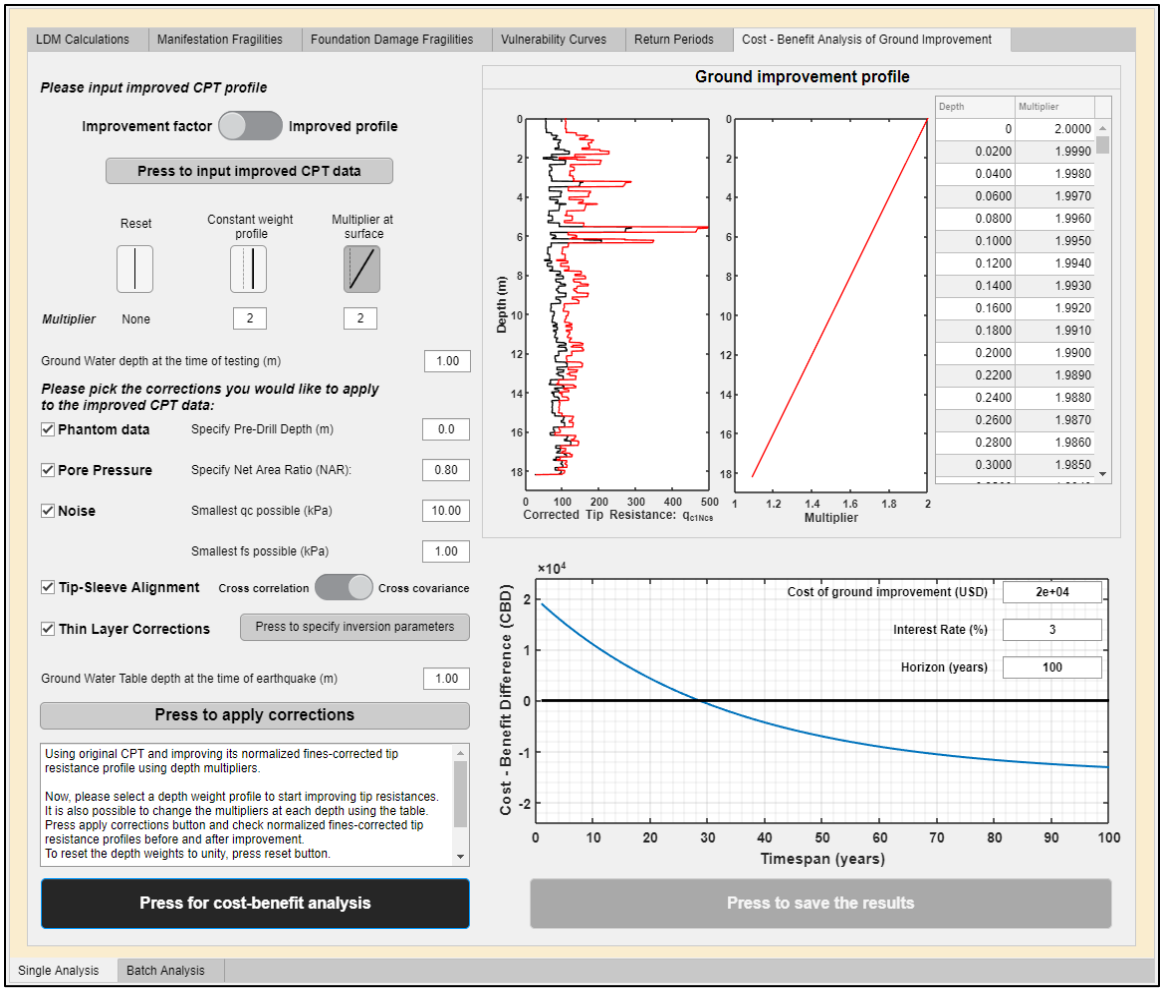


Figure 22. Horizon interface: cost-benefit analyses for liquefaction risk mitigation options using ground improvement. interface incorporates pre- and post-improvement cpt sensing data.

BIBLIOGRAPHY

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

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, University of California, Davis, CA.

Boulanger, R.W. and DeJong, J.T. (2018). "Inverse filtering procedure to correct cone penetration data for thin-layer and transition effects." Cone Penetration Testing 2018, Hicks, Pisano, and Peuchen, eds., Delft University of Technology, The Netherlands, 25-44.

Bradtmueller, J.P., McFarland, A.E., and Foley, S.P. (2015). "A comparison of house size and foundation type for U.S. residential homes." In: Proceedings of the 51st ASC Annual International Conference, The Associated Schools of Construction.

Bray, J. D., Sancio, R., Kammerer, A. M., Merry, S., Rodriguez-Marek, A., Khazai, B., Chang, S., Bastani, A., Collins, B., Hausler, E., Dreger, D., Perkins, W.J., 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 Center, UCB, University of Arizona, Washington State University, Shannon and Wilson Inc., and Leighton and Associates.

Bullock, Z., Karimi, Z., Dashti, S., Porter, K., Liel, A. B., & Franke, K. W. (2019). A physics-informed semi-empirical probabilistic model for the settlement of shallow-founded structures on liquefiable ground. Géotechnique, 69(5), 406-419.

CERA (2013) Purpose and scope of the Canterbury geotechnical database. Canterbury Earthquake Recovery Authority, Information Note, Revision D. Christchurch, New Zealand.

Chleborad, A. F., and R. L. Schuster. 1990.Ground failure associated with the Puget Sound region earthquakes of April 13, 1949, and April 29,1965. USGS Open-File Rep. 90-687. Washington, DC: USGS

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.

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

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.

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

Kramer, S.L. (2008). "Performance-Based Earthquake Engineering: Opportunities and Implications for Geotechnical Engineering Practice," Keynote Paper, Fourth Conference on Geotechnical Earthquake Engineering and Soil Dynamics, ASCE, Sacramento, 32 pp.

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.

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

National Research Council (NRC) (2016) State of the Art and Practice in the Assessment of Earthquake

NZGD (2020) New Zealand Geotechnical Database. New Zealand Earthquake Commission. <https://www.nzgd.org.nz/>, (Accessed Jan 01, 2020).

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.

Ramirez, C.M., Liel, A.B., Mitrani-Reiser, J., Haselton, C.B., Spear, A.D., Steiner, J., Dierlein, G.G, Miranda, E. (2012). "Expected earthquake damage and repair costs in reinforced concrete frame buildings." Earthquake Eng and Structural Dynamics. 41:1455-1475.

Robertson, P.K., and Wride, C.E., 1998. Evaluating cyclic liquefaction potential using cone penetration test. Canadian Geotechnical Journal 35(3), 442-459.

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.

Zhang, G., Robertson, P. K., and Brachman, R.W.I., 2002. Estimating liquefaction-induced ground settlements from CPT for level ground. Canadian Geotechnical Journal 39(5), 1168-1180.

Zhu, J., Daley, D., Baise, L.G., Thompson, E.M., Wald, D.J., and Knudsen, K.L., 2015. A geospatial liquefaction model for rapid response and loss estimation. Earthquake Spectra 31(3), 1813-1837.

Zhu, J., Baise, L.G., and Thompson, E.M., 2017. An updated geospatial liquefaction model for global application. BSSA 107(3), 1365-1385.

APPENDIX A

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

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology,* In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

CPT-BASED LIQUEFACTION CASE HISTORIES COMPILED FROM THREE EARTHQUAKES IN CANTERBURY, NEW ZEALAND

Mertcan Geyin,[a)] M.EERI, Brett W. Maurer,[a)] M.EERI, Brendon A. Bradley,[b)] M.EERI, Russell A. Green,[c)] M.EERI, and Sjoerd van Ballegooy[d)]

Earthquakes occurring over the last decade in the Canterbury region of New Zealand have resulted in liquefaction case-history data of unprecedented quantity. This provides the profession with a unique opportunity to advance the prediction of liquefaction occurrence and consequences. Towards that end, this paper presents a curated dataset containing ~15,000 cone-penetration-test-based liquefaction case histories compiled from three earthquakes in Canterbury. The compiled, post-processed data is presented in a dense array structure, allowing researchers to easily access and analyze a wealth of information pertinent to free-field liquefaction response (i.e., triggering and surface manifestation). Research opportunities using this data include, but are not limited to, the training or testing of new and existing liquefaction-prediction models. The many methods used to obtain and process the case-history data are detailed herein, as is the structure of the compiled digital file. Lastly, recommendations for analyzing the data are outlined, including nuances and limitations that users should carefully consider.

a) University of Washington, Seattle, Washington, United States
b) University of Canterbury, Christchurch, New Zealand
c) Virginia Tech, Blacksburg, Virginia, United States
d) Tonkin + Taylor Ltd., Auckland, New Zealand

Introduction

Within the six years following the 4 Sept. 2010 M_w7.1 Darfield earthquake, which triggered widespread liquefaction in the city of Christchurch, New Zealand, and its environs, 21 additional $M_w \geq 5$ earthquakes occurred within ~20 km of the city's center. While *some* liquefaction was observed in at least 10 of these events (Quigley et al., 2013), damaging liquefaction was most notably triggered by ruptures on 22 Feb 2011, 13 June 2011, 23 Dec 2011, and 14 Feb 2016. A comprehensive summary of the first three of these, including tectonic and geologic settings, seismology, and effects, is provided by Quigley et al. (2016). Specific to liquefaction, observed consequences included: damage to low-, mid-, and high-rise structures, resulting in widespread loss of building stock (e.g., Cubrinovski et al., 2011; van Ballegooy et al., 2014a; Bray et al., 2014); failure of water, wastewater, power, and communications networks (e.g., O'Rourke et al., 2014; Kwasinski et al., 2014; Tang et al., 2014); loss of road, rail, bridge, and levee functionality (e.g., Green et al. 2011; Wotherspoon et al., 2011; Cubrinovski et al., 2014); and impairment of port infrastructure (e.g., Chalmers et al., 2013).

The presence of these effects in a major urban center facilitated and motivated the collection of vast amounts of data, including seismologic, hydrologic, geospatial, and geotechnical measurements, much of which was uploaded to the open access Canterbury Geotechnical Database (CERA, 2013), now the New Zealand Geotechnical Database (NZGD, 2020). These bulk, raw ingredients constitute the makings of an unprecedented quantity of liquefaction case histories, which can be used to train or test predictive models. While several "tiers" of liquefaction prediction model exist (Geyin et al., 2020a), most prevalent models in practice are based on in-situ geotechnical tests, among which the cone-penetration-test (CPT) has important advantages (NRC, 2016). Yet, while such models are widely used to predict liquefaction, they have to-date been

trained on relatively modest datasets. For example, the CPT-based liquefaction triggering model of Boulanger and Idriss (2014), when developed, was trained on essentially all published case histories from all earthquakes combined, or 255 datapoints. Accordingly, this study compiles a curated digital dataset of approximately 14,500 – 15,500 CPT-based case-histories from three earthquakes in Canterbury – namely the Sept. 2010 M_w7.1, Feb. 2011 M_w6.2, and Feb. 2016 M_w5.7 earthquakes – with the exact total depending on criteria discussed subsequently. The post-processed data is presented in a structure array (i.e., a single file), allowing researchers to readily access and analyze a wealth of information pertinent to free-field liquefaction response. As shown in Figure 1, this considerably augments the data available for model training and testing (by at least 50x), presenting the profession with a unique opportunity to advance the science of liquefaction prediction.

In the following sections, the methods used to obtain, process, and populate the database are first detailed, with each "datapoint" including: (i) identifying information (e.g., geographic coordinates); (ii) processed CPT data, both with and without inverse-filtering and interface correction (Boulanger and DeJong, 2018); (iii) peak ground acceleration (*PGA*) and earthquake magnitude (M_w); (iv) groundwater table (*GWT*) depth; and (v) the classified occurrence and severity of liquefaction manifestation at the ground surface, with explicit focus on free-field level ground sites. Next, the structure and formatting of the resulting data array are described, and lastly, recommendations for analyzing the data are discussed, including nuances, uncertainties, and limitations that users should carefully consider.

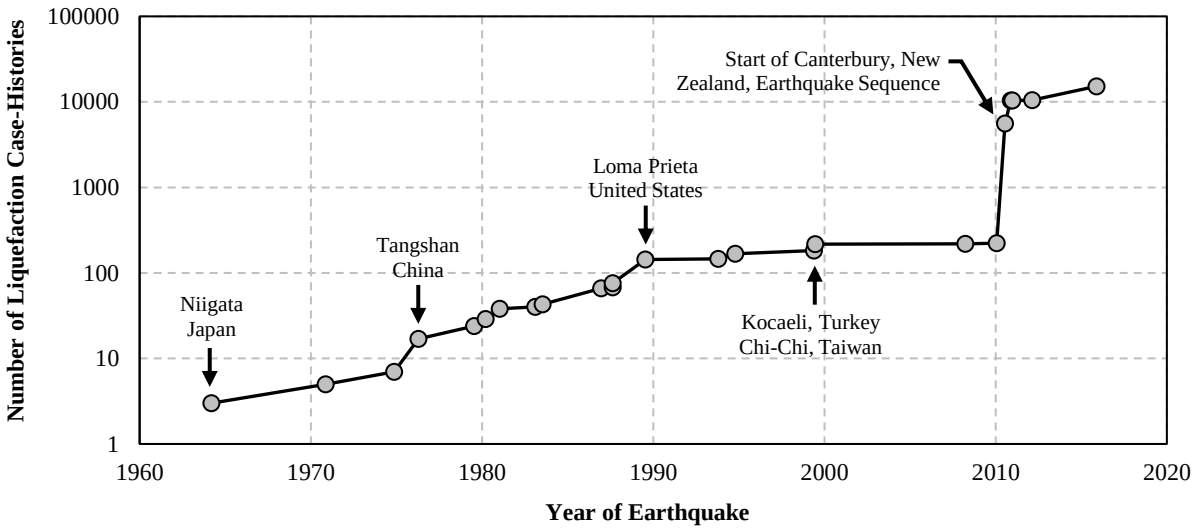


Figure 1. Chronology of CPT-based liquefaction case histories, as compiled by Geyin et al. (2020a).

Methodology

Case histories were compiled from the Sept. 2010 M_w7.1, Feb. 2011 M_w6.2, and Feb. 2016 M_w5.7 earthquakes. This effort built upon successive compilations (Maurer et al., 2014, 2015a), augmenting the largest by more than 50%. While data could potentially also be compiled from the aforementioned events of 13 June and 23 Dec 2011, these events are complicated by the occurrence of multiple, similar-magnitude ruptures only minutes-to-hours apart (Bradley, 2016). As a result, reconnaissance captured the compounded effects of multiple events (complicating observations of response) and pore pressures were elevated at the start of latter events (complicating predictions of response). We thus choose not to present these data, focusing instead on three events without this obfuscating circumstance.

CPT Data and Processing

CPT data was obtained from the New Zealand Geotechnical Database (NZGD, 2020) at sites where liquefaction manifestations could be reliably classified, as discussed subsequently. During this process, CPTs were rejected if inferred from geospatial autocorrelation analyses (Anselin, 1995) to have terminated prematurely (e.g., due to impedance from gravel), such that liquefiable soils potentially exist at greater depth. The local geology of Christchurch is well characterized, with dense, non-liquefiable soils typically found at a certain depth and unlikely to be underlain by looser soils that contribute to liquefaction hazard. In particular, beach, estuarine, and coastal swamp sediments were deposited across Christchurch as sea level rose during the late Pleistocene and Holocene, reaching a peak ~6,500 years before present, with the coastline located 1-2 km west of the present-day city center (Brown et al., 1995). Since then, alluvial deposition has resulted in progradation of the coast to its present location (Brown et al., 1995). Collectively, the deposits resulting from coastline transgression and progradation are known as the Christchurch formation and overlay Pleistocene gravels (i.e., the Riccarton Gravel formation). The terrestrial thickness of the Christchurch formation is greatest beneath the present-day coastline and tapers from east to west, terminating around the mid-Holocene coastal highstand, beyond which the surface geology is characterized by the Springston formation of alluvial gravels and sands (Begg and Jones, 2012). Thus, where the Springston formation dominates (and in some areas of the Christchurch formation), gravelly soils force CPT termination at shallow depth (< 20 m).

Figure 2 maps the expected, surficial geologic units as described in Table 1, and the locations and termination depths of CPT soundings. The termination-depth trends shown in Figure 2b generally agree with the known geologic profile, such that these depths diminish from east to west. While the possibility of liquefiable soils at greater depths exists, it was assumed for this study that their limits are generally defined by CPT termination depths. However, the database was first

parsed using an Anselin Local Morans I analysis (Anselin, 1995) to identify and remove outliers with sounding depths statistically less than the spatial average (i.e., soundings more likely to have prematurely terminated before reaching the Riccarton Gravel formation).

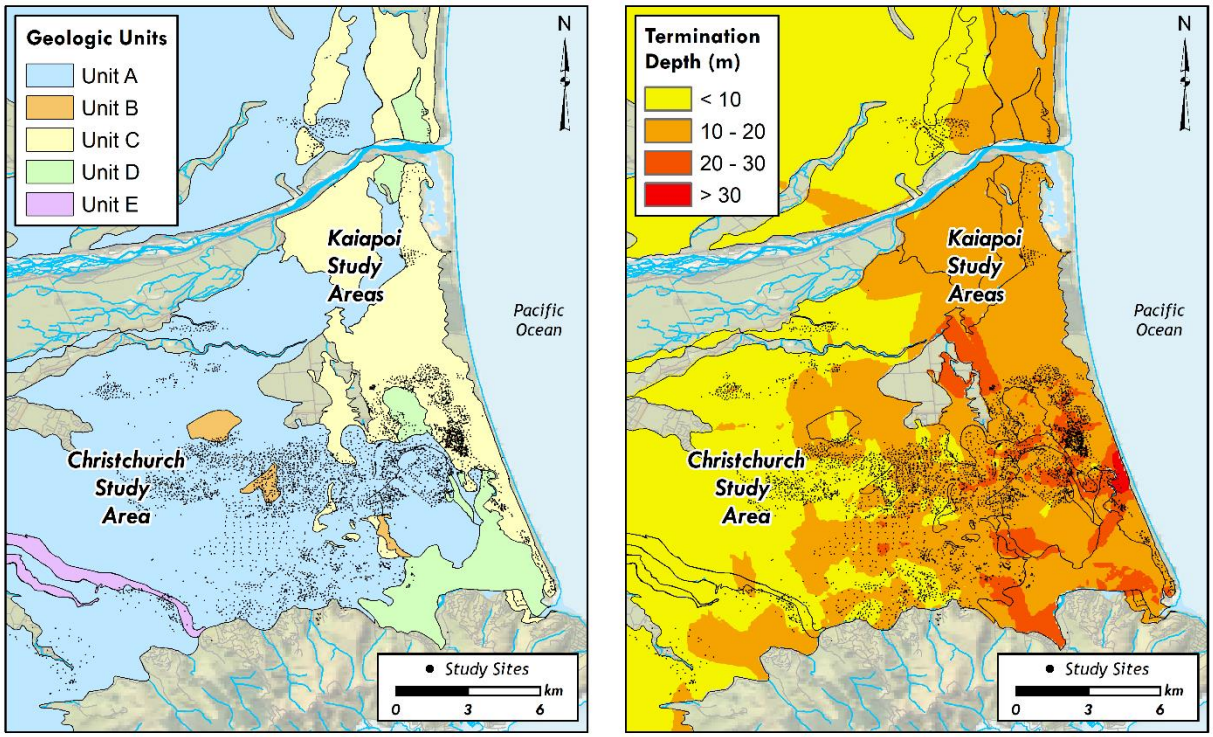


Figure 2. Case-history locations in context of: **(a)** surifical geologic units, as described in Table 1; and **(b)** CPT sounding termination depths, as discussed in the text.

Table 1. Surficial Geologic Units of Study Area.

Geologic Unit	Description	Source	Percent (%) of CPTs in Unit
A	Alluvial sand and silt of overbank deposits	Brown (1975)	59.39
B	Peat swamps now drained	Brown & Weeber (1992)	2.24
C	Fixed dune sand and beach deposits	Brown (1975)	35.34
D	Saline sand, silt and peat of drained lagoons and estuaries	Brown & Weeber (1992)	2.26
E	Fluviatile gravel, sand, and silt of historic river flood channels	Brown & Weeber (1992)	0.43

While the CPT offers advantages among in-situ tests used to predict liquefaction, it is still limited by the volume of soil mobilized around the cone. As an intermediate-to-large-strain penetration test, this mobilized zone acts as a physical "low-pass filter" on the true soil stratigraphy, removing information from the low spatial wavelengths, such as the data defining a thin soil stratum or the interface between two disparate soils. These spatial smoothing effects, which are commonly referred to as "thin layer" and "transition" effects, have long been recognized and studied (e.g., Treadwell 1976; Lunne et al. 1997; Ahmadi and Robertson 2005; Robertson 2011; van der Linden 2018). While chart-based methods exist for manually correcting these effects on CPT data, Boulanger and DeJong (2018) proposed the first programmable procedure. This methodology, referred to as an "inverse filtering and interface detection" procedure, predicts the "true" CPT profile from measured CPT values. Since these measured values reflect a filtered view

of reality, their correction would improve subsurface characterization. As a demonstration of the methodology, CPT data from Christchurch is shown in Figure 3, both with and without correction.

While the performance of Boulanger and DeJong's (2018) procedure is currently being evaluated (e.g., Yost et al. 2020), its use can change a site's perceived liquefaction hazard, with the direction and magnitude of change dependent on numerous factors. Considering this potential influence, and that the Boulanger and DeJong (2018) procedure might prove to be efficacious, both measured and "true" CPT data are provided in the database. While the reader is referred to Boulanger and DeJong (2018) for complete details, the procedure's "baseline" parameters were used to compute "true" CPT data. This was the case for both the subroutine that inverts tip resistance and sleeve friction, and that which detects and corrects stratigraphic interfaces. These defaults can conceivably be calibrated via site-specific study (e.g., from borings adjacent to a CPT), but the information compiled for this study either was insufficient to attempt calibration or provided insufficient statistical support to justify it. As part of the processing methodology, CPT tip- and sleeve-measurements were aligned using statistical cross-correlation (Buck et al. 2002), both for measured and "true" CPT data. In addition, CPT data was infilled in the "pre-drill" zone (i.e., where borings were used to safety bypass pavements or utilities, most often to a depth of ~1 m where applicable (~40% of CPTs were pre-drilled)). In the absence of this correction, the recorded data is that of noise as the cone penetrates an open boring. Accordingly, CPT data was sampled 15 cm beyond the recorded depth of pre-drill, then uniformly applied to the pre-drill interval. While this provides reasonable data for approximating soil unit weights, and by corollary, in-situ stresses below the pre-drill zone, users should consider the relative depths of pre-drill and groundwater when analyzing case histories, as further discussed herein. As part of this process, CPTs with unknown pre-drill depth were preemptively removed from the dataset, as were CPTs

with pre-drill depth exceeding 2.5 m. All CPT processing was completed using the software *Horizon* (Geyin and Maurer, 2020a), a freely-available program developed by the authors.

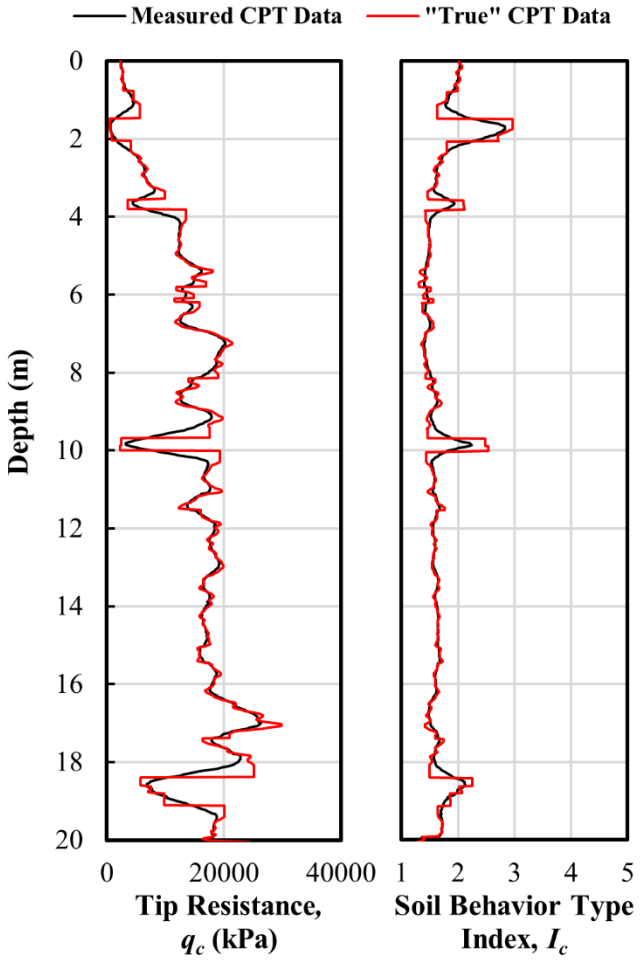


Figure 3. Example CPT data with (i.e. "true") and without (i.e. "measured") inverse-filtering and interface correction via the Boulanger and DeJong (2018) procedure, as implemented in the software *Horizon* (Geyin and Maurer, 2020a).

Liquefaction Manifestations

Emphasis was placed on compiling case histories from free field level-ground sites, with the occurrence and severity of surface manifestation defined primarily by liquefaction ejecta. In this

respect, sites with other indicators of liquefaction (e.g., evidence from ground-motions or foundation settlements) were expressly omitted. While ~7% of case histories were characterized by a predominance of lateral spreading, the majority were compiled from level-ground sites. In particular, surface manifestations were observed at CPT sites following at least one of the three aforementioned earthquakes and manually classified by the authors as "none," "marginal," "moderate," "severe," "lateral spreading," or "severe lateral spreading" using criteria modified from Green et al. (2014) and given in Table 2; the identifying codes assigned herein to each classification are also provided. This was accomplished using high-resolution satellite imagery and reconnaissance reports available in the New Zealand Geotechnical Database (NZGD, 2020). Classifications were based on a circular sample area, centered on each CPT, with approximate radius of 10 m. Representative observations of manifestation classifications are provided in Appendix A. Sites where surface manifestations could not be reliably classified following an event are denoted as "unknown" and coded "10" (i.e., sites where manifestations were classified following at least one earthquake, but not all three, which was the case for ~18% of study sites). Of the resulting 15,890 compiled case histories, 61% were classified as "none" and 39% are cases in which manifestations were observed and classified in accordance with Table 2. Owing to nuances that will be discussed subsequently, the quantity of data best suited for model training and testing is ultimately reduced to ~14,500 – 15,500 cases.

Table 2. Criteria used to classify liquefaction manifestations (after Green et al. 2014).

Classification	Severity ID	Criteria
None	0	No observed liquefaction ejecta or lateral spreading
Minor	1	Small, isolated liquefaction features less than a vehicle width; <5% of ground surface is covered by ejecta; no lateral spreading.
Moderate	2	Groups of liquefaction features greater than a vehicle width; 5-40% of ground surface is covered by ejecta; streets are generally passable; no lateral spreading.
Severe	3	Adjoining large liquefaction features that are greater than a vehicle width; >40% of ground surface is covered by ejecta; streets are generally impassable; no lateral spreading.
Lateral Spreading	4	Ejection of liquefied material at the ground surface may be observed, but lateral spreading is the predominant manifestation and damage mechanism. Measured crack-displacement widths are less than 200 mm.
Severe Lateral Spreading	5	Ejection of liquefied material at the ground surface may be observed, but lateral spreading is the predominant manifestation and damage mechanism. Measured crack-displacement widths exceed 200 mm.
Unknown	10	Insufficient information to reliably classify: out of bounds, no reliable documentation, obscured or otherwise ambiguous imagery.

Hydrologic Data

GWT depths at CPT locations were obtained from the time-dependent models of van Ballegooy et al. (2014b). These models, which reflect seasonal and local fluctuations across the region, were derived in part using long-term monitoring data from a network of ~1000 monitoring wells and

provide a best estimate of *GWT* depths at the time of each earthquake. Well measurements were corrected for elevation changes caused by the earthquakes using digital elevation models derived from ground-based surveys and airborne LiDAR (Light Detection and Ranging). River and coastline data were used to shape and position *GWT* contours at places of significant groundwater-surface water interaction (van Ballegooy et al., 2014b). The median *GWT* diminishes from 10+ m elevation (relative to sea level) west of Christchurch to less than 1 m elevation in the eastern suburbs (i.e., near the coast), roughly consistent with the change in ground elevation. The *GWT* depth is generally 1-2.5 m beneath much of the study area but reaches 5 m west of the city center. A histogram of *GWT* depth for all compiled case histories is shown in Figure 4a.

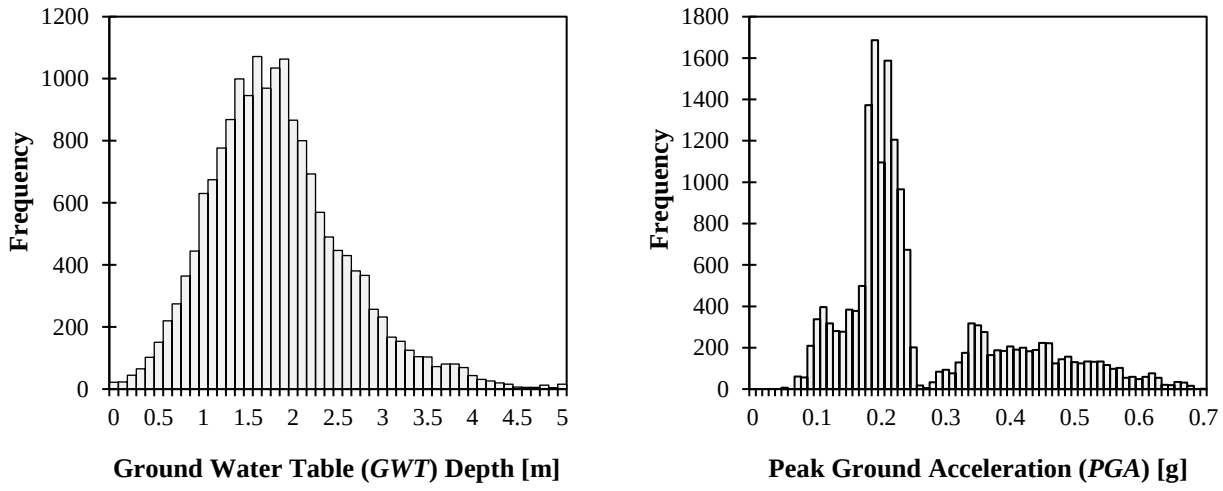


Figure 4. Histograms of: **(a)** ground water table (*GWT*) depth; and **(b)** peak ground acceleration (*PGA*) for case histories compiled in the curated dataset.

Peak Ground Accelerations (PGAs)

PGA is to-date the most common ground-motion intensity measure (IM) for quantifying seismic demand in liquefaction models. Among other standard IMs, it has been shown to be the most

efficient predictor of pore-pressure generation and the initiation of liquefaction (Sideras, 2019). In this study, *PGAs* were estimated at CPT sites via the Bradley (2014) method, which has been used widely in research related to the Canterbury earthquakes (e.g., van Ballegooy et al. 2015; Geyin et al. 2020a; Geyin and Maurer 2020b). This method geostatistically coalesces instrumentally recorded *PGA*s with predictions from ground-motion models (GMMs), where the former were recorded by more than 20 near-source strong-motion stations (SMS) (e.g., Bradley and Cubrinovski, 2011; Bradley, 2012). Using this approach, the *PGA* at SMS, *i*, is expressed as:

$$\ln(PGA_i) = \mu_{\ln PGA_i}(\text{Site, Rup}) + \eta + \varepsilon_i, \quad (1)$$

where $\ln(PGA_i)$ is the natural logarithm of the observed *PGA* at SMS *i*; $\mu_{\ln PGA_i}(\text{Site, Rup})$ is the mean of the natural logarithm of *PGA* at SMS *i* predicted by a GMM, which is a function of site and rupture parameters; η is the inter-event residual; and ε_i is the intra-event residual. Within Equation 1, a GMM predicts a *PGA* distribution:

$$\ln(PGA_i) \sim N(\mu_{\ln PGA_i}, \sigma_\eta^2 + \sigma_\varepsilon^2), \quad (2)$$

where $X \sim N(\mu_X, \sigma_X^2)$ is shorthand notation for *X* having a normal distribution with mean, μ_X, and variance, σ_X^2. By definition, all *PGA*s recorded in a given earthquake have the same inter-event residual, η. Conversely, the intra-event residual, ε_i, varies from site to site but is correlated spatially due to similarities in path and site effects. Accordingly, *PGA*s at SMS locations can be used to compute conditional distributions of *PGA*s at CPT locations. First, the Bradley (2013) New Zealand GMM was used to compute the unconditional distribution of *PGA*s at SMS locations. A mixed-effects regression was then used to determine the inter-event residual, η, and the intra-event residuals, ε_i's, for each strong-motion station (e.g., Abrahamson and Youngs 1992; Pinheiro et al. 2008). Second, the covariance matrix of intra-event residuals was computed by accounting for the

spatial correlation between SMS locations and a test site of interest. The joint distribution of intra-event residuals at a site of interest and the SMS is given as:

$$\begin{bmatrix} \varepsilon^{site} \\ \boldsymbol{\varepsilon^{SMstation}} \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\varepsilon^{site}} & \boldsymbol{\Sigma_{12}} \\ \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{bmatrix} \right), \tag{3}$$

where $X \sim N(\boldsymbol{\mu_X}, \boldsymbol{\Sigma})$ is shorthand notation for X having a multivariate normal distribution with mean $\boldsymbol{\mu_X}$ and covariance matrix $\boldsymbol{\Sigma}$ (i.e., same as above, but in vector form, with bold denoting vectors or matrices); and $\sigma^2_{\varepsilon^{site}}$ is the variance of the intra-event residual at the site of interest. In Equation 3, the covariance matrix has been expressed in a partitioned fashion to elucidate the subsequent computation of the conditional distribution of ε^{site}. The individual elements of the covariance matrix were computed from:

$$\boldsymbol{\Sigma}(i, j) = \rho_{i,j}\, \sigma_{\varepsilon i} \sigma_{\varepsilon j} \tag{4}$$

where $\rho_{i,j}$ is the spatial correlation of intra-event residuals between two locations i and j, and $\sigma_{\varepsilon i}$ and $\sigma_{\varepsilon j}$ are the standard deviations of the intra-event residual at locations i and j. Based on the joint distribution of intra-event residuals given by Equation 3, the conditional distribution of $\varepsilon^{\text{site}}$ was computed from Johnson and Wichern (2007):

$$\begin{aligned} \left[\varepsilon^{site} | \boldsymbol{\varepsilon^{SMstation}}\right] &= N(\boldsymbol{\Sigma_{12}} \cdot \boldsymbol{\Sigma_{22}^{-1}} \cdot \boldsymbol{\varepsilon^{SMstation}}, \sigma^2_{\varepsilon^{site}} - \boldsymbol{\Sigma_{12}} \cdot \boldsymbol{\Sigma_{22}^{-1}} \cdot \boldsymbol{\Sigma_{21}}) \\ &= N(\mu_{\varepsilon^{site} | \, \varepsilon^{SMstation}}, \sigma^2_{\varepsilon^{site} | \, \varepsilon^{SMstation}}) \end{aligned} \tag{5}$$

Using the conditional distribution of the intra-event residual given by Equation 5 and substituting into Equation 2, the conditional distribution of *PGA* at a site of interest, PGA_{site}, is:

$$[ln\, PGA_{site} \mid ln\, PGA_{SMstation}] = N(\mu_{\ln PGA_i} + \eta + \mu_{\varepsilon^{site} | \, \varepsilon^{SMstation}}, \sigma^2_{\varepsilon^{site} | \, \varepsilon^{SMstation}}) \tag{6}$$

That is, the conditional distribution of *PGA* is a lognormal random variable completely defined by the conditional median and conditional uncertainty (i.e., lognormal standard deviation). Intuitively, in cases where a CPT is located far from any SMS, the conditional distribution (i.e., final estimate of *PGA*) is similar to the unconditional distribution (i.e., GMM estimate of *PGA*), and for a CPT very near to a SMS, the conditional distribution approaches the value observed at the SMS. The conditional median and conditional uncertainty of *PGA*, both of which are given in the dataset, were computed at CPT sites using the spatial correlation model of Goda and Hong (2008). A histogram of the median *PGA* values for all case histories is shown in Figure 4b.

One benefit of the adopted approach is that uncertainty is explicitly computed, rather than subjectively assigned, which prior compilations and studies of liquefaction case-history data have been resigned to. With respect to the computed values, GMMs generally have standard deviations of ~0.45-0.55 in the natural log of *PGA*, though it can be larger in some cases (e.g., Bradley 2013). In the compiled database, this parameter varies from 0.045 to 0.55, with a median of ~0.33. Thus, the *PGA* uncertainty is generally much less than that which would be obtained using a GMM alone. Any percentile, *x*, of the conditional *PGA* distribution may be computed as:

$$PGA_x = PGA_{50} * \exp\left(n\sigma_{lnPGA}\right) \tag{7}$$

where PGA_x is the value of *PGA* for the x^{th} percentile, 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, which is the number of standard deviations from the median. As an example, the 16th and 84th percentiles of a *PGA* distribution are computed using n values of -1 and 1, respectively, and the PGA_{50} and σ_{lnPGA} values computed for each case history. Maps of these values are available in the New Zealand Geotechnical Database (NZGD, 2020). Additional information describing the preceding methodology is provided by Bradley and Hughes (2012).

Data Structure

The compiled post-processed data is presented in a structure array (i.e., a single file), with each case history including identifying information (e.g., ID, geographic coordinates); CPT data, both with and without inverse-filtering and interface correction; earthquake magnitude (M_w); the median and uncertainty (i.e., lognormal standard deviation) of the conditional *PGA*; *GWT* depth; and the classified occurrence and severity of surficial liquefaction manifestation.

The curated dataset is available via the NHERI DesignSafe Cyber-Infrastructure data depot at Geyin et al. (2020b) (https://doi.org/10.17603/ds2-tygh-ht91) and provided in both Matlab data format and as a python data frame. The data fields, classes, contents, and their units are described in Table 3 and its accompanying footnotes. The structure of the data array is depicted in Figure 5 and arranged such that case histories are principally sorted by a CPT identification (ID) number, wherein multiple liquefaction case histories may be accessed. Event-specific data fields (e.g., M_w and *PGA*) are 3 x 1 arrays containing information from the 2010, 2011, and 2016 earthquakes, respectively. CPT measurements, which are depth-dependent but not event-specific, are l_i x 1 arrays, where l_i is the length of CPT i. Other fields include CPT ID, geographic coordinates, pre-drill depth, and test date. To provide users with a benchmark against which data re-use scripts may be verified, all relevant data is provided for one case-history site in Appendix B. Recommendations for analyzing these data are next discussed, including important nuances and limitations that users should consider prior to analysis.

Table 3. Data fields, typologies, descriptions, and units.

Field	Class	Content
ID	cell	[1]CPT identifier assigned in the field (non-unique)
CPTname	char	CPT name assigned by the current authors (unique)
FILEname	char	File name from which CPT data was obtained (unique)
Date	datetime	Date CPT conducted
NorthingNZMG	double	NZMG – y coordinate
EastingNZMG	double	NZMG – x coordinate
NorthingWGS84	double	WGS – y coordinate
Easting WGS84	double	WGS – x coordinate
depth	double	Depth below the ground surface [m]
qc	double	Measured tip resistance [kPa]
qc_inv	double	[2]True tip resistance [kPa]
fs	double	Measured sleeve friction resistance [kPa]
fs_inv	double	[2]True sleeve friction resistance [kPa]
u2	double	CPT pore pressure measurement, if present [kPa]
pd	double	Pre-drill depth [m]
Magnitude	struct	[3]Earthquake moment magnitude (M_w)
PGA	struct	[3]Event-specific conditional median peak ground acceleration [g]
PGAsigma	struct	[3]Event-specific conditional lognormal standard deviation of PGA
GWT	struct	[3]Event-specific groundwater table depth [m]
Manifestation	struct	[3,4]Classified type/severity of surface manifestation

[1]CPT IDs from original CPT data (.csv or .xlsx files), which were recorded by field engineers and are non-unique in the dataset.
[2]Processed per the Boulanger and DeJong (2018) procedure; specifically, using the "baseline" inversion model.
[3]There are three earthquakes within the fields from which event-specific data is compiled: M_w7.1 4 Sept 2010 (Yr2010), M_w6.2 22 Feb 2011 (Yr2011), and M_w5.7 14 Feb 2016 (Yr2016).
[4]The occurrence/severity of surface manifestation was manually classified for each CPT location in each earthquake per the criteria in Table 2. Classifications are based on a circular sample area, centered on each CPT, with approximate radius of 10 m.

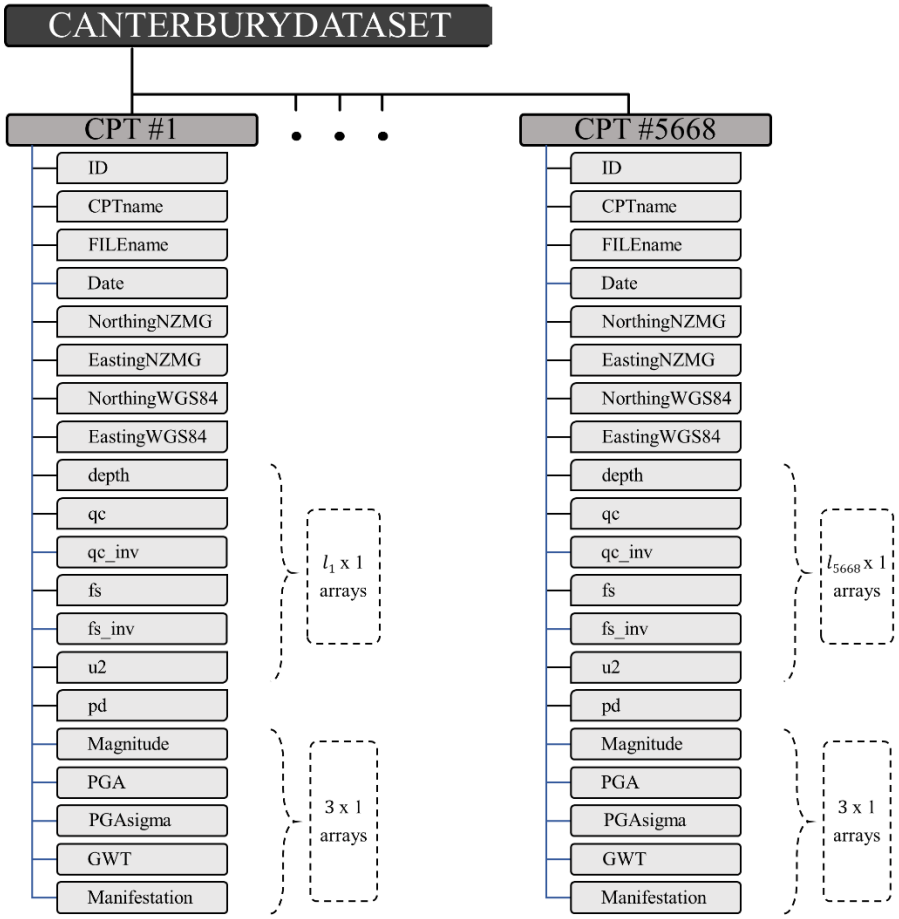


Figure 5. Depiction of the Canterbury case-history dataset structure array.

Discussion: Data Nuances, Alternatives, and Analysis

The compiled, post-processed data allows researchers to easily access and analyze a wealth of information pertinent to in-situ site characterization and free-field liquefaction response (i.e., triggering and surface manifestation). Research opportunities using this data include, but are not limited to: (i) methods for quantifying/simulating subsurface spatial variability; (ii) training or testing of new and existing liquefaction-prediction models; (iii) temporal assessment of CPT data during shaking sequences, including use of aging-correction factors for liquefaction prediction (i.e., K_{DR}); and (iv) evaluation of CPT inversion filters in the context of liquefaction model performance. Prior to such analyses, however, users should

carefully consider several important data nuances, alternatives, and limitations. These topics are discussed as follows and ordered by: (i) additional data exclusion criteria; (ii) alternative sources of data; (iii) correlations and decisions for analysis; and (iv) lingering uncertainties.

Additional Data Exclusion Criteria

GWT and CPT Pre-drill Depths. As discussed previously, CPTs in which the pre-drill depth exceeded 2.5 m were preemptively removed from the dataset. For the remainder of CPTs, the pre-drill interval (typically ~1 m when present) was infilled with CPT data from just below the pre-drill (i.e., where the sensors began penetrating undisturbed soil). While this provides reasonable data for estimating in-situ stresses, users performing liquefaction studies should consider the relative depths of pre-drill and groundwater. Case histories in which the depth of pre-drill exceeds that of the groundwater have additional uncertainty, given that CPT data below the expected water table is extrapolated, rather than measured. A histogram of the *GWT* depth minus pre-drill depth is shown in Figure 6 for the 15,890 compiled case histories. Of these, 1,503 have pre-drill depth exceeding the *GWT* depth. However, this differential exceeds 0.5 m for just 420 cases, and exceeds 1 m for just 52 cases. Nonetheless, analysts might exclude some or all such cases to avoid near-surface site-characterization uncertainty. Some 1D liquefaction manifestation models, for example, are especially sensitive to this uncertainty owing to depth-weighting functions (e.g., Ballegooy et al., 2014a; Maurer et al., 2015b).

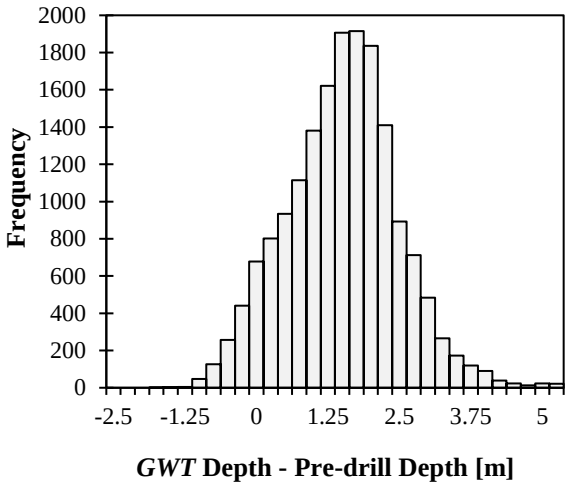


Figure 6. Histogram of ground water table (GWT) depth minus pre-drill depth, for compiled case histories.

PGAs. Profiles subjected to a *PGA* less than the expected threshold for inducing pore pressure (Dobry et al. 1982) might not provide meaningful data for testing liquefaction analytics. That is, if the expected peak strain is less than the volumetric threshold shear-strain for a very loose soil, the absence of liquefaction is easily predicted by judgment. Using such cases to test liquefaction models could thus increase the prediction efficiency in a misleading manner. Accordingly, analysts might select a site-specific *PGA* threshold for excluding data, or a general threshold considering the most susceptible soil that could be encountered (e.g., de Magistris et al., 2013). Considering all compiled cases, 0.096 g was the lowest *PGA* for which surface manifestation of liquefaction was observed, albeit this is distinctly different from that which may induce pore pressure at depth. Of the 15,890 compiled case histories, 98 have *PGA* less than 0.075 g (see Figure 4b).

Model Applicability. Lateral spreading is a distinct manifestation of liquefaction influenced by topographic factors that were not compiled in this study, but which could be sampled (e.g., ground slope, distance to free face). However, users should consider whether such case histories are appropriate for model training or testing. As an example, 1D liquefaction manifestation models

may not fully account for the factors known to cause lateral spreading and can thus predict it poorly (e.g., Maurer et al. 2015c; Rashidian and Gillins 2018). Accordingly, for some purposes, it may be most appropriate to exclude such cases from analysis. Of the 15,890 compiled case histories, 1,110 are cases in which lateral spreading was the predominant manifestation of liquefaction. For further coverage of lateral spreading in Canterbury, see Cubrinovski and Robinson (2016).

Alternative Sources of Data

Ground Motions. As previously outlined, *PGA*s were obtained by statistically coalescing strong-motion records with GMM predictions (Bradley 2014), the general concept of which is common, and which could be used to obtain other IMs of interest. Notably, liquefaction likely occurred at some SMS sites during the Canterbury earthquakes, potentially effecting measured *PGA*s and in-turn the adopted approach. In particular, evidence of liquefaction was observed in several SMS records from the 22 Feb 2011 earthquake (e.g., Bradley and Cubrinovski 2011), namely: (i) high-frequency acceleration spikes, inferred to result from cyclic mobility/dilation response (Kramer et al. 2016); and (ii) subsequent reduction in high-frequency motion, inferred to result from the softening of liquefaction (Kramer et al. 2016). One such example is shown in Figure 7. It can be seen that a recorded *PGA*, if associated with a high-frequency dilation spike, could exceed the peak acceleration prior to liquefaction, and possibly, that which would have occurred in its absence (i.e., from the time of liquefaction onward). Wotherspoon et al. (2015) identified four such SMS records (station codes CBGS, CCCC, NNBS, and REHS) and proposed reducing *PGA*s to those observed prior to interpreted dilation spikes. Adopting these values within a liquefaction analysis, Upadhyaya et al. (2019) suggested that existing prediction models performed slightly better using the corrected values. However, given that liquefaction obscures SMS records following its onset, the "true" *PGA*s cannot be known. Whereas dilation spikes may inflate the *PGA*, selecting a peak

value prior to any evidence of liquefaction may artificially depress it. Nonetheless, users should be aware of this issue.

Regardless of which *PGA*s are used within the adopted Bradley (2014) approach, they could be less accurate when complex local phenomena are not captured by empirical predictions (e.g., the effects of rupture directivity, basin-generated surface waves, and near-surface stratigraphic and topographic features). In contrast, physics-based simulations can provide insight into these phenomena via explicit modeling of kinematic fault rupture, wave propagation, and the subsurface velocity structure, thereby predicting IM patterns more accurately. Users may thus be interested in the physics-based simulations of Bradley et al. (2017), which predict both common IMs and full acceleration time-series for each of the three earthquakes in the dataset. These may be obtained at case-history coordinates via the SeisFinder (2020) web portal.

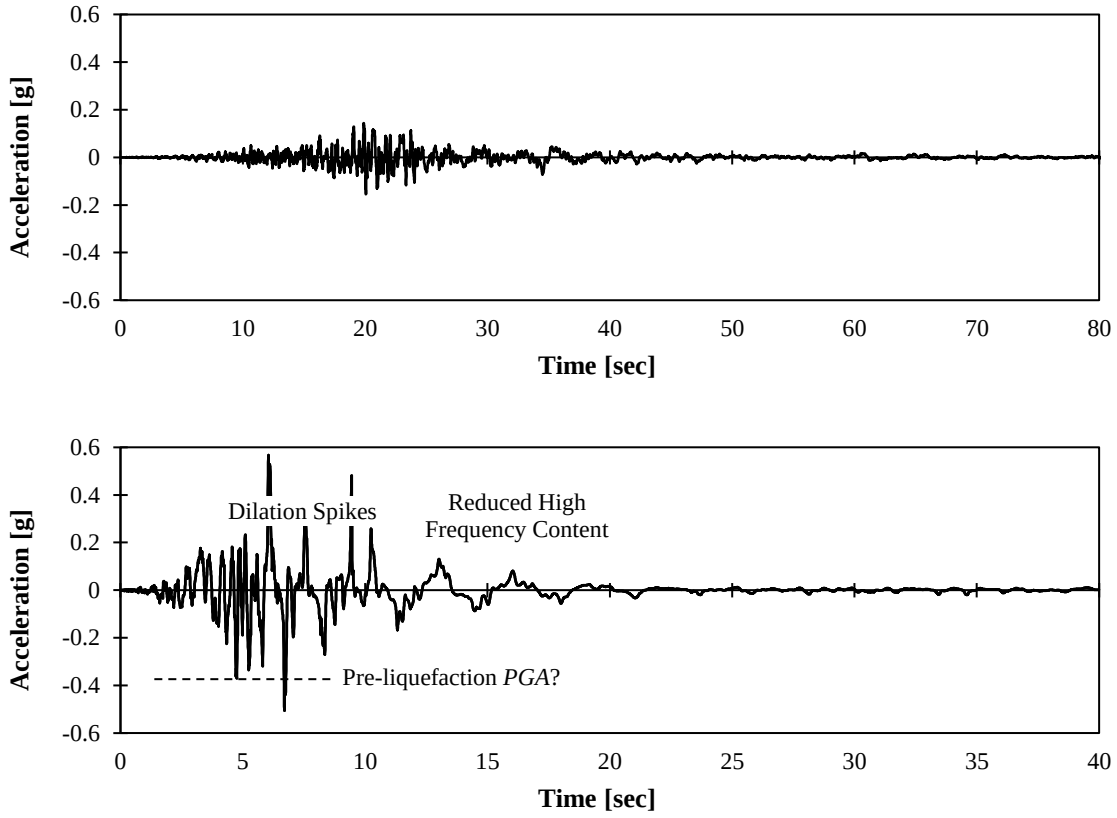


Figure 7. Ground-motion records (SMS code NNBS) during the: **(a)** M_w7.1 Sept 2010 Darfield; and **(b)** M_w6.2 Feb 2011 Christchurch earthquakes, showing the effects of liquefaction on recorded *PGA*s.

Liquefaction Manifestations. As previously discussed, surface manifestations were manually classified at individual CPT sites using a circular sample with 15 m radius. Inherent to this process, which required hundreds-to-thousands of hours to complete, sites where manifestations could not be reliably classified are denoted "unknown." These sites either lacked ground reconnaissance data, were beyond the bounds of high-resolution satellite imagery, had obscured or otherwise ambiguous imagery, or lacked sufficient and consistent information to support classification. Following a similar approach, but without concern for CPT locations or compiling case histories, Townsend et al. (2016) presented maps for the Sept 2010 and Feb 2011 earthquakes wherein

observed manifestations are enclosed by polygons and assigned a confidence rating of "certain," "probable," "possible," or "uncertain." These polygons could supplement/replace the classifications made herein or could provide additional quality control. While Townsend et al. (2016) provide a high-quality dataset, caveats for use include: (i) polygons and ratings are for positive observations only (i.e., they do not explicitly delineate negative observations or assign the confidence therein, although a lack of liquefaction may be inferred where polygons are not present); (ii) the mapping does not classify the severity of liquefaction, which may be useful for model training/testing; and (iii) due to the scale of polygons (e.g., that of building parcels), a polygon may be classified as positive but lack manifestations over some or much of its surface area. For these reasons, the classifications made herein may differ from those of Townsend et al. (2016). Nonetheless, users should be aware of this excellent database and consider its use.

Correlations and Decisions for Analysis

Liquefaction Susceptibility. Existing "simplified stress-based" triggering models (e.g., Robertson and Wride 1998; Moss et el. 2006; Boulanger and Idriss 2014; Green et al. 2019) are not intended to be applied to high plasticity, fine grained, "non-liquefiable" soils, which could result in less accurate predictions of cyclic response, and for which other, more appropriate methods exist (e.g., Boulanger and Idriss 2007). Soils not susceptible to liquefaction triggering are thus generally identified and screened from analysis, consistent with the development of the models. Various criteria based on lab indices have been proposed for this purpose (e.g., Polito 2001; Seed et al. 2003; Bray and Sancio 2006; Boulanger and Idriss 2006), an overview of which is given by Green and Ziotopoulou (2015). However, while soil samples may be obtained using a CPT push-sampler, continuous sampling and testing to assess susceptibility is prohibitively expensive. For this reason, the CPT soil-behavior-type index (I_c) proposed by Jeffries and Davies (1993) and modified by

Robertson and Wride (1998) is generally used to assess susceptibility by way of correlations with lab criteria. For example, an $I_c = 2.6$ threshold is common, such that soils with $I_c < 2.6$ are inferred to be liquefiable (Robertson and Wride 1998). However, because I_c boundaries between soil types are approximate, regional refinement may be needed for optimal efficiency (e.g., Pease 2010; Li et al. 2007). Accordingly, analysts of the data compiled herein may be interested in the susceptibility correlations of Maurer et al. (2019), developed specifically for soils in Christchurch. Using these correlations, the probability that a soil is "susceptible" to liquefaction is:

$$P_{susceptible}(I_c) = 1 - \Phi\left[\frac{\ln(I_c/x_m)}{\beta}\right] \quad (8)$$

where Φ is the Gaussian cumulative distribution function; x_m is the median value of the distribution (the value of I_c corresponding to 50% probability); and β is the logarithmic standard deviation. Using this form, Maurer et al. (2019) correlated I_c to four criteria based on Atterberg limits, the coefficients for which are provided in Table 4, and an example of which is shown in Figure 8. Here, "susceptible" generically refers to the varying definitions adopted by the respective works. For example, the Boulanger and Idriss (2006) criterion was explicitly developed to determine the most appropriate analysis procedure for predicting cyclic response, based on whether the soil's expected behavior is "sand-like" or "clay-like." For deterministic analyses in which a single I_c threshold is desired, the median value of the probability distribution (x_m) is recommended, such that soils with I_c exceeding x_m are not susceptible per the underlying criterion.

Table 4. Model coefficients for I_c-susceptibility relationship (Equation 8) (Maurer et al. 2019).

Susceptibility Criterion	β	x_m
Boulanger and Idriss (2006)	0.0851	2.5031
Polito (2001)	0.0988	2.5474
Seed et al. (2003)	0.1348	2.6214
Bray and Sancio (2006)	0.1275	2.7315

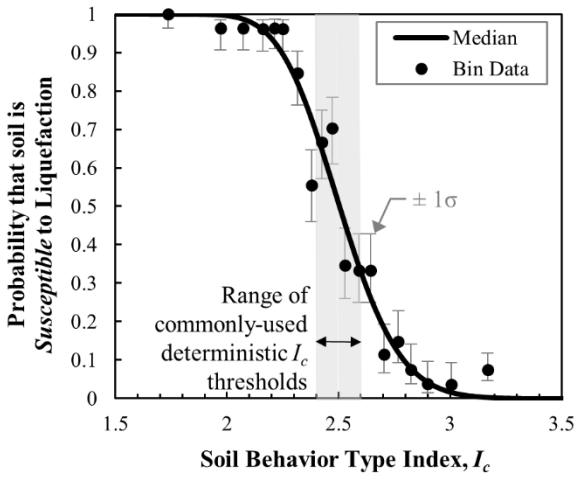


Figure 8. The probability of liquefaction susceptibility per the Boulanger and Idriss (2006) criterion as a function of measured I_c. The range of deterministic I_c thresholds commonly used in practice is also highlighted (Maurer et al., 2019).

Fines Content. Some liquefaction models use fines content (*FC*) as a predictive variable. As with liquefaction susceptibility, *FC* is best measured directly, but continuous sampling and measurement is not feasible for a large CPT campaign. Accordingly, CPT correlations developed from global data are commonly used to estimate *FC* but can often be improved via regional calibration. Analysts testing or training response models may thus be interested in the regional I_c

– *FC* correlations of Lees et al. (2015a) and Maurer et al. (2019). Lees et al. (2015a) calibrated the general I_c – *FC* correlation of Boulanger and Idriss (2014), wherein *FC* (%) is estimated as:

$$FC = 80\ (I_c + C_{FC}) - 137 \tag{9}$$

where C_{FC} is a calibration parameter that may adjust the general correlation (i.e., $C_{FC} = 0$) to region-specific conditions. Analyzing 2,600 *FC* measurements from Canterbury, Lees et al. (2015a) proposed that $C_{FC} = 0.2$ was optimal. Using a similar amount of data but different processing and regression methods, Maurer et al. (2019) proposed that *FC* be estimated as:

$$\mu_{FC} = 80.645\ I_c - 128.5967 \tag{10}$$

where μ_{FC} is the mean estimate of *FC* (%), limited to $0\% \leq FC\ (\%) \leq 100\%$. Guidance on using this correlation probabilistically is given in Maurer et al. (2019). A comparison of the Maurer et al. (2019) correlation and others, along with data from Christchurch, is shown in Figure 9.

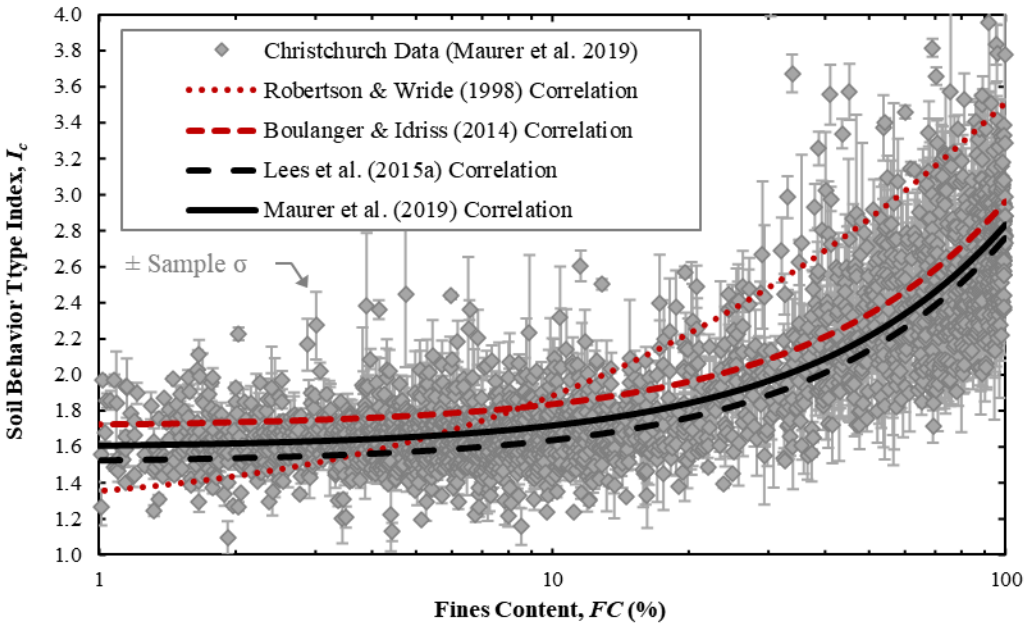


Figure 9. Canterbury I_c – *FC* data and correlations (Lees et al. 2015a; Maurer et al. 2019), and comparison with the generic Robertson and Wride (1998) and Boulanger and Idriss (2014) correlations.

Soil Density. Total and effective vertical stresses are integral to CPT data processing and a common input to liquefaction models. While the authors are unaware of calibrated, Canterbury-specific correlations for estimating soil unit weights, several global correlations are available, including Mayne et al. (2010) and Robertson and Cabal (2010), with the latter being used in previous Canterbury earthquake research by the authors (e.g., Green et al. 2018; Geyin and Maurer 2019). While liquefaction models may be relatively insensitive to the adopted correlation, users should consider constraints, as needed and reasonable, to limit physically indefensible values.

GWT at Time of Testing. As discussed, the database contains event specific *GWT* depths, which are an estimate of conditions immediately prior to each earthquake. These may be used to infer the depth of saturation for assessing liquefaction susceptibility (to be further discussed) but may differ from the *GWT* depths at the time of testing, which are needed for CPT stress-normalization as part of routine data processing. While regional hydrologic models are unavailable for CPT test dates, which are shown in Figure 10a, most CPTs (~90%) were performed between 22 Feb 2011 and 30 Sept 2013. In addition, it can be seen in Figures 10b and 10c that *GWT* depths in Feb 2011 were typically ~1% shallower relative to Sept 2010 and ~14% deeper relative to Feb 2016. In the absence of more rigorous modeling, adopting a *GWT* depth either interpolated from the three estimates available, or simply averaged from the Feb 2011 and Feb 2016 values (since most CPTs were performed during this time), may provide a reasonable estimate for CPT stress normalization. Estimates could also be obtained from analyses of the CPT u2 data, although the reliability of this data due to issues with porous stone saturation, etc. is unknown for the compiled database.

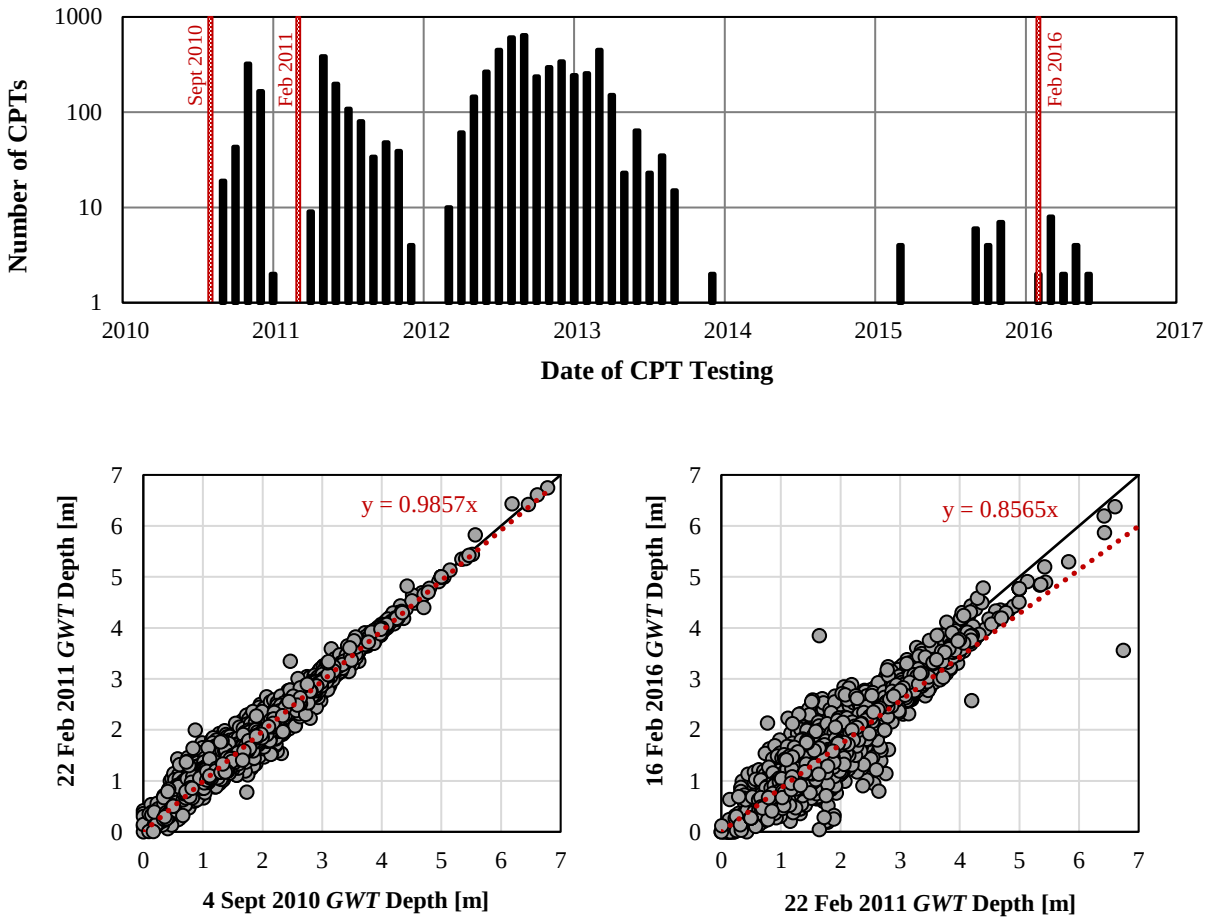


Figure 10. Case-history database statistics: **(a)** Monthly histogram of CPT test dates; **(b)** *GWT* depths from Sept 2010 versus Feb 2011; **(c)** *GWT* depths from Feb 2011 versus Feb 2016.

Model Training, Testing, and Bias. Historically, publications proposing, calibrating, or evaluating liquefaction models often lack standard test metrics (e.g., Maurer et al. 2015d), hindering quantifications and comparisons of model performance. Receiver-operating-characteristic (ROC) analyses (e.g., Fawcett 2006) are ubiquitous in medical diagnostics and data science (Zou 2007), and increasingly, are being adopted in geotechnical modelling (e.g., Oommen et al. 2010; Zhu et al. 2017; Green et al. 2017; Upadhyaya et al. 2020). ROC analyses provide a standard and objective assessment of prediction efficiency via the area-under-the-ROC-curve (AUC) and are relatively

insensitive to sampling imbalance (i.e., unequal positives and negatives). Analysts using the curated data to test or train liquefaction models should similarly adopt standard, objective, and repeatable measures of performance, be it ROC analyses or some other. In addition, while the Canterbury earthquakes resulted in a wealth of data, this data nonetheless samples the geologic and seismologic setting of one region, the findings from which may or may not translate elsewhere. Analysts should carefully consider sampling bias and weigh results from Canterbury with those from global case-histories (e.g., Brandenberg et al. 2020). Given that the compiled Canterbury database is much larger than that resulting from all other earthquakes combined, the finite-sample uncertainty of model performance should be computed for each respective database (e.g., via bootstrap sampling) and used to test for statistical significance. That is, to illustrate the sensitivity of performance to the data available for analysis, and to assess whether differences could arise from chance (i.e., due to finite sampling) and not because one model is better than another. As an example, p-values specific to ROC analyses may be computed per DeLong et al. (1988), an application of which is presented in Geyin et al. (2020a) for liquefaction case-history data.

Lingering Uncertainties

GWT Depth and Saturation. It is established that liquefaction resistance and degree of saturation are inversely related, all else being equal (e.g., Ishihara and Tsukamoto 2004; Hossain et al. 2013), and that soil beneath the apparent *GWT* can conceivably be less than 100% saturated (e.g., due to seasonal or tidal fluctuations, or to biologic activity). This phenomenon has been inferred from crosshole p-wave velocities (Cox et al. 2018) at select locations in Christchurch and investigated as a possible cause of observed mispredictions of liquefaction by popular models (McLaughlin 2017; Boulanger et al. 2018; Yost et al. 2019; Ntritsos and Cubrinovski, 2020). One detailed study of this issue is that of McLaughlin (2017), who analyzed 31 cases in Christchurch and computed

liquefaction potential index (*LPI*) values with and without various corrections. These included corrections for partial saturation, site-specific *FC*, and inverse filtering and interface correction. While evidence of partial saturation was found at some locations, the corrections to *LPI* were typically minor compared to those made for *FC* and inverse filtering. The results of McLaughlin (2017) indicate that partial saturation beneath the *GWT* could potentially be important at some sites, but in general, does not sufficiently or consistently explain mispredictions of liquefaction. Nonetheless, uncertainties pertaining to partial saturation persist, but could only be adequately addressed via extensive additional in-situ testing and/or regional hydrologic modeling. Owing to the rarity of p-wave measurements in parallel with CPTs, it is unknown whether partial saturation beneath the *GWT* is present in other case-histories, previously collected elsewhere globally.

CPT Spacing in Time. As discussed, multiple case-histories were often developed from a single site (i.e., CPT) affected by multiple earthquakes, wherein the event-specific *GWT*, *PGA*, and response were known. This raises the question of whether a CPT performed at one moment in time (predominantly between Feb 2011 and Sept 2013; see Figure 10a) is representative of a soil profile at multiple other times when earthquakes occurred? This will be addressed in three parts.

First, does the approach taken break from precedent? When considering all liquefaction case histories published to-date (e.g., Boulanger and Idriss 2014), in-situ testing has been performed: (i) well in advance of an earthquake; (ii) months-to-years after an earthquake; (iii) decades after an earthquake; and (iv) all scenarios in between. Additionally, between the time of in-situ testing and the occurrence of an earthquake, or between the time of the occurrence of an earthquake and in-situ testing, it is often the case that multiple other earthquakes of varying intensity have affected a site. In the authors' opinion, there has historically been no standard, or best practice, for the relative timing of in-situ testing when publishing liquefaction case-histories. The Boulanger and

Idriss (2014) global database contains cases representing each of the four scenarios above, with multiple case histories based on the same CPT. Of the 255 case histories compiled therein, 25% are cases in which one CPT was used to develop multiple case histories. As an example, four case histories were developed from one CPT effected by earthquakes occurring over a 10-year span.

Second, does CPT data change over time once deposited or disturbed? Increases in the strength and stiffness of sands over time, or "aging effects," have been widely investigated. Temporal gains have been discerned both from penetration resistance, with reported gains of 3-7% per log-cycle in years (Mesri et al. 1990; Kulhawy and Mayne 1990), and from liquefaction resistance (i.e., CRR), with reported gains of 9-17% per log-cycle in years (Arango et al. 2000; Hayati and Andrus 2009; Saftner et al. 2015). It has thus been proposed that aging effects may be resolved into gains measurable by intermediate-to-large-strain penetration data and gains in liquefaction resistance, where the latter is influenced by small-strain fabric phenomena difficult to detect at large strain (Leon et al. 2006). Of relevance to the compilation of case histories, CPT measurements could conceivably vary with time, particularly over short time scales following liquefaction. For example, assuming the rates above, and that a soil is "reset" following liquefaction, CPT resistance measured 1 month after an earthquake could be 3-7% less than if measured 1 year later. While such changes are plausible, they would be difficult to distinguish from site variability and measurement uncertainty, and to-date, have not been considered in case-history publications.

Third, does CPT data change due to repeated disturbance from shaking, and if so, does the magnitude and direction of change (e.g., an increase or decrease in penetration resistance) depend on whether liquefaction did or did not occur? A closely related, but different question, is whether liquefaction resistance changes due to prior shaking/liquefaction, even if CPT data does not

change? Researchers have sought answers to these questions using a variety of approaches: (i) CPT testing in the field before and after shaking/liquefaction (e.g., Lees et al. 2015b; Finno et al. 2016); (ii) CPT testing in centrifuge and shaking-table models, before and after shaking/liquefaction (e.g. Darby et al. 2016; Dobry et al. 2019); and (iii) cyclic triaxial and cyclic simple-shear tests wherein samples were subjected to multiple shaking/liquefaction sequences (e.g., Ha et al. 2011; Wang et al. 2013). Various conclusions were collectively drawn from these experiments, including: (i) penetration resistance increases; (ii) penetration resistance decreases; (iii) penetration resistance does not change, even after severe liquefaction; (iv) penetration resistance changes in some parts of the profile but not others; (v) the magnitude and direction of the change in penetration resistance depends on the number of previous shaking cycles, and on the pore pressure generated by those cycles; and (vi) liquefaction resistance may change, independent of whether this change is detected via CPT data. Most relevant to the current effort, perhaps, is the work of Lees et al. (2015b), who studied pairs of CPTs performed at 30 locations before and after the 22 Feb 2011 Christchurch earthquake, concluding that CPT measurements did not change in a statistically significant manner.

In summary, the approach taken by this study is consistent with past precedent. Additionally, questions pertaining to CPT data and soil response during earthquake sequences – which are very worthy of investigation – have not been adequately resolved to suggest when CPTs should or should not be used to compile liquefaction case histories. However, the compiled dataset could potentially be analyzed to further study these issues in the field, making use of the provided CPT coordinates and test dates.

Conclusions

Earthquakes occurring over the last decade in Canterbury, New Zealand, resulted in liquefaction case-history data of unprecedented quantity. Accordingly, this paper presented a curated dataset containing ~15,000 CPT-based liquefaction case-histories compiled from three earthquakes in this sequence. The compiled, post-processed data was provided in a dense array structure, allowing researchers to easily access and analyze information pertinent to CPT-based site characterization and free-field liquefaction response. Research opportunities using this data include, but are not limited to, the training or testing of new and existing liquefaction-prediction models. The many methods used to obtain and process the case-history data were detailed herein, as was the structure of the compiled file. Numerous recommendations for analyzing the data were also outlined, including nuances and limitations that users should carefully consider prior to analysis.

Acknowledgements

The authors gratefully acknowledge Jason Motha of the University of Canterbury, New Zealand, for converting the case-history dataset into a python-compatible data frame.

Data Availability

The Canterbury dataset is available in digital format through the NEHRI DesignSafe Data Depot at https://doi.org/10.17603/ds2-tygh-ht91. Additionally, Appendices A and B are provided as an electronic supplement to this paper.

Funding

The presented study is based on work supported by the National Science Foundation (NSF), US Geological Survey (USGS), and Pacific Earthquake Engineering Research Center (PEER) under Grants CMMI-1751216, CMMI-1825189, CMMI-1937984, G18AP-00006, and 1132-NCTRBM, respectively. This work was also partially supported by QuakeCoRE, the New Zealand Centre for Earthquake Resilience, and is QuakeCoRE publication 0616. The authors also acknowledge the countless people who contributed to the liquefaction case-history data studied herein, which was collected under the auspices the New Zealand Earthquake Commission (EQC). 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, USGS, PEER, QuakeCoRE, or EQC.

References

Abrahamson NA and Youngs, RR (1992) A stable algorithm for regression analyses using the random effects model. *Bulletin of the Seismological Society of America* 82(1): 505–510.

Ahmadi MM and Robertson PK (2005) Thin-layer effects on the CPT q_c measurement. *Canadian Geotechnical Journal* 42(5): 1302-1317.

Anselin L (1995) Local indicators of spatial association—LISA. *Geographical analysis* 27(2): 93-115.

Arango I, Lewis MR and Kramer C. (2000) Updated liquefaction potential analysis eliminates foundation retrofitting of two critical structures. *Soil Dynamics and Earthquake Engineering* 20: 17-25.

Begg J and Jones K (2012) 20,000 years of time in Christchurch. *Presentation to New Zealand Geotechnical Society*, Canterbury Branch, June 28, 2012, New Zealand Geotechnical Society, Wellington, New Zealand.

Boulanger RW and DeJong JT (2018) Inverse filtering procedure to correct cone penetration data for thin-layer and transition effects. *Cone Penetration Testing 2018*, Hicks, Pisano, and Peuchen, eds., Delft University of Technology, The Netherlands: 25-44.

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 (2007) Evaluation of cyclic softening in silts and clays. *Journal of Geotechnical and Geoenvironmental Engineering* 133(6): 641–652.

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, Khosravi M, Cox BR and DeJong JT (2019) Liquefaction Evaluation for an Interbedded Soil Deposit: St. Teresa's School, Christchurch, New Zealand. *IACGE 2018: Geotechnical and Seismic Research and Practices for Sustainability*: 686-704.

Bradley BA (2012) Strong ground motion characteristics observed in the 4 September 2010 Darfield, New Zealand earthquake. *Soil Dynamics and Earthquake Engineering* 42: 32–46.

Bradley BA (2013) A New Zealand-specific pseudo spectral acceleration ground-motion prediction equation for active shallow crustal earthquakes based on foreign models. *Bulletin of the Seismological Society of America* 103(3):1801–1822.

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.

Bradley BA (2016) Strong ground motion characteristics observed in the 13 June 2011 Mw6.0 Christchurch, New Zealand earthquake. *Soil Dynamics and Earthquake Engineering* 91: 23-38.

Bradley BA and Cubrinovski M (2011) Near-source strong ground motions observed in the 22 February 2011 Christchurch earthquake. *Seismological Research Letters* 82(6): 853–865.

Bradley BA, and Hughes, M (2012) Conditional peak ground accelerations in the Canterbury earthquakes for conventional liquefaction assessment. *University of Canterbury, Christchurch, New Zealand*.

Bradley BA, Savarimuthu S, Lagrava D, Huang J, Motha J, Polak V and Bae S (2017) SeisFinder: A web application for extraction of data from computationally-intensive earthquake resilience calculations. Poster Presentation at *2017 SCEC Annual Meeting*. Palm Springs, CA.

Brandenberg SJ, Zimmaro P, Stewart JP, Kwak DY, Franke KW, Moss RE, … Kramer SL. (2020) Next-generation liquefaction database. *Earthquake Spectra*, Doi: 10.1177/8755293020902477.

Bray JD and Sancio RB (2006) Assessment of the liquefaction susceptibility of fine-grained soils. *Journal of Geotechnical and Geoenvironmental Engineering* 132(9): 1165-1177.

Bray J, Cubrinovski M, Zupan J, and Taylor M (2014) Liquefaction effects on buildings in the central business district of Christchurch. *Earthquake Spectra* 30(1): 85–109.

Brown LJ (1975) Water well data. Sheet 576/7-8, Belfast-Styx. New Zealand Geological Survey, Report 72.

Brown LJ and Weeber JH (1992) Geology of the Christchurch urban area: Institute of Geological and Nuclear Sciences Geological Map 1. *Institute of Geological and Nuclear Sciences Limited, Lower Hutt, New Zealand, scale*, *1*(25,000), 1.

Brown LJ, Beetham RD, Paterson BR and Weeber JH (1995) Geology of Christchurch, New Zealand. *Environmental Engineering and Geoscience* 1(4): 427–488.

Buck JR, Daniel MM and Singer AC (2002) *Computer Explorations in Signals and Systems Using MATLAB®*, 2nd Edition. Upper Saddle River, NJ: Prentice Hall.

CERA (2013) Purpose and scope of the Canterbury geotechnical database. *Canterbury Earthquake Recovery Authority*, Information Note, Revision D. Christchurch, New Zealand.

Chalmers G, McLennan N and Barsanti L (2013) Lyttelton port of Christchurch seismic resilience from an owners' perspective. *Proceedings of Ports 2013: Success Through Diversification*: 1405-1414.

Cox BR, Stolte AC, Stokoe KH and Wotherspoon LM. (2018) A Direct-Push Crosshole (DPCH) Test Method for the In Situ Evaluation of High-Resolution P-and S-Wave Velocities. *Geotechnical Testing Journal* 42(5): 1101-1132.

Cubrinovski M, Bray JD, Taylor M, Giorgini S, Bradley BA, Wotherspoon L and Zupan J (2011) Soil liquefaction effects in the central business district during the February 2011 Christchurch earthquake. *Seismological Research Letters* 82: 893–904.

Cubrinovski M, Winkley A, Haskell J, Palermo A, Wotherspoon L, Robinson K, Bradley B, Brabhaharan P and Hughes, M (2014) Spreading-induced damage to short-span bridges in Christchurch, New Zealand. *Earthquake Spectra* 30(1): 57–83.

Cubrinovski, M, and Robinson, K, (2016) Lateral spreading: Evidence and interpretation from the 2010–2011 Christchurch earthquakes. *Soil Dynamics and Earthquake Engineering* 91: 187-201.

Darby KM, Bronner JD, Parra Bastidas AM, Boulanger RW and DeJong JT (2016) Effect of shaking history on the cone penetration resistance and cyclic strength of saturated sand. In *Geotechnical and Structural Engineering Congress 2016*: 1460-1471.

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.

Dobry R, Ladd RS, Yokel FY, Chung RM and Powell D (1982) *Prediction of pore water pressure buildup and liquefaction of sands during earthquakes by the cyclic strain method* (Vol. 138). Gaithersburg, MD: National Bureau of Standards.

Dobry R, Thevanayagam S, El-Sekelly W, Abdoun T and Huang Q (2019) Large-Scale Modeling of Preshaking Effect on Liquefaction Resistance, Shear Wave Velocity, and CPT Tip Resistance of Clean Sand. *Journal of Geotechnical and Geoenvironmental Engineering* 145(10): 04019065.

Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874.

Finno RJ, Gallant AP and Sabatini PJ (2016) Evaluating ground improvement after blast densification: performance at the Oakridge landfill. *Journal of Geotechnical and Geoenvironmental Engineering* 142(1): 04015054.

Geyin M and Maurer BW (2019) An analysis of liquefaction-induced free-field ground settlement using 1,000+ case-histories: observations vs. state-of-practice predictions. *Geocongress 2019: Earthquake Engineering and Soil Dynamics (C.L. Meehan, S. Kumar, M.A. Pando, and J.T. Coe, eds.),* Geotechnical Special Publication 308: 489-498.

Geyin M, Baird AJ and Maurer BW (2020a) 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., Bradley, B.A., Green, R.A., and van Ballegooy, S. (2020b). "CPT-Based Liquefaction Case Histories Resulting from the 2010-2016 Canterbury, New Zealand, Earthquakes: A Curated Digital Dataset (Version 2)." DesignSafe-CI. https://doi.org/10.17603/ds2-tygh-ht91.

Geyin M and Maurer BW (2020a) Horizon: CPT-based liquefaction risk assessment and decision software. DesignSafe-CI, doi: 10.17603/ds2-2fky-tm46.

Geyin M and Maurer BW (2020b) Fragility functions for liquefaction induced ground failure. *Journal of Geotechnical and Geoenvironmental Engineering,* doi: 10.1061/(ASCE)GT.1943-5606.0002416.

Goda K and Hong HP (2008) Spatial correlation of peak ground motions and response spectra. *Bulletin of the Seismological Society of America* 98(1): 354–365.

Green RA, Allen A, Wotherspoon L, Cubrinovski M, Bradley B, Bradshaw A, Cox B and Algie T (2011) Performance of levees (stopbanks) during the 4 September M_w7.1 Darfield and 22 February 2011 M_w6.2 Christchurch, New Zealand, earthquakes. *Seismological Research Letters* 82(6): 939–949.

Green RA, Cubrinovski M, Cox B, Wood C, Wotherspoon L, Bradley B and Maurer B (2014)

Select Liquefaction Case Histories from the 2010-2011 Canterbury Earthquake Sequence. *Earthquake Spectra* 30(1): 131-153.

Green RA and Ziotopoulou K (2015). Overview of screening criteria for liquefaction triggering susceptibility. *Proceedings of the 10th Pacific Conference on Earthquake Engineering*, Nov 6-8, Sydney, Australia. Australian Earthquake Engineering Society; Paper No. 35. 2015.

Green RA, Upadhyaya S, Wood C, Maurer BW, Cox BR, Wotherspoon L, Bradley BA and Cubrinovski M (2017) Relative efficacy of CPT- versus Vs – based simplified liquefaction evaluation procedures. *19th International Conference on Soil Mechanics and Geotechnical Engineering*, 1521-1524.

Green RA, Maurer BW and van Ballegooy S (2018) The influence of the non-liquefied crust on the severity of surficial liquefaction manifestations: case history from the 2016 Valentine's Day earthquake in New Zealand. *Geotechnical Earthquake Engineering and Soil Dynamics V: Liquefaction Triggering, Consequences, and Mitigation (S.J. Brandenberg and M.T. Manzari, eds.)*, Geotechnical Special Publication 290: 21-32.

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.

Ha IS, Olson SM, Seo MW and Kim MM (2011) Evaluation of reliquefaction resistance using shaking table tests. *Soil dynamics and earthquake engineering* 31(4): 682-691.

Hayati H and Andrus RD (2009) Updated liquefaction resistance correction factors for aged sands. *Journal of geotechnical and geoenvironmental engineering* 135(11): 1683-1692.

Hossain AM, Andrus RD and Camp III WM (2013) Correcting liquefaction resistance of

unsaturated soil using wave velocity. *Journal of Geotechnical and Geoenvironmental Engineering* 139(2): 277-287.

Ishihara K and Tsukamoto Y (2004) Cyclic strength of imperfectly saturated sands and analysis of liquefaction. *Proceedings of the Japan Academy, Series B*, 80(8): 372-391.

Jeffries MG and Davies MP (1993) Use of CPTu to estimate equivalent SPTN60. *Geotechnical Testing Journal* 16(4): 458–468.

Johnson RA and Wichern DW (2007) Applied multivariate statistical analysis, Pearson Prentice Hall, Upper Saddle River, NJ.

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.

Kulhawy FH and Mayne PW (1990) Manual on estimating soil properties for foundation design. Final Rep. 1493-6, EL-6800, Electric Power Research Institute, Palo Alto, Calif

Kwasinski A, Eidinger J, Tang A, and Tudo-Bornarel C (2014) Performance of electric power systems in the 2010–2011 Christchurch, New Zealand, earthquake sequence. *Earthquake Spectra* 30(1): 205–230.

Lees J, van Ballegooy S and Wentz FJ (2015a) Liquefaction susceptibility and fines content correlations of the Christchurch soils. *Proceedings of the 6th international conference on earthquake geotechnical engineering*. Nov 2-4; Christchurch, New Zealand; Paper No. 491.

Lees J, Ballagh RH, Orense RP and Van Ballegooy S (2015b) CPT-based analysis of liquefaction and re-liquefaction following the Canterbury earthquake sequence. *Soil Dynamics and Earthquake Engineering* 79: 304-314.

Leon E, Gassman SL and Talwani P (2006) Accounting for soil aging when assessing liquefaction potential. *Journal of Geotechnical and Geoenvironmental Engineering* 132(3): 363-377.

Li DK, Juang CH, Andrus, RD and Camp WM (2007) Index properties-based criteria for liquefaction susceptibility of clayey soils: a critical assessment. *Journal of Geotechnical and Geoenvironmental Engineering* 133(1): 110-115.

Lunne T, Robertson PK and Powell JM (1997) *Cone Penetration Testing in Geotechnical Practice*. Blackie Academic & Professional, London, U.K.

Mayne PW, Peuchen J and Bouwmeester D (2010) Soil unit weight estimation from CPTs. *2nd International symposium on cone penetration testing*, 8p

McLaughlin K (2017) Investigation of false-positive liquefaction case history sites in Christchurch, New Zealand. M.S. Thesis. The University of Texas at Austin, Austin, TX.

Maurer BW, Green RA, Cubrinovski M and Bradley BA (2014) 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 (2015a) 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 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 and Bradley B (2015c) Assessment of CPT-based methods for liquefaction evaluation in a liquefaction potential index framework. *Géotechnique* 65(5): 328-336.

Maurer BW, Green RA, Cubrinovski M and Bradley BA (2015d) 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, 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.

Mesri G, Feng T and Benak JM (1990) Postdensification penetration resistance of clean sands. *Journal of Geotechnical Engineering* 116(7): 1095–1115.

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 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.

Ntritsos N and Cubrinovski M (2020) A CPT based effective stress analysis procedure for liquefaction assessment. *Soil Dynamics and Earthquake Engineering* 131: 106063.

NZGD (2020) New Zealand Geotechnical Database. *New Zealand Earthquake Commission.* <https://www.nzgd.org.nz/>, (Accessed Jan 01, 2020).

O'Rourke T, Jeon SS, Toprak S, Cubrinovski M, Hughes M, van Ballegooy S and Bouziou D (2014) Earthquake response of underground pipeline networks in Christchurch, NZ. *Earthquake Spectra* 30(1): 183–204.

Pease JW (2010) Misclassification in CPT liquefaction evaluation. *2nd International Symposium on Cone Penetration Testing*. Huntington Beach, CA, Paper # 3-23. 2010.

Pinheiro J, Bates DM, DebRoy S, Sarkar D and the R Core Team. (2008). NLME: Linear and nonlinear mixed effects models, R Corepackage, Ver. 3.1, http://www.r-project.org (Mar. 26, 2014).

Polito C (2001) Plasticity Based Liquefaction Criteria. *Proceedings of the Fourth International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*, San Diego, California, March 26-31. Paper 25.

Quigley MC, Bastin S and 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.

Rashidian V and Gillins DT (2018) Modification of the liquefaction potential index to consider the topography in Christchurch, New Zealand. *Engineering Geology* 232: 68-81.

Robertson PK (2011) Automated detection of CPT transition zones. *Geotechnical News*, June: 35-38.

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

Robertson PK and Cabal KL (2010) Estimating soil unit weight from CPT. *2nd Int. Symp. on Cone Penetration Testing*, 8p.

Saftner DA, Green RA and Hryciw RD (2015) Use of Explosives to Investigate Liquefaction Resistance of Aged Sand Deposits. *Engineering Geology* 199: 140-147.

Seed RB, Cetin KO, Moss RES, Kammerer AM, Wu J, Pestana JM, Riemer MF, Sancio RB, Bray JD, Kayen RE and Faris A (2003) Recent advances in soil liquefaction engineering: a unified and consistent framework. *26th Annual ASCE Los Angeles Geotechnical Spring Seminar*. 30 April, Long Beach, CA.

Seisfinder (2020). *A web application to enable the extraction of high-fidelity outputs from computationally intensive earthquake resilience calculations*. Accessed 1 August 2020 <https://quakecoresoft. canterbury.ac.nz/seisfinder/>

Sideras S (2019) Evolutionary Intensity Measures for More Accurate and Informative Evaluation of Liquefaction Triggering. PhD Dissertation, University of Washington, Seattle, WA.

Tang A, Kwasinski A, Eidinger J, Foster C, and Anderson P (2014) Telecommunication systems' performance: Christchurch earthquakes. *Earthquake Spectra* 30(1): 231–252.

Townsend D, Lee JM, Strong DT, Jongens R, Smith Lyttle B, Ashraf S, ... and Taylor ML (2016) Mapping surface liquefaction caused by the September 2010 and February 2011 Canterbury earthquakes: A digital dataset. *New Zealand Journal of Geology and Geophysics* 59(4): 496-513.

Treadwell DD (1976) The influence of gravity, prestress, compressibility, and layering on soil resistance to static penetration. PhD Dissertation, Univ. of California, Berkeley, CA.

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)*.

Upadhyaya S, Maurer BW, Green RA and Rodriguez-Marek A (2020) Selecting the Optimal Factor of Safety or Probability of Liquefaction Triggering for Engineering Projects Based on Misprediction Costs. *Journal of Geotechnical and Geoenvironmental Engineering*, (*in review*).

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.

van Ballegooy S, Cox SC, Thurlow C, Rutter HK, Reynolds T, Harrington G, Fraser J and Smith T (2014b) Median water table elevation in Christchurch and surrounding area after the 4 September 2010 Darfield earthquake: Version 2. *GNS Science Report 2014/18*, 2014b.

van Ballegooy S, Green RA, Lees J, Wentz F and Maurer BW (2015) Assessment of various CPT based liquefaction severity index frameworks relative to the Ishihara (1985) H_1-H_2 boundary curves. *Soil Dynamics and Earthquake Engineering* 79: 347-364.

van der Linden TI, De Lange DA and Korff M (2018) Cone penetration testing in thinly inter-layered soils. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering* 171(3): 215-231.

Wang S, Yang J and Onyejekwe S (2013) Effect of previous cyclic shearing on liquefaction resistance of Mississippi River Valley silt. *Journal of materials in civil engineering* 25(10): 1415-1423.

Wotherspoon L, Bradshaw A, Green RA, Wood C, Palermo A, Cubrinovski M and Bradley BA (2011) Performance of bridges during the 2010 Darfield and 2011 Christchurch earthquakes. *Seismological Research Letters* 82(6): 950–964.

Wotherspoon LM, Orense RP, Green RA, Bradley BA, Cox BR and Wood CM (2015) Assessment of liquefaction evaluation procedures and severity index frameworks at Christchurch strong motion stations. *Soil Dynamics and Earthquake Engineering* 79: 335-346.

Yost KM, Cox BR, Wotherspoon L, Boulanger RW, van Ballegooy S and Cubrinovski M (2019) In Situ Investigation of False-Positive Liquefaction Sites in Christchurch, New Zealand: Palinurus Road Case History. *Geo-Congress 2019: Earthquake Engineering and Soil Dynamics* 436-451.

Yost KM, Green RA, Upadhyaya S, Maurer BW, Yerro-Colom A and Martin E (2020) Assessment of the Efficacies of Correction Procedures for Multiple Thin Layer Effects on Cone Penetration Tests. *Soil Dynamics and Earthquake Engineering*, (*in preparation*).

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.

Zou KH (2007) Receiver operating characteristic (ROC) literature research. *On-line bibliography*, http://www.spl.harvard.edu/archive/spl-pre2007/pages/ppl/zou/roc.html

APPENDIX B

Publication 2 (Dataset 1), highlighted below, is provided on the following pages in full. Specifically, the "ReadMe" file is provided (not the database file).

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology*, In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

CPT-BASED LIQUEFACTION CASE HISTORIES FROM GLOBAL EARTHQUAKES: A DIGITAL DATASET (VERSION 1).

Mertcan Geyin,[a)] M.EERI, Brett W. Maurer,[a)]

This dataset contains 275 CPT-based liquefaction case histories from 21 global earthquakes, as summarized in Table 1. These cases are sourced from the literature, to include observations of manifestation severity, CPT soundings, and estimation of GWT depth and ground-motion IMs, as generally reported by original investigators. When available, recent refinements are adopted from the literature. Of the 275 global cases, ~58% are "Manifestation" and ~42% are "No Manifestation." To properly recognize all sources of data used to compile this dataset, a reference list appears near the end of this paper, parsed by earthquake; a table of data for each case-history is also provided. In this regard, the case history assemblages of Moss et al. (2003) and Boulanger and Idriss (2014) are acknowledged for greatly assisting the present compilation. These data, or subsets of these data, have been analyzed in the literature (e.g., Geyin and Maurer, 2020a) and are provided here to replicate published results.

a) University of Washington, Seattle, Washington, United States

Table 1. Summary of Liquefaction Case Histories

Date	Earthquake	Country	Magnitude (M_w)	Number of Case Histories
16/6/1964	Niigata	Japan	7.60	3
9/2/1971	San Fernando	USA	6.60	2
4/2/1975	Haicheng	China	7.00	2
27/7/1976	Tangshan	China	7.60	10
15/10/1979	Imperial Valley	USA	6.53	7
9/6/1980	Victora (Mexicali)	Mexico	6.33	5
26/4/1981	Westmoreland	USA	5.90	9
26/5/1983	Nihonkai-Chubu	Japan	7.70	2
28/10/1983	Borah Peak	USA	6.88	3
2/3/1987	Edgecumbe	New Zealand	6.60	23
24/11/1987	Elmore Ranch	USA	6.22	2
24/11/1987	Superstition Hills	USA	6.54	8
18/10/1989	Loma Prieta	USA	6.93	67
17/1/1994	Northridge	USA	6.69	3
16/1/1995	Hyogoken-Nambu	Japan	6.90	21
17/8/1999	Kocaeli	Turkey	7.51	16
20/9/1999	Chi-Chi	Taiwan	7.62	34
8/6/2008	Achaia-Ilia	Greece	6.40	2
4/4/2010	Baja	Mexico	7.20	3
11/3/2011	Tohoku	Japan	9.00	7
20/5/2012	Emilia	Italy	6.10	46
			Total	**275**

CPT Processing Methodology

While the CPT offers advantages among in-situ tests used to predict liquefaction, it is still limited by the volume of soil mobilized around the cone. As an intermediate-to-large-strain penetration test, this mobilized zone acts as a physical “low-pass filter” on the true soil stratigraphy, removing information from the low spatial wavelengths, such as the data defining a thin soil stratum or the interface between two disparate soils. These spatial smoothing effects, which are commonly referred to as “thin layer” and “transition” effects, have long been recognized

and studied (e.g., Treadwell 1976; Lunne et al. 1997; Ahmadi and Robertson 2005; Robertson 2011; van der Linden 2018). While chart-based methods exist for manually correcting these effects on CPT data, Boulanger and DeJong (2018) proposed the first programmable procedure. This methodology, referred to as an "inverse filtering and interface detection" procedure, predicts the "true" CPT profile from measured CPT values. As a demonstration of the methodology, CPT data is shown in Fig. 1, both with and without correction. Since measured values reflect a filtered view of reality, their correction would improve subsurface characterization. Accordingly, both measured and "true" CPT data are provided in the database.

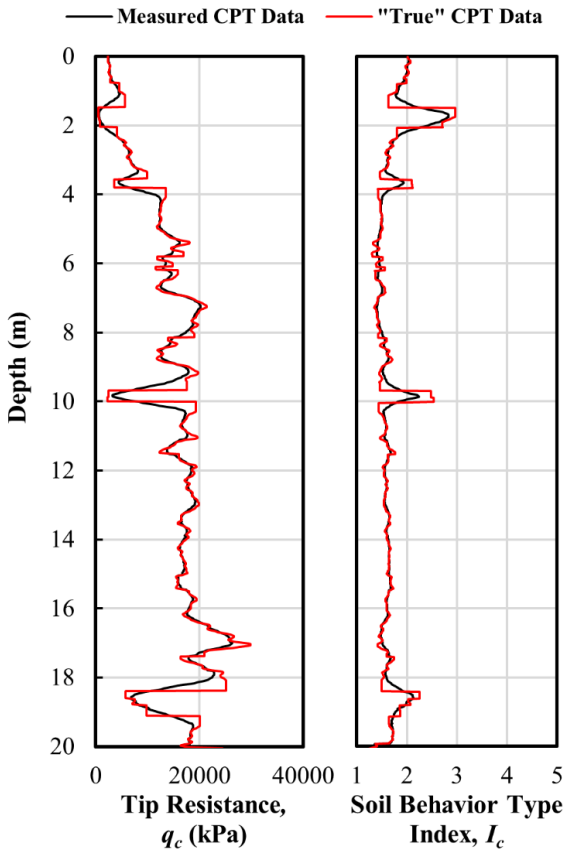


Fig. 1. Example CPT data with (i.e. "true") and without (i.e. "measured") inverse-filtering and interface correction via the Boulanger and DeJong (2018) procedure, as implemented in the software *Horizon* (Geyin and Maurer, 2020b).

While the reader is referred to Boulanger and DeJong (2018) for complete details, the procedure is summarized as follows. The correction of CPT data can be idealized as the conversion of measured q_c and f_s values, or q^m and f^m, to the "true" values q^t and f^t, which would be obtained if measurements reflected conditions at discrete points. This can be viewed as an inverse problem, with the goal of determining the "true" values via inversion of what was measured. As illustrated in Fig. 2 (in this case for tip resistance), q^m may deviate from q^t near layer interfaces because the measurements are influenced by materials both above and below the interface, even though the measurement is recorded at a discrete depth. In the case shown in Fig. 2, q^m is artificially increased in the upper soft clay as the cone approaches, and is influenced by, the underlying dense sand. The opposite occurs in the dense sand as the cone approaches the underlying soft clay. That is, q^m is artificially decreased when the cone begins to sense the softer material. These errors are commonly referred to as "transistion" or "interface" effects. It can also be seen in Fig.2 that should the dense sand be insufficiently thick for the cone to receive no influence, at least momentarily, from the soft clay either above or below it, then q^m in the dense sand will never reach its true value. Of course, the opposite would occur in a soft, loose stratum sandwhiched by denser strata. While these latter errors have the same root cause as transition effects (i.e., measurement over a volume) they are often separetely referred to as "thin layer" effects.

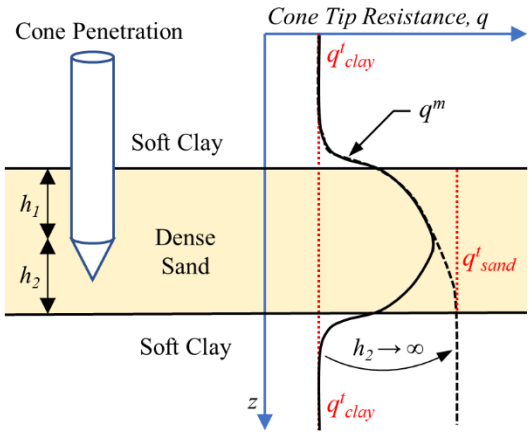


Fig. 2. Conceptual schematic showing cone penetration in layered soil. The measured cone-tip resistance (q^m) deviates from the "true" resistance (q^t) that would be measured in each material if measurements reflected conditions at discrete points (modified from Mo et al. 2017).

Nomenclature aside, the Boulanger and DeJong (2018) procedure aims to correct these collective effects by viewing q^m as the convolution of q^t with a physical low-pass filter (w_c) over a zone of influence (a depth window 60 times that of the cone diameter (d_c), centered at the cone tip). With this approach, q^m is computed as:

$$q^m(z) = q^t(z) * w_c(z) \tag{1}$$

where q^m, q^t, and w_c are a function of depth, z, and the asterisk indicates convolution, which is the integral of the point-wise multiplication of $q^t(z)$ and $w_c(z)$, as a function of the amount that one of the functions is shifted relative to the other. While complete details will not be given here, $w_c(z)$ is a function of two other terms: w_1, which decreases the relative influence of any soil as its distance from the cone tip increases; and w_2, which weights the influence of soil above or below the cone tip based on whether those soils are stronger or weaker than that immediately at the tip. Specifically, it is assumed that q^m receives more influence from the soil immediately near the tip when that soil is relatively weaker than the surrounding soil. Conversely, it is assumed that q^m

receives less influence from the soil immediately near the tip when that soil is relatively stronger than the surrounding soil. In essence, the Boulanger and DeJong (2018) technique identifies the q^t that, when convolved with w_c, best predicts q^m. As part of this procedure, additional filters are used to smooth high frequency noise, thereby increasing the speed and likelihood of convergence to an optimal solution.

Following estimation of q^t via inversion of q^m, Boulanger and DeJong (2018) propose, at least for the time being, that f^t be estimated from q^t. Specifically, the proposed "inversion" of f^m follows the assumption that both inverted and measured pairs of normalized tip resistance (Q) and normalized sleeve friction ratio (F) lie along the same radial line originating from the origin of the Soil Behavior Type Index, I_c, proposed by Robertson and Wride (1998), which maps in Q-F space. In effect, this approach changes I_c near interfaces, and in-turn the inferred soil type (e.g., the inferred susceptibility to liquefaction) but results in relatively minimal changes to I_c otherwise (i.e., away from interfaces). Following the initial development of q^t and f^t profiles, a separate procedure detects and corrects interfaces based on the rate of change of q^t with respect to depth. In total, the Boulanger and DeJong (2018) procedure has five parameters, for which recommended "baseline" values were provided: $z'_{50,ref} = 4.2$; $m_{z50} = 0.5$; $m_z = 3$; $m_q = 2$; and $m_t = 0.1$. While the reader is referred to Boulanger and DeJong (2018) for complete details, these baseline parameters were herein adopted to compute "true" CPT data. These defaults can conceivably be calibrated via site-specific study (e.g., using high resolution borings adjacent to a CPT) to change the magnitude of correction and the sensing and development distances, which may improve identification of thin layers. However, the information compiled for this study either was insufficient to attempt calibration or, when available, provided insufficient statistical support to justify it. As part of the CPT processing methodology, tip- and sleeve-measurements were aligned using statistical cross-

correlation (Buck et al. 2002), both for measured and "true" CPT data. All CPT processing, including implementation of the Boulanger and DeJong (2018) procedure, was completed using the software *Horizon* (Geyin and Maurer, 2020b), a freely available program developed by the authors.

Data Structure

The compiled post-processed data is presented in a structure array (i.e., a single file), with each case history including: identifying information (e.g., ID, geographic coordinates); CPT data, both with and without inverse-filtering and interface correction; earthquake magnitude (M_w); median estimates of the peak ground acceleration (*PGA*) and groundwater table (*GWT*) depth; and the classified occurrence of surficial liquefaction manifestation. The curated dataset is provided in Matlab data format (GLOBALDATASETV1.mat). The data fields, classes, contents, and their units are described in Table 2 and its accompanying footnotes. The structure of the data array is depicted in Fig. 3 and arranged such that case histories are principally sorted by a CPT identification (ID) number. CPT measurements, which are depth-dependent, are l_i x 1 arrays, where l_i is the length of CPT *i*. Other fields include CPT ID, geographic coordinates, M_w, and *GWT*.

Table 2. Data fields, typologies, descriptions, and units.

Field	Class	Content
ID	cell	CPT identifier
NorthingWGS84	double	WGS – y coordinate
Easting WGS84	double	WGS – x coordinate
depth	double	Depth below the ground surface [m]
qc	double	Measured tip resistance [kPa]
qc_inv	double	[1]True tip resistance [kPa]
fs	double	Measured sleeve friction resistance [kPa]
fs_inv	double	[1]True sleeve friction resistance [kPa]
Magnitude	double	Earthquake moment magnitude, M_w
PGA	double	Median peak ground acceleration, PGA [g]
GWT	double	Event-specific groundwater table depth [m]
Manifestation	double	[2]Classified surface manifestation

[1]Processed per the Boulanger and DeJong (2018) procedure; specifically, using the "baseline" inversion model.
[4]The occurrence of surface manifestation was binomially classified as no (coded 0) or yes (coded 2).

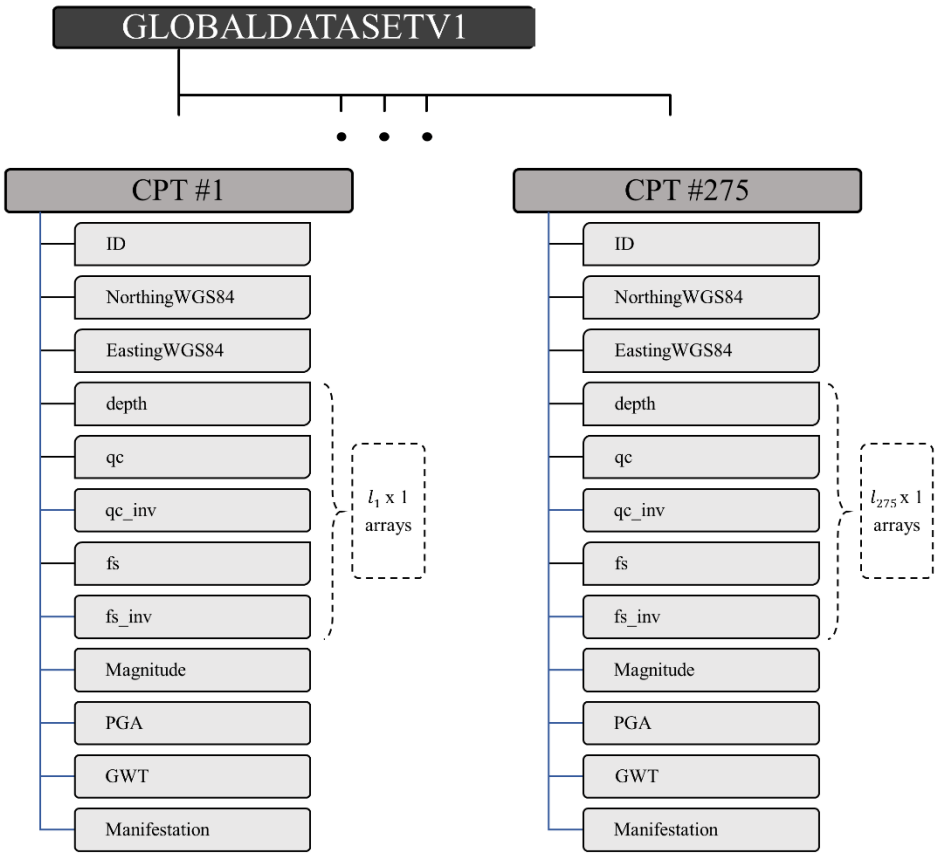


Fig. 3. Depiction of the case-history dataset structure array.

Data Sources

Table 3 summarizes all non-array data contained in the dataset. That is, all information aside from CPT data. The dataset was compiled from the following sources, parsed by event: **1964 M_w7.6 Niigata, JPN** - Ishihara & Koga (1981), Farrar (1990), Moss et al. (2003); **1971 M_w6.6 San Fernando, USA** - Bennett et al., 1998, Toprak and Holzer (2003); **1975 M_w7.0 Haicheng, CHN** - Arulandan et al. (1986), Shengcong & Tatsuoka (1984); **1976 M_w7.6 Tangshan, CHN** - Shibata & Teparaska (1988), Moss et al. (2009; 2011); **1979 M_w6.53 Imperial Valley, USA** - Diaz-Rodriguez (1984), Diaz-Rodriguez and Armijo-Palaio (1991), Moss et al. (2003); **1981 M_w5.9 Westmoreland, USA** - Bennett et al. (1984), Seed et al. (1984), Cetin et al. (2000), Moss et al. (2005); **1983 M_w7.7 Nihonkai-Chubu,**

JPN - Farrar (1990); **1983 M_w6.88 Borah Peak, USA** - Andrus (1986), Andrus & Youd (1987), Moss et al. (2003); **1987 M_w6.6 Edgecumbe, NZ** - Christensen (1995), Moss et al. (2003); **1987 M_w6.54 Superstition Hills, USA** - Bennett et al. (1984), Cetin et al. (2000), Toprak & Holzer (2003), Moss et al. (2005), Holzer & Youd (2007); **1989 M_w6.93 Loma Prieta, USA** - Mitchell et al. (1994), Pass (1994), Bennett & Tinsely (1995), Boulanger et al. (1995; 1997), Kayen et al. (1998), Toprak & Holzer (2003), Youd & Carter (2005); **1994 M_w6.69 Northridge, USA** - Abdel-Haq & Hryciw (1998), Bennett et al., 1998, Holzer et al. (1999), Moss et al. (2003); **1995 M_w6.9 Hyogoken-Nambu, JPN** - Suzuki et al. (2003); **1999 M_w7.51 Kocaeli, TUR** - PEER (2000a), Youd et al. (2009); **1999 M_w7.62 Chi-Chi, TWN** - Lee et al. (2000), PEER (2000b); **2008 M_w6.4 Achaia-Ilia, GRC** - Batilas et al. (2014); **2008 M_w7.2 El Mayor-Cucapah, MEX** - Moss et al. (2005); CESMD (2016), Turner et al. (2016); **2011 M_w9 Tohoku, JPN** - Cox et al. (2013), Boulanger & Idriss (2014); **2012 M_w6.1 Emilia, ITA** - Papathanassiou et al. (2015), Facciorusso et al. (2015), Servizio Geologico (2016).

Table 3. Non-Array Components of the Dataset

CPT ID	Event Date	Event Name	Event Country	Magnitude (M_w)	PGA (g)	GWT (mbg)	Y/N	References
D - Kawagisho-cho	16-Jun-64	Niigata	Japan	7.6	0.162	1.1	Y	Ishihara and Koga (1981); Farrar (1990); Moss et al. (2003)
E - Showa Bridge (Left Bank)	16-Jun-64	Niigata	Japan	7.6	0.162	1.4	Y	Ishihara and Koga (1981); Farrar (1990); Moss et al. (2003)
F - Showa Bridge (Right Bank)	16-Jun-64	Niigata	Japan	7.6	0.162	1.7	N	Ishihara and Koga (1981); Farrar (1990); Moss et al. (2003)
Balboa Boulevard (BAL-10)	9-Feb-71	San Fernando	United States	6.6	0.45	7.2	N	Bennett et al., 1998; Toprak and Holzer (2003)
Balboa Boulevard (BAL-11)	9-Feb-71	San Fernando	United States	6.6	0.45	7.6	N	Bennett et al., 1998; Toprak and Holzer (2003)
17th Middle School	4-Feb-75	Haicheng	China	7	0.3	1	N	Arulandan et al. (1986); Shengcong and Tatsuoka (1984)
Paper Mill	4-Feb-75	Haicheng	China	7	0.3	1.52	Y	Arulandan et al. (1986); Shengcong and Tatsuoka (1984)
Tangshan (T1)	27-Jul-76	Tangshan	China	7.6	0.64	3.7	Y	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T4)	27-Jul-76	Tangshan	China	7.6	0.64	1.1	N	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T6)	27-Jul-76	Tangshan	China	7.6	0.64	1.5	Y	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T7)	27-Jul-76	Tangshan	China	7.6	0.64	3	Y	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T8)	27-Jul-76	Tangshan	China	7.6	0.64	2.2	Y	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T9-2)	27-Jul-76	Tangshan	China	7.6	0.64	1.1	N	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T10)	27-Jul-76	Tangshan	China	7.6	0.64	1.5	Y	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T11)	27-Jul-76	Tangshan	China	7.6	0.61	0.9	Y	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T13)	27-Jul-76	Tangshan	China	7.6	0.58	1.1	Y	Shibata and Teparaska (1988); Moss et al. (2009; 2011)
Tangshan (T16)	27-Jul-76	Tangshan	China	7.6	0.26	3.5	N	Shibata and Teparaska (1988); Moss et al. (2009; 2011)

Kornbloom (KOR4)	15-Oct-79	Imperial Valley	United States	6.53	0.13	2.7	N	Bennett et al. (1984); Moss et al. (2003)
Kornbloom (KOR5)	15-Oct-79	Imperial Valley	United States	6.53	0.13	2.7	N	Bennett et al. (1984); Moss et al. (2003)
McKim Ranch (MCK4)	15-Oct-79	Imperial Valley	United States	6.53	0.51	1.5	N	Bennett et al. (1984); Seed et al. (1984)
McKim Ranch (MCK7)	15-Oct-79	Imperial Valley	United States	6.53	0.51	1.5	Y	Bennett et al. (1984); Seed et al. (1984)
Radio Tower (Rad2)	15-Oct-79	Imperial Valley	United States	6.53	0.2	2.1	Y	Bennett et al. (1984); Seed et al. (1984)
Radio Tower (Rad4)	15-Oct-79	Imperial Valley	United States	6.53	0.2	2.1	N	Bennett et al. (1984); Seed et al. (1984)
River Park (RVP002)	15-Oct-79	Imperial Valley	United States	6.53	0.16	0.3	Y	Moss et al. (2005)
Delta Site 1	9-Jun-80	Victoria (Mexicali)	Mexico	6.33	0.19	2.3	N	Diaz-Rodriguez (1984); Diaz-Rodriguez and Armijo-Palaio (1991); Moss et al. (2003)
Delta Site 2	9-Jun-80	Victoria (Mexicali)	Mexico	6.33	0.19	2.2	Y	Diaz-Rodriguez (1984); Diaz-Rodriguez and Armijo-Palaio (1991); Moss et al. (2003)
Delta Site 3	9-Jun-80	Victoria (Mexicali)	Mexico	6.33	0.19	2	Y	Diaz-Rodriguez (1984); Diaz-Rodriguez and Armijo-Palaio (1991); Moss et al. (2003)
Delta Site 3'	9-Jun-80	Victoria (Mexicali)	Mexico	6.33	0.19	2.2	Y	Diaz-Rodriguez (1984); Diaz-Rodriguez and Armijo-Palaio (1991); Moss et al. (2003)
Delta Site 4	9-Jun-80	Victoria (Mexicali)	Mexico	6.33	0.19	2	Y	Diaz-Rodriguez (1984); Diaz-Rodriguez and Armijo-Palaio (1991); Moss et al. (2003)
Wildlife B (3Cgb)	26-Apr-81	Elmore Ranch	United States	6.22	0.133	1.2	N	Bennett et al. (1984); Cetin et al. (2000)
Wildlife B (3Cgc)	26-Apr-81	Westmorland	United States	5.9	0.26	1.2	Y	Bennett et al. (1984); Cetin et al. (2000)
Herber Road (HEB001)	26-Apr-81	Westmorland	United States	5.9	0.17	1.8	N	Moss et al. (2005)
Kornbloom (KOR4)	26-Apr-81	Westmorland	United States	5.9	0.32	2.7	Y	Bennett et al. (1984); Seed et al. (1984)
Kornbloom (KOR5)	26-Apr-81	Westmorland	United States	5.9	0.32	2.7	Y	Bennett et al. (1984); Seed et al. (1984)
McKim Ranch (MCK4)	26-Apr-81	Westmorland	United States	5.9	0.09	1.5	N	Bennett et al. (1984); Seed et al. (1984)
McKim Ranch (MCK7)	26-Apr-81	Westmorland	United States	5.9	0.09	1.5	N	Bennett et al. (1984); Seed et al. (1984)
Radio Tower (Rad2)	26-Apr-81	Elmore Ranch	United States	6.22	0.09	2.1	N	Bennett et al. (1984)
Radio Tower (Rad2)	26-Apr-81	Westmorland	United States	5.9	0.2	2.1	Y	Bennett et al. (1984); Seed et al. (1984)
Radio Tower (Rad4)	26-Apr-81	Westmorland	United States	5.9	0.2	2.1	N	Bennett et al. (1984); Seed et al. (1984)
River Park (RVP002)	26-Apr-81	Westmorland	United States	5.9	0.17	0.3	N	Moss et al. (2005)

Akita B	26-May-83	Nihonkai-Chubu	Japan	7.7	0.17	1	Y	Farrar (1990)
Akita C	26-May-83	Nihonkai-Chubu	Japan	7.7	0.17	2.4	N	Farrar (1990)
Whiskey Springs Site 1 (CPT-1a)	28-Oct-83	Borah Peak	United States	6.88	0.5	0.8	Y	Andrus & Youd (1987), Moss et al. (2003)
Whiskey Springs Site 2 (CPT-2)	28-Oct-83	Borah Peak	United States	6.88	0.5	2.4	Y	Andrus & Youd (1987), Moss et al. (2003)
Whiskey Springs Site 3 (CPT-3)	28-Oct-83	Borah Peak	United States	6.88	0.5	6.8	Y	Andrus & Youd (1987), Moss et al. (2003)
Awaroa Farm (AWA004)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.37	1.1	Y	Christensen (1995)
Brady Farm (BDY001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.4	1.65	Y	Christensen (1995), Moss et al. (2003)
Brady Farm (BDY004)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.4	1.53	N	Christensen (1995), Moss et al. (2003)
Edgecumbe Pipe Breaks (EPB001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.39	3.5	Y	Christensen (1995), Moss et al. (2003)
Gordon Farm (GDN001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.43	0.5	Y	Christensen (1995), Moss et al. (2003)
Gordon Farm (GDN002)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.43	0.9	N	Christensen (1995), Moss et al. (2003)
Whakatane Hospital (HSP001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.26	4.4	N	Christensen (1995), Moss et al. (2003)
James Street Loop (JSL006)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.28	2	Y	Christensen (1995), Moss et al. (2003)
Keir Farm (KER001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.31	2.5	Y	Christensen (1995), Moss et al. (2003)
Landing Road Bridge LRB007)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.27	1.2	Y	Christensen (1995), Moss et al. (2003)
Morris Farm (MRS001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.42	1.6	Y	Christensen (1995), Moss et al. (2003)
Morris Farm (MRS002)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.42	1.89	N	Christensen (1995)
Morris Farm (MRS003)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.41	2.08	N	Christensen (1995), Moss et al. (2003)
Robinson Farm (RBN001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.44	0.8	Y	Christensen (1995), Moss et al. (2003)
Robinson Farm (RBN002)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.44	0.7	N	Christensen (1995)

Robinson Farm (RBN003)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.44	0.9	N	Christensen (1995)
Robinson Farm (RBN004)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.44	0.61	Y	Christensen (1995), Moss et al. (2003)
Sewage Pumping Station (SPS001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.26	1.3	Y	Christensen (1995), Moss et al. (2003)
Whakatane Board Mill (WBM001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.27	1.44	N	Christensen (1995), Moss et al. (2003)
Whakatane Board Mill (WBM002)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.27	1.44	N	Christensen (1995), Moss et al. (2003)
Whakatane Pony Club (WPC001)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.27	2.35	Y	Christensen (1995), Moss et al. (2003)
Whakatane Pony Club (WPC002)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.27	2.2	Y	Christensen (1995)
Whakatane Pony Club (WPC003)	2-Mar-87	Edgecumbe	New Zealand	6.6	0.27	2.2	N	Christensen (1995)
Wildlife B (3Cg)	24-Nov-87	Superstition Hills	United States	6.54	0.206	1.2	Y	Bennett et al. (1984); Holzer and Youd (2007); Cetin et al. (2000)
Herber Road (HEB001)	24-Nov-87	Superstition Hills	United States	6.6	0.16	1.8	N	Moss et al. (2005)
Kornbloom (KOR4)	24-Nov-87	Superstition Hills	United States	6.54	0.174	2.7	N	Bennett et al. (1984); Cetin et al. (2000)
Kornbloom (KOR5)	24-Nov-87	Superstition Hills	United States	6.54	0.174	2.7	N	Bennett et al. (1984); Cetin et al. (2000)
McKim Ranch (MCK4)	24-Nov-87	Superstition Hills	United States	6.54	0.2	1.5	N	Bennett et al. (1984); Toprak and Holzer (2003)
Radio Tower (Rad2)	24-Nov-87	Superstition Hills	United States	6.54	0.2	2.1	N	Bennett et al. (1984); Cetin et al. (2000)
Radio Tower (Rad4)	24-Nov-87	Superstition Hills	United States	6.54	0.18	2.1	N	Bennett et al. (1984); Cetin et al. (2000)
River Park (RVP002)	24-Nov-87	Superstition Hills	United States	6.6	0.19	0.3	N	Moss et al. (2005)
Alameda Bay Farm Island (Dike)	18-Oct-89	Loma Prieta	United States	6.93	0.24	5.5	N	Mitchell et al. (1994)
Model Airport (AIR-18)	18-Oct-89	Loma Prieta	United States	6.93	0.26	2.4	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Model Airport (AIR-21)	18-Oct-89	Loma Prieta	United States	6.93	0.26	2.4	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Marine Lab (C4)	18-Oct-89	Loma Prieta	United States	6.93	0.28	2.8	Y	Boulanger et al. (1995; 1997)

Miller Farm (CMF-3)	18-Oct-89	Loma Prieta	United States	6.93	0.36	4.9	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Miller Farm (CMF-5)	18-Oct-89	Loma Prieta	United States	6.93	0.36	4.9	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Miller Farm (CMF-8)	18-Oct-89	Loma Prieta	United States	6.93	0.36	4.9	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Miller Farm (CMF-10)	18-Oct-89	Loma Prieta	United States	6.93	0.36	3	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
MBARI 4 (CPT-1)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.9	N	Boulanger et al. (1995; 1997)
General Fish (CPT-6)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.7	N	Boulanger et al. (1995; 1997)
Farris Farm (FAR-58)	18-Oct-89	Loma Prieta	United States	6.93	0.36	4.8	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Farris Farm (FAR-59)	18-Oct-89	Loma Prieta	United States	6.93	0.36	4.8	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Farris Farm (FAR-61)	18-Oct-89	Loma Prieta	United States	6.93	0.36	4.2	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Granite Construction (GRA-123)	18-Oct-89	Loma Prieta	United States	6.93	0.34	5	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Alameda Bay Farm Island (HBI-P6))	18-Oct-89	Loma Prieta	United States	6.93	0.24	3	Y	Mitchell et al. (1994)
Jefferson Ranch (JRR-141)	18-Oct-89	Loma Prieta	United States	6.93	0.21	2.1	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Jefferson Ranch (JRR-148)	18-Oct-89	Loma Prieta	United States	6.93	0.21	3	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
KETT (KET-74)	18-Oct-89	Loma Prieta	United States	6.93	0.47	1.5	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Leonardini (LEN-37)	18-Oct-89	Loma Prieta	United States	6.93	0.22	2.5	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Leonardini (LEN-39)	18-Oct-89	Loma Prieta	United States	6.93	0.22	1.9	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Leonardini (LEN-51)	18-Oct-89	Loma Prieta	United States	6.93	0.22	1.8	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Leonardini (LEN-52a)	18-Oct-89	Loma Prieta	United States	6.93	0.22	2.7	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Leonardini (LEN-53)	18-Oct-89	Loma Prieta	United States	6.93	0.22	2.1	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Martella (MAR-110)	18-Oct-89	Loma Prieta	United States	6.93	0.13	1.8	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)

Martella (MAR-111)	18-Oct-89	Loma Prieta	United States	6.93	0.13	1.7	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
McGowan Farm (MCG-136)	18-Oct-89	Loma Prieta	United States	6.93	0.26	2.4	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
McGowan Farm (MCG-138)	18-Oct-89	Loma Prieta	United States	6.93	0.26	1.8	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Woodward Marine (14-A)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.2	Y	Boulanger et al. (1995; 1997)
Woodward Marine (15-A)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.3	Y	Boulanger et al. (1995; 1997)
Marinovich (MRR-65)	18-Oct-89	Loma Prieta	United States	6.93	0.4	5.6	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Marinovich (MRR-67)	18-Oct-89	Loma Prieta	United States	6.93	0.4	6.2	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Port of Oakland (POO7-2)	18-Oct-89	Loma Prieta	United States	6.93	0.28	3	Y	Mitchell et al. (1994); Kayen et al. (1998)
Port of Oakland (POO7-3)	18-Oct-89	Loma Prieta	United States	6.93	0.28	3	N	Mitchell et al. (1994); Kayen et al. (1998)
Pajaro Dunes (PD1-44)	18-Oct-89	Loma Prieta	United States	6.93	0.22	3.4	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Port of Richmond (POR-2)	18-Oct-89	Loma Prieta	United States	6.93	0.18	2.4	Y	Mitchell et al. (1994); Kayen et al. (1998)
Port of Richmond (POR-3)	18-Oct-89	Loma Prieta	United States	6.93	0.18	2.4	Y	Mitchell et al. (1994); Kayen et al. (1998)
Port of Richmond (POR-4)	18-Oct-89	Loma Prieta	United States	6.93	0.18	2.4	Y	Mitchell et al. (1994); Kayen et al. (1998)
Radovich (RAD-98)	18-Oct-89	Loma Prieta	United States	6.93	0.38	3.5	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Radovich (RAD-99)	18-Oct-89	Loma Prieta	United States	6.93	0.38	4.1	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
MBARI 3 (RC-6)	18-Oct-89	Loma Prieta	United States	6.93	0.28	2.6	N	Boulanger et al. (1995; 1997)
MBARI 3 (RC-7)	18-Oct-89	Loma Prieta	United States	6.93	0.28	3.7	N	Boulanger et al. (1995; 1997)
Sea Mist (SEA-31)	18-Oct-89	Loma Prieta	United States	6.93	0.22	0.8	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
SFO Bay Bridge (SFOBB-1)	18-Oct-89	Loma Prieta	United States	6.93	0.28	3	Y	Mitchell et al. (1994); Kayen et al. (1998)
SFO Bay Bridge (SFOBB-2)	18-Oct-89	Loma Prieta	United States	6.93	0.28	3	Y	Mitchell et al. (1994); Kayen et al. (1998)
Silliman (SIL-68)	18-Oct-89	Loma Prieta	United States	6.93	0.38	3.5	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)

Southern Pacific Bridge (SPR-48)	18-Oct-89	Loma Prieta	United States	6.93	0.33	5.3	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Salinas River Bridge (SRB-116)	18-Oct-89	Loma Prieta	United States	6.93	0.12	6.4	N	Bennett & Tinsely (1995); Moss et al. (2003)
Salinas River Bridge (SRB-117)	18-Oct-89	Loma Prieta	United States	6.93	0.12	6.4	N	Bennett & Tinsely (1995); Moss et al. (2003)
Tanimura (TAN-103)	18-Oct-89	Loma Prieta	United States	6.93	0.13	5	Y	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Tanimura (TAN-105)	18-Oct-89	Loma Prieta	United States	6.93	0.13	4.2	N	Bennett & Tinsely (1995); Toprak & Holzer (2003)
Treasure Island Fire Station (CPTU1)	18-Oct-89	Loma Prieta	United States	6.93	0.16	1.5	N	Pass (1994), Youd and Carter (2005)
Marine Lab (UC-1)	18-Oct-89	Loma Prieta	United States	6.93	0.28	2.4	Y	Boulanger et al. (1995; 1997)
Sandhold Road (UC-2)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.7	N	Boulanger et al. (1995; 1997)
Sandhold Road (UC-3)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.7	N	Boulanger et al. (1995; 1997)
Sandhold Road (UC-6)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.7	N	Boulanger et al. (1995; 1997)
Marine Lab (UC-7)	18-Oct-89	Loma Prieta	United States	6.93	0.28	2.4	Y	Boulanger et al. (1995; 1997)
Woodward Marine (UC-8)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.3	Y	Boulanger et al. (1995; 1997)
Woodward Marine (UC-9)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.2	Y	Boulanger et al. (1995; 1997)
Woodward Marine (UC-10)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1	Y	Boulanger et al. (1995; 1997)
Woodward Marine (UC-11)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1	Y	Boulanger et al. (1995; 1997)
Harbor Office (UC-12)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.9	Y	Boulanger et al. (1995; 1997)
Harbor Office (UC-13)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.9	Y	Boulanger et al. (1995; 1997)
State Beach Kiosk (UC-14)	18-Oct-89	Loma Prieta	United States	6.93	0.28	1.8	Y	Boulanger et al. (1995; 1997)
State Beach Path (UC-16)	18-Oct-89	Loma Prieta	United States	6.93	0.28	2.5	Y	Boulanger et al. (1995; 1997)
State Beach (UC-18)	18-Oct-89	Loma Prieta	United States	6.93	0.28	3.4	N	Boulanger et al. (1995; 1997)

Harbor Office (UC-20)	18-Oct-89	Loma Prieta	United States	6.93	0.28	3	Y	Boulanger et al. (1995; 1997)
Harbor Office (UC-21)	18-Oct-89	Loma Prieta	United States	6.93	0.28	2.7	Y	Boulanger et al. (1995; 1997)
Balboa Boulevard (BAL-10)	17-Jan-94	Northridge	United States	6.69	0.84	7.2	Y	Bennett et al., 1998; Holzer et al. (1999); Moss et al. (2003)
Balboa Boulevard (BAL-11)	17-Jan-94	Northridge	United States	6.69	0.84	7.6	Y	Bennett et al., 1998; Holzer et al. (1999); Moss et al. (2003)
Rory Lane (M-27)	17-Jan-94	Northridge	United States	6.69	0.8	3.4	Y	Abdel-Haq & Hryciw (1998)
Dust Management Facility (DMC)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.37	2	Y	Suzuki et al. (2003)
Fukuzumi Park (FUP-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.65	3.1	N	Suzuki et al. (2003)
Hamakoshienn Housing Area (HAH-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.5	2	Y	Suzuki et al. (2003)
Honjyo Central Park (HCP-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.7	2.5	N	Suzuki et al. (2003)
Imazu Elementary School (IES-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	1.4	Y	Suzuki et al. (2003)
Kobe Art Institute (KAI-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.5	3	N	Suzuki et al. (2003)
Kobe Customs Maya Office A (KMO-A)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	1.8	Y	Suzuki et al. (2003)
Kobe Customs Maya Office A (KMO-B)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	1.8	Y	Suzuki et al. (2003)
Kobe Port Construction Office (KOP-2)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	2.5	Y	Suzuki et al. (2003)
Kobe Pump Station (KPS-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.45	2.6	Y	Suzuki et al. (2003)
Mikuska Park (MIP-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.65	2	Y	Suzuki et al. (2003)
Nagashi Park (NAP-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.65	1	N	Suzuki et al. (2003)
Nisseki Kobe Oil Tank A (NKO-2)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	2.4	Y	Suzuki et al. (2003)

Nisseki Kobe Oil Tank B (NKO-3)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	2.4	Y	Suzuki et al. (2003)
New Port No. 6 Pier (NPP-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	2.5	Y	Suzuki et al. (2003)
New Wharf Construction Offices (NWC-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.45	2.6	Y	Suzuki et al. (2003)
Sumiyoshi Elementary (SES-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.6	1.9	N	Suzuki et al. (2003)
Shimonakajima Park (SHP-2)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.65	2	N	Suzuki et al. (2003)
Shiporex Kogyo Osaka Factory (SKF-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.4	1.5	Y	Suzuki et al. (2003)
Tokuyama Concrete Factory (TCF-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.5	2	Y	Suzuki et al. (2003)
Yoshida Kogyo Factory (YKF-1)	16-Jan-95	Hyogoken-Nambu	Japan	6.9	0.5	3	N	Suzuki et al. (2003)
Adapazari Site C2 (CPT-C4)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	0.4	Y	PEER (2000a)
Adapazari Site K (CPT-K1)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	0.8	Y	PEER (2000a)
Adapazari Site B	17-Aug-99	Kocaeli	Turkey	7.51	0.4	3.3	Y	PEER (2000a)
Adapazari Site D (CPT-D1)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	1.5	Y	PEER (2000a)
Degirmendere DN-1	17-Aug-99	Kocaeli	Turkey	7.51	0.4	1.7	Y	Youd et al. (2009)
Degirmendere DN-2	17-Aug-99	Kocaeli	Turkey	7.51	0.4	2.5	N	Youd et al. (2009)
Adapazari Site E (CPT-E1)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	0.5	Y	PEER (2000a)
Adapazari Site F (CPT-F1)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	0.5	Y	PEER (2000a)
Adapazari Site G (CPT-G1)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	0.5	Y	PEER (2000a)
Adapazari Site H (CPT-H1)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	1.7	Y	PEER (2000a)
Adapazari Site J (CPT-J2)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	0.6	Y	PEER (2000a)

Adapazari Site L (CPT-L1)	17-Aug-99	Kocaeli	Turkey	7.51	0.4	1.72	Y	PEER (2000a)
Police Station PS-1	17-Aug-99	Kocaeli	Turkey	7.51	0.4	1	Y	PEER (2000a)
Soccer Field SF-5	17-Aug-99	Kocaeli	Turkey	7.51	0.37	1	Y	PEER (2000a)
Hotel Spanca SH-4	17-Aug-99	Kocaeli	Turkey	7.51	0.37	0.5	Y	PEER (2000a)
Yalova Harbor YH-3	17-Aug-99	Kocaeli	Turkey	7.51	0.37	1	Y	PEER (2000a)
Nantou Site C (CPT-1)	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	1	Y	PEER (2000b)
Nantou Site C2	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	5	Y	PEER (2000b)
Nantou Site C3	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	1.5	Y	PEER (2000b)
Nantou Site C7 (CPT-7)	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	1	Y	PEER (2000b)
Nantou Site C8 (CPT-8)	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	1	Y	PEER (2000b)
Nantou Site C13	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	2	N	PEER (2000b)
Nantou Site C16	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	5.3	N	PEER (2000b)
Nantou Site K1	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	2.5	Y	PEER (2000b)
Nantou Site K5	20-Sep-99	Chi-Chi	Taiwan	7.62	0.38	2.5	Y	PEER (2000b)
WuFeng Site C7 NCREE	20-Sep-99	Chi-Chi	Taiwan	7.62	0.6	3.2	Y	PEER (2000b)
WuFeng Site C8 NCREE	20-Sep-99	Chi-Chi	Taiwan	7.62	0.6	3.2	N	PEER (2000b)
WuFeng Site C10 NCREE	20-Sep-99	Chi-Chi	Taiwan	7.62	0.6	1.4	Y	PEER (2000b)
WuFeng Site C15 NCREE	20-Sep-99	Chi-Chi	Taiwan	7.62	0.6	2.5	Y	PEER (2000b)
WuFeng Site B (WBC-1)	20-Sep-99	Chi-Chi	Taiwan	7.62	0.6	1.1	Y	PEER (2000b)
WuFeng Site C (WCC-6)	20-Sep-99	Chi-Chi	Taiwan	7.62	0.6	1.2	Y	PEER (2000b)
WuFeng Site K5	20-Sep-99	Chi-Chi	Taiwan	7.62	0.6	1	Y	Lee et al. (2000)
Yanlin Site C2	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	0.56	Y	PEER (2000b)
Yanlin Site C4	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	0.7	Y	PEER (2000b)
Yanlin Site C5	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	2.4	N	PEER (2000b)

Yanlin Site C7	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.7	N	PEER (2000b)
Yanlin Site C8	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.3	N	PEER (2000b)
Yanlin Site C9	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.2	N	PEER (2000b)
Yanlin Site C10	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.9	N	PEER (2000b)
Yanlin Site C11	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	2.9	N	PEER (2000b)
Yanlin Site C13	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.1	N	PEER (2000b)
Yanlin Site C15	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.5	N	PEER (2000b)
Yanlin Site C16	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	2.5	N	PEER (2000b)
Yanlin Site C19	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	0.6	Y	PEER (2000b)
Yanlin Site C22	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.1	Y	PEER (2000b)
Yanlin Site C24	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.2	Y	PEER (2000b)
Yanlin Site C25	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	3.5	Y	PEER (2000b)
Yanlin Site C32	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	0.7	Y	PEER (2000b)
Yanlin Site C38	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.4	N	PEER (2000b)
Yanlin Site C44	20-Sep-99	Chi-Chi	Taiwan	7.62	0.25	1.4	N	PEER (2000b)
Site I (CPT-2)	8-Jun-08	Achaia-Ilia	Greece	6.4	0.18	0.4	Y	Batilas et al. (2014)
Site II (CPT-5)	8-Jun-08	Achaia-Ilia	Greece	6.4	0.18	0.4	N	Batilas et al. (2014)
San Felipito Bridges (CPT-1)	4-Apr-10	Baja	Mexico	7.2	0.265	2	Y	Turner et al. (2016)
Hinode Minami Elementary School	11-Mar-11	Tohoku	Japan	9	0.17	1.1	N	Cox et al. (2013), Boulanger and Idriss (2014)
Hosoyama Nekki	11-Mar-11	Tohoku	Japan	9	0.18	2.5	N	Cox et al. (2013), Boulanger and Idriss (2014)
Takasu Chuou Park	11-Mar-11	Tohoku	Japan	9	0.21	1.1	Y	Cox et al. (2013), Boulanger and Idriss (2014)
Takasu Kaihin Park	11-Mar-11	Tohoku	Japan	9	0.22	1.3	Y	Cox et al. (2013), Boulanger and Idriss (2014)
Akemi Elementary School	11-Mar-11	Tohoku	Japan	9	0.169	1.2	N	Cox et al. (2013), Boulanger and Idriss (2014)
Hinode Elementary School	11-Mar-11	Tohoku	Japan	9	0.199	1.2	Y	Cox et al. (2013), Boulanger and Idriss (2014)
Irifune Nursery School	11-Mar-11	Tohoku	Japan	9	0.256	1.6	Y	Cox et al. (2013), Boulanger and Idriss (2014)
184050U015	20-May-12	Emilia	Italy	6.1	0.22	4.2	N	Papathanassiou et al. (2015); Servizio Geologico (2016)

184050U016	20-May-12	Emilia	Italy	6.1	0.22	4.2	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184050U017	20-May-12	Emilia	Italy	6.1	0.22	2.65	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184050U018	20-May-12	Emilia	Italy	6.1	0.22	3.45	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184060U001	20-May-12	Emilia	Italy	6.1	0.31	3.8	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U047	20-May-12	Emilia	Italy	6.1	0.27	4.2	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U048	20-May-12	Emilia	Italy	6.1	0.27	3.7	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U049	20-May-12	Emilia	Italy	6.1	0.27	3.25	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U050	20-May-12	Emilia	Italy	6.1	0.27	3.2	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U051	20-May-12	Emilia	Italy	6.1	0.25	2.5	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U052	20-May-12	Emilia	Italy	6.1	0.25	2.5	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U053	20-May-12	Emilia	Italy	6.1	0.19	4.9	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U054	20-May-12	Emilia	Italy	6.1	0.19	4.9	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
184090U055	20-May-12	Emilia	Italy	6.1	0.23	3.4	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130B501	20-May-12	Emilia	Italy	6.1	0.59	3	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130B502	20-May-12	Emilia	Italy	6.1	0.6	3	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130B503	20-May-12	Emilia	Italy	6.1	0.59	3	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130B504	20-May-12	Emilia	Italy	6.1	0.59	3	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U006	20-May-12	Emilia	Italy	6.1	0.73	3	N	Papathanassiou et al. (2015); Servizio Geologico (2016)

185130U022	20-May-12	Emilia	Italy	6.1	0.58	1.8	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U505	20-May-12	Emilia	Italy	6.1	0.6	5	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U506	20-May-12	Emilia	Italy	6.1	0.6	4.84	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U507	20-May-12	Emilia	Italy	6.1	0.6	5.16	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U508	20-May-12	Emilia	Italy	6.1	0.59	4.22	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U509	20-May-12	Emilia	Italy	6.1	0.59	2.37	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U510	20-May-12	Emilia	Italy	6.1	0.59	5.23	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U511	20-May-12	Emilia	Italy	6.1	0.59	4.65	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U512	20-May-12	Emilia	Italy	6.1	0.58	4.4	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U513	20-May-12	Emilia	Italy	6.1	0.58	4.18	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185130U514	20-May-12	Emilia	Italy	6.1	0.58	4.55	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
185140B001	20-May-12	Emilia	Italy	6.1	0.54	2	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185140U002	20-May-12	Emilia	Italy	6.1	0.53	2.3	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185140U003	20-May-12	Emilia	Italy	6.1	0.54	3.6	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185140U004	20-May-12	Emilia	Italy	6.1	0.54	3.3	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
185140U005	20-May-12	Emilia	Italy	6.1	0.56	4.9	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
203010U001	20-May-12	Emilia	Italy	6.1	0.65	4	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
203010U002	20-May-12	Emilia	Italy	6.1	0.65	4.9	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)

203010U005	20-May-12	Emilia	Italy	6.1	0.66	2.4	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
203010U006	20-May-12	Emilia	Italy	6.1	0.65	2.4	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
203010U509	20-May-12	Emilia	Italy	6.1	0.72	1.9	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
203020U095	20-May-12	Emilia	Italy	6.1	0.48	4.5	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
203020U096	20-May-12	Emilia	Italy	6.1	0.48	1.6	Y	Papathanassiou et al. (2015); Servizio Geologico (2016)
203020U097	20-May-12	Emilia	Italy	6.1	0.5	2.05	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
203020U098	20-May-12	Emilia	Italy	6.1	0.5	1.9	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
203020U099	20-May-12	Emilia	Italy	6.1	0.51	1.8	N	Papathanassiou et al. (2015); Servizio Geologico (2016)
CPTU1	20-May-12	Emilia	Italy	6.1	0.31	1.4	Y	Papathanassiou et al. (2015); Servizio Geologico (2016); Facciorusso et al. (2015)
Wildlife B (3Cgd)	4-Apr-10	Baja	Mexico	7.2	0.094	1.2	N	Bennett et al. (1984); Holzer and Youd (2007); UCSB (2016)
Herber Road (HEB001)	4-Apr-10	Baja	Mexico	7.2	0.23	1.8	N	Moss et al. (2005); CESMD (2016)

References

Abdel-Haq, A., and Hryciw, R. D., 1998. Ground settlement in Simi Valley following the Northridge earthquake, *Journal of Geotechnical and Geoenvironmental* **124**(1), 80-89.

Ahmadi, M.M. and Robertson, P.K. 2005. Thin-layer effects on the CPT q_c measurement. *Canadian Geotechnical Journal* **42**(5): 1302-1317.

Andrus, R. D. and Youd, T. L. 1987. Subsurface investigation of a liquefaction-induced lateral spread Thousand Springs Valley, Idaho. Misc. paper GL-87-8, U.S. Army Corps of Engineers.

Arulanandan, K., Yogachandran, C., Meegoda, N.J., Ying, L., and Zhauji, S., 1986. Comparison of the SPT, CPT, SV and electrical methods of evaluating earthquake induced liquefaction susceptibility in Ying Kou City during the Haicheng Earthquake. *Use of in situ tests in geotechnical engineering*, *Geotechnical Special Publication* **6**: 389–415.

Batilas, A., Pelekis, P., Vlachakis, V., and Athanasopoulos, G., 2014. Soil Liquefaction/ Nonliquefaction in the Achaia-Ilia (Greece) 2008 Earthquake: field evidence, site characterization and ground motion assessment, *International Journal of Geoengineering Case histories* **2**(4), 270-287.

Bennett, M.J., and Tinsley, J.C., III 1995. Geotechnical data from surface and subsurface samples outside of and within liquefaction-related ground failures caused by the October 17, 1989, Loma Prieta earthquake, Santa Cruz and Monterey Counties, CA. U.S. Geological Survey Open-File Report 95-663.

Bennett, M.J., McLaughlin, P.V., Sarmiento, J.S., and Youd, T. L., 1984. Geotechnical investigation of liquefaction sites, Imperial Valley, CA. U.S. Geological Survey, Open-File Report 84-252.

Bennett, M.J., Ponti, D.J., Tinsley, J.C., III, Holzer, T.L., and Conaway, C.H., 1998. Subsurface geotechnical investigations near sites of ground deformations caused by the January 17, 1994, Northridge, California, earthquake. U.S. Geological Survey, Open-File Report 98-373.

Boulanger, R.W. and DeJong, J.T. (2018). Inverse filtering procedure to correct cone penetration data for thin-layer and transition effects. *Cone Penetration Testing 2018*, Hicks, Pisano, and Peuchen, eds., Delft University of Technology, The Netherlands: 25-44.

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.

Boulanger, R.W., Idriss, I.M., and Mejia, L.H., 1995. Investigation and evaluation of liquefaction related ground displacements at Moss Landing during the 1989 Loma Prieta earthquake. Report No. UCD/CGM-95/02, Center for Geotechnical Modeling, Dept. of Civil & Environmental Engineering, University of California, Davis, 231 pp.

Boulanger, R.W., Mejia, L.H., and Idriss, I.M., 1997. Liquefaction at Moss Landing during Loma Prieta earthquake. *Journal of Geotechnical & Geoenvironmental Engineering* **123**(5), 453–67.

Buck, J.R., Daniel, M.M, and Singer, A.C, 2002. *Computer Explorations in Signals and Systems Using MATLAB®*, 2nd Edition. Upper Saddle River, NJ: Prentice Hall.

CESMD (Center for Engineering Strong Motion Data), 2016. U.S. structural and ground response data; www.strongmotioncenter.org/

Cetin, K.O., Seed, R.B., Moss, R.E.S., Der Kiureghian, A.K., Tokimatsu, K., Harder, L.F., and Kayen, R.E., 2000. Field Performance Case Histories for SPT-Based Evaluation of Soil Liquefaction Triggering Hazard, Geotechnical Engineering Research Report No. UCB/GT-2000/09, UC Berkeley.

Christensen, S.A., 1995. Liquefaction of cohesionless soils in the March 2, 1987 Edgecumbe earthquake, Bay of Plenty, New Zealand, and other earthquakes, Master of Engineering Thesis, Department of Civil Engineering, University of Canterbury, Christchurch, New Zealand.

Cox, B.R., Boulanger, R.W., Tokimatsu, K., Wood, C., Abe, A., Ashford, S., Donahue, J., Ishihara, K., Kayen, R., Katsumata, K., Kishida, T., Kokusho, T., Mason, B., Moss, R., Stewart, J., Tohyama, K., and Zekkos, D., 2013. Liquefaction at strong motion stations in Urayasu City during the 2011 Tohoku-Oki earthquake." *Earthquake Spectra* **29**(S1), S55-S80.

Diaz-Rodriquez, J.A., 1984. Liquefaction in the Mexicali Valley during the earthquake of June 9, 1980, *Eighth World Conference on Earthquake Engineering*, San Francisco, 223-230.

Diaz-Rodriquez, J.A., and Armijo-Palacio, G., 1991. Liquefaction potential of fine cohesionless soils using the CPT, *Soils and Foundations* **31**(3), 111-119.

Farrar, J.A., 1990. Study of in situ testing for evaluation of liquefaction resistance, R-90-06, U.S. Department of the Interior, Bureau of Reclamation, Geotechnical Services Branch, Denver Office.

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

Geyin, M. and Maurer, B.W. 2020b. Horizon: CPT-based liquefaction risk assessment and decision software. DesignSafe-CI, doi: 10.17603/ds2-2fky-tm46.

Holzer, T.L., and Youd, T.L., 2007. Liquefaction, ground oscillation, and soil deformation at the wildlife array, California, *Bulletin of the Seismological Society of America* **97**(3), 961-976.

Holzer, T.L., Bennett, M.J., Ponti, D.J., and Tinsley, J.C., 1999. Liquefaction and soil failure during 1994 Northridge earthquake, *Journal of Geotechnical and Geoenvironmental Engineering*, **125**(6), 438-452.

Ishihara, K., and Koga, Y., 1981. Case studies of liquefaction in the 1964 Niigata earthquake, *Soils and Foundations* **21**(3), 35-52.

Kayen, R.E., Mitchell, J.K., Seed, R.B., and Nishio, S., 1998. Soil liquefaction in the east bay during the earthquake, *The Loma Prieta, California, Earthquake of October 17, 1989 – Liquefaction.* Thomas L. Holzer, editor, USGS Professional Paper 1551-B, B61-B86.

Lee, D.H., Ku, C.S., and Juang, C.H., 2000. Preliminary investigation of soil liquefaction in the 1999 Chi-Chi, Taiwan, earthquake, *International Workshop on Annual Commemoration of Chi-Chi Earthquake, Vol. III – Geotechnical Aspect,* C.H. Loh and W.I. Liao, eds., National Center for Research on Earthquake Engineering, Taipei, Taiwan, pp. 140-151.

Lunne, T., Robertson, P.K. and Powell, J.M. (1997). *Cone Penetration Testing in Geotechnical Practice*. Blackie Academic & Professional, London, U.K.

Mitchell, J.K., Lodge, A.L., Coutinho, R.Q., Kayen, R.E., Seed, R.B., Nishio, S., and Stokoe, K. H., II 1994. Insitu test results from four Loma Prieta earthquake liquefaction sites: SPT, CPT, DMT and shear wave velocity. Report No. UCB/EERC-94/04. Earthquake Engineering Research Center, UC Berkeley.

Mo, P.Q., Marshall, A.M. and Yu, H.S., 2017. Interpretation of cone penetration test data in layered soils using cavity expansion analysis. *Journal of Geotechnical and Geoenvironmental Engineering*, ***143***(1), p.04016084.

Moss, R.E.S., Collins, B.D., and Whang, D.H., 2005. Retesting of liquefaction/nonliquefaction case histories in the imperial valley. *Earthquake Spectra* **21**(1), 179–196.

Moss, R.E.S., Kayen, R.E., Tong, L.-Y., Liu, S.-Y., Cai, G.-J., and Wu, J., 2009. Re-investigation of liquefaction and nonliquefaction case histories from the 1976 Tangshan earthquake. Rep. No. 209/102, Pacific Earthquake Engineering Research (PEER) Center, Berkeley, CA.

Moss, R.E.S., Seed, R.B., Kayen, R.E., Stewart, J. P., Youd, T.L., and Tokimatsu, K. 2003. Field case histories for CPT-based in situ liquefaction potential evaluation. Geoengineering Research Rep. UCB/GE-2003/04.

Papathanassiou, G., Mantovani, A., Tarabusi, G., Rapti, D., and Caputo, R., 2015. Assessment of liquefaction potential for two liquefaction prone areas considering the May 20, 2012 Emilia (Italy) earthquake, *Engineering Geology*, **189**, 1-16.

Pass, D.G., 1994. Soil characterization of the deep accelerometer site at Treasure Island, San Francisco, California. MS thesis, University of New Hampshire.

PEER (Pacific Earthquake Engineering Research Center), 2000a. Documenting incidents of ground failure resulting from the 1999 Kocaeli, Turkey Earthquake.

PEER (Pacific Earthquake Engineering Research Center), 2000b. Documentation of soil conditions at liquefaction sites from 1999 Chi-Chi, Taiwan Earthquake.

Robertson, P.K. 2011. Automated detection of CPT transition zones. *Geotechnical News*, June: 35-38.

Robertson, P.K., and Wride, C.E., 1998. Evaluating cyclic liquefaction potential using cone penetration test. Canadian Geotechnical Journal **35**(3), 442-459.

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

Servizio Geologico, 2016. Database of the Emilia-Romagna region: geological, seismic, and soil Survey." http://ambiente.regione.emiliaromagna.it/geologia/cartografia/webgis-banchedati/.

Shengcong, F. and Tatsuoka, F., 1984. Soil liquefaction during Haicheng and Tangshan earthquake in China; a review, *Soils and Foundations* **24**(4), 11-29.

Shibata, T., and Teparaska, W., 1988. Evaluation of Liquefaction Potential of Soils Using Cone Penetration Testing. *Soils and Foundations* **28**(2), 49-60.

Suzuki, Y., Tokimatsu, K., Moss, R.E.S., Seed, R.B., and Kayen, R.E., 2003. CPT-based liquefaction case histories from the 1995 Hyogoken-Nambu (Kobe) earthquake, Japan. Geotechnical Engineering Research Report No. UCB/GE-2003/03.

Toprak, S., and Holzer, T.L., 2003. Liquefaction potential index: field assessment, *Journal of Geotechnical and Geoenvironmental Engineering* **129(**4), 315-322.

Treadwell, D.D. (1976). The influence of gravity, prestress, compressibility, and layering on soil resistance to static penetration. PhD Dissertation, Univ. of California, Berkeley, CA.

Turner, B., Brandenberg, S., and Stewart, J. (2016). Case Study of Parallel Bridges Affected by Liquefaction and Lateral Spreading. *Journal of Geotechnical and Geoenvironmental Engineering* **142**(7), 05016001.

UCSB, 2016. *Geotechnical Array Monitoring Program* < http://nees.ucsb.edu/data-portal>.

van der Linden, T.I., De Lange, D.A. and Korff, M. (2018). Cone penetration testing in thinly inter-layered soils. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering* **171**(3): 215-231.

Youd, T.L., DeDen, D.W., Bray, J.D., Sancio, R., Cetin, K.O., and Gerber, T.M., 2009. Zero displacement lateral spreads, 1999 Kocaeli, Turkey, earthquake. *Journal of Geotechnical & Geoenvironmental Engineering*, **135**(1), 46-61.

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	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology,* In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

FIELD ASSESSMENT OF LIQUEFACTION PREDICTION MODELS BASED ON GEOTECHNICAL VS. GEOSPATIAL DATA, WITH LESSONS FOR EACH

Mertcan Geyin,a) M.EERI, Alex J. Baird,a) and Brett W. Maurer,a) M.EERI

Abstract

Semi-empirical models based on in-situ geotechnical tests have been the standard-of-practice for predicting soil liquefaction since 1971. More recently, prediction models based on free, readily available data were proposed. These "geospatial" models rely on satellite remote-sensing to infer subsurface traits without in-situ tests. Using 15,223 liquefaction case-histories from 24 earthquakes, this study assesses the performance of 23 models based on geotechnical or geospatial data using standardized metrics. Uncertainty due to finite sampling of case histories is accounted for and used to establish statistical significance. Geotechnical predictions are significantly more efficient on a global scale, yet successive models proposed over the last twenty years show little or no demonstrable improvement. In addition, geospatial models perform equally well for large subsets of the data – a provocative finding given the relative time- and cost-requirements underlying these predictions. Through this performance comparison, lessons for improving each class of model are elucidated in detail.

Introduction

Since the inception of the so-called "simplified stress-based procedure" for predicting liquefaction triggering (Seed & Idriss, 1971; Whitman, 1971), variants based on several in-situ geotechnical measurements have been developed. These include cone penetration test (CPT) indices, standard penetration test (SPT) blow counts, and shear-wave velocity (Vs), among others.

In conjunction with such measurements, “simplified” liquefaction triggering models have been used in virtually every seismic zone on Earth. The outputs from these triggering models are also often used cooperatively with other models that predict surface manifestations, such as ground settlement (e.g., Cetin et al., 2009), lateral spreading (e.g., Zhang et al., 2004), liquefaction ejecta (e.g., Maurer et al., 2017), and foundation movements (e.g., Bray and Macedo, 2017; Bullock et al., 2018). While this approach to modelling liquefaction occurrence and consequence is popular worldwide, it requires field measurements that can be costly and time consuming to perform, especially over large areas. More recently, “geospatial” models were proposed to predict liquefaction using data that is freely available via satellite remote sensing (Zhu et al., 2015, 2017). Like the geotechnical approach, geospatial models characterize liquefaction demand using ground-motion intensity measures (IMs). But, in lieu of characterizing liquefaction resistance via in-situ measurements, geospatial models use surface parameters to infer subsurface traits. Examples include surface slope, mineralogy, roughness, and wetness; distance to and elevation above rivers, streams, and other water bodies; and compound-topographic-index, which can all be derived from satellite data. Geospatial models are particularly suited for applications in which time- and cost-considerations outweigh the required, expected model accuracy (e.g., regional earthquake simulations; planning and policy development; post-event response and reconnaissance; and hazard assessments in regions that lack geotechnical testing). Implicit to this statement, of course, is the assumption that geotechnical models are more accurate than geospatial models. But are they?

The efficacies of these two model classes (i.e., geotechnical and geospatial) have not been directly compared using a consistent set of case histories and standardized performance metrics. Such an assessment could elucidate pathways to improve each model class (e.g., what can each class teach the other?) and inform ensemble-modeling approaches that statistically coalesce

multiple predictions (Bradley et al., 2018). In addition, while numerous "simplified" geotechnical models have been proposed over the last twenty years, the popular question of "which performs best?" remains contentious – the answer obscured, in part, by prior paucity and inconsistency of test cases, as well as use of differing, non-standard, non-objective performance metrics. Further, though often ignored, the performance of any model is intimately tied to the site-specific consequences, or "economies" of misprediction. We should thus ask not only "which model performs best?" but also "which model performs best for particular misprediction economies?"

Accordingly, the objective of this study is to rigorously assess and compare the performance of 18 geotechnical models and 5 geospatial models using 15,223 liquefaction case-histories compiled from 24 earthquakes in 9 countries. As part of this assessment, standardized and objective metrics will be used to evaluate model performance, both in an overall, comprehensive sense, as well as for particular misprediction economies that may be encountered.

Geotechnical and Geospatial Liquefaction Models

The hierarchy of liquefaction models may be loosely defined by three tiers: (T1) wholly-empirical models that require only geologic or geospatial data and are accessible to a broad userbase (e.g., Youd and Hoose, 1977; Kramer, 2008; FEMA, 2013; Zhu et al., 2017); (T2) semi-mechanistic "simplified stress-based" models that require in-situ measurements and are generally limited to use by geoengineers (e.g., Kayen et al., 2013; Boulanger and Idriss, 2014); and (T3) wholly-mechanistic constitutive models, which typically require many soil and model parameters, and which are generally limited to use by geoengineers trained in computational mechanics (e.g., Cubrinovski and Ishihara, 1998; Byrne et al., 2004; Ziotopoulou and Boulanger, 2016). While advances to "T3" models and the enabling technologies have grown their use, their application to

predicting liquefaction triggering remains relatively rare. This is due to the required inputs and operator skill, but also to uncertainties about model adequacy and the interpretation of results (NRC, 2016). As such, these models are not evaluated in this study, but are mentioned to provide proper context to the assessment of "T1" and "T2" models that is presented herein.

For brevity, the "T2, simplified stress-based models" are henceforth referred to as "geotechnical models." Six such liquefaction triggering models will be used in this study: Robertson and Wride (1998), Architectural Institute of Japan (2001), Moss et al. (2006), Idriss and Boulanger (2008), Boulanger and Idriss (2014), and Green et al. (2018). All six are based on the CPT, which offers significant advantages over other in-situ tests (NRC, 2016). However, because triggering models predict liquefaction at-depth within a soil profile, an evaluation of their performance requires direct investigation of the subsurface (i.e., to assess the agreement between predicted and actual responses in various strata). This might be accomplished proactively using downhole instrument-arrays (e.g., Holzer et al., 2007) or reactively using vision penetrometers (Raschke and Hryciw, 1997) or geoslicers (Nakata and Shimazaki, 1997). Yet, such investigations are generally expensive, exceedingly rare, and may not result in definitive interpretations (e.g., Takada and Atwater, 2004). Thus, nearly all existing case-histories simply document whether manifestations of liquefaction were observed at the surface. Accordingly, to compare predictions of liquefaction at-depth against surface observations, the six triggering models will each be used in series with three separate manifestation models: Iwasaki et al. (1978), van Ballegooy et al. (2014a), and Maurer et al. (2015a), who respectively proposed models named *LPI*, *LSN*, and LPI_{ISH}. It must therefore be understood that "geotechnical model" refers to the combined use of a triggering model and a manifestation model. There is simply no practical or objective way to isolate and assess the independent performance of triggering models. While *LPI*, *LSN*, and LPI_{ISH}

were each proposed as general "ground failure" predictors, they have been calibrated almost exclusively on the observed occurrence or nonoccurrence of liquefaction ejecta.

The 18 CPT-based geotechnical models (6 triggering models · 3 manifestation models) will be tested against the performance of 5 "T1" geospatial models. While several types of "T1" model exist, those of Zhu et al. (2015, 2017) are arguably the most rigorously-formulated and well-trained to date; they were also recently implemented into US Geological Survey post-earthquake data products to provide automated content on possible impacts (e.g., Allstadt et al., 2019). These models were trained on observations of ground-failure and thus inherently merge liquefaction triggering and manifestation. Three of the five models to be assessed are region specific - developed specifically for Canterbury, New Zealand – and will be referred to as regional geospatial models (RGMs). The remaining two were trained on successively larger datasets from global earthquakes and will be referred to as global geospatial models (GGMs). A summary of the 23 models to be evaluated, and the symbols henceforth used to identify them, is provided in Table 1. Additional model details are provided subsequently in the *Methodology*.

Table 1. Summary of Geotechnical and Geospatial Liquefaction Models Evaluated in this Study

Geotechnical Models (18)				**Geospatial Models (5)**	
Triggering Model	**Symbol**	**Manifestation Model**	**Symbol**	**Triggering/Manifestation Model**	**Symbol**
Robertson & Wride (1998)	RW98	Iwasaki et al. (1978)	*LPI*	Zhu et al. (2015) Regional 1	*RGM1*
Arch. Institute Japan (2001)	AIJ01	van Ballegooy et al. (2014a)	*LSN*	Zhu et al. (2015) Regional 2	*RGM2*
Moss et al. (2006)	Mea06	Maurer et al. (2015a)	LPI_{ISH}	Zhu et al. (2015) Regional 3	*RGM3*
Idriss & Boulanger (2008)	IB08			Zhu et al. (2015) Global	*GGM1*
Boulanger & Idriss (2014)	BI14			Zhu et al. (2017) Global	*GGM2*
Green et al. (2019)	Gea19				

Data

This study analyzes 15,223 liquefaction case histories resulting from 24 earthquakes, as summarized in Table 2. However, because most of these cases were compiled from three earthquakes in the Canterbury region of New Zealand, results are separately presented for these and the other 21 earthquakes, henceforth respectively referred to as the "Canterbury dataset" and "global dataset." The details of these case-history datasets are discussed next.

Table 2. Summary of Liquefaction Case-Histories Analyzed

Date	Earthquake	Country	Magnitude (M_w)	Number of Case Histories
16/6/1964	Niigata	Japan	7.60	3
9/2/1971	San Fernando	USA	6.60	2
4/2/1975	Haicheng	China	7.00	2
27/7/1976	Tangshan	China	7.60	10
15/10/1979	Imperial Valley	USA	6.53	7
9/6/1980	Victora (Mexicali)	Mexico	6.33	5
26/4/1981	Westmoreland	USA	5.90	9
26/5/1983	Nihonkai-Chubu	Japan	7.70	2
28/10/1983	Borah Peak	USA	6.88	3
2/3/1987	Edgecumbe	New Zealand	6.60	23
24/11/1987	Elmore Ranch	USA	6.22	2
24/11/1987	Superstition Hills	USA	6.54	8
18/10/1989	Loma Prieta	USA	6.93	67
17/1/1994	Northridge	USA	6.69	3
16/1/1995	Hyogoken-Nambu	Japan	6.90	21
17/8/1999	Kocaeli	Turkey	7.51	16
20/9/1999	Chi-Chi	Taiwan	7.62	34
8/6/2008	Achaia-Ilia	Greece	6.40	2
4/4/2010	Baja	Mexico	7.20	3
4/10/2010	Darfield	New Zealand	7.10	5371
22/2/2011	Christchurch	New Zealand	6.20	4806
11/3/2011	Tohoku	Japan	9.00	7
20/5/2012	Emilia	Italy	6.10	46
14/2/2016	Christchurch	New Zealand	5.70	4771
			Total	**15,223**

CANTERBURY EARTHQUAKE DATASET

Earthquakes occurring over the last decade in the Canterbury region of New Zealand have resulted in case-history data of unprecedented quantity and quality. A comprehensive summary of these earthquakes, to include tectonic and geologic settings, seismology, and effects, is provided by Quigley et al. (2016). The present study compiles case-histories from the M_w7.1, 4 Sept. 2010 Darfield earthquake, the M_w6.2, 22 Feb. 2011 Christchurch earthquake, and the M_w5.7, 14 Feb. 2016 Christchurch earthquake. This effort built upon a series of successive compilations (Maurer et al., 2014, 2015b, 2019), augmenting the largest of these by more than 50% and resulting in a total of 14,948 case histories. These consist of ground-motion intensity measures, geotechnical and hydrological data, readily available geospatial information, and classifications of liquefaction manifestations. These components are succinctly summarized as follows.

Liquefaction Manifestations

All liquefaction models will be evaluated on their ability to predict free-field surface manifestations on level ground – specifically liquefaction ejecta - rather than any other indicator, such as evidence from ground-motions, foundation movements, or lateral spreading. Sites with these indicators were expressly removed from the study because the models to be evaluated are not intended to predict them. Observations of the occurrence and severity of liquefaction ejecta were compiled by the authors and classified as “none,” “minor,” “moderate,” and “severe” using criteria from Green et al. (2014) and a 15-m sampling radius centered on each CPT site. This was accomplished using ground-reconnaissance reports and high-resolution satellite imagery available in the New Zealand Geotechnical Database (CERA, 2012). Cases in which manifestations could not be reliably classified are not included in this study. To facilitate model assessment, the cases

are reclassified binomially as “No Manifestation” and “Manifestation,” where the latter are sites with at least “minor” manifestations per Green et al. (2014). Of the resulting 14,948 case histories compiled from Canterbury, 65% are “No Manifestation” and 35% are “Manifestation.”

Ground-Motion Intensity Measures

The 23 liquefaction models use either peak ground acceleration, *PGA* (all geotechnical models and some geospatial models) or peak ground velocity, *PGV* (some geospatial models). For this study, *PGA*s were estimated with the Bradley (2013) procedure, which has been used in previous Canterbury earthquake research (e.g., van Ballegooy et al. 2015), and which geostatistically merges *PGA*s recorded at strong-motion stations with *PGA*s estimated by ground-motion prediction equations. *PGV*s were estimated using USGS ShakeMap (Worden and Wald, 2016), consistent with the formulation of geospatial liquefaction models that use *PGV*.

Geotechnical, Hydrological, and Geospatial Data

This study analyzes CPT soundings available in the New Zealand Geotechnical Database (CERA, 2012) and performed at sites where liquefaction manifestations were classified as described above. In the process of compiling case-histories, CPTs were rejected: (1) if the depth of “pre-drill” significantly exceeded the depth of the ground water table; and (2) if inferred to have prematurely terminated on shallow gravels from a geospatial autocorrelation analysis (Anselin, 1995). Prior to processing, CPT tip- and sleeve-measurements were aligned using cross-correlation (Buck et al., 2002). Extended coverage of CPT data and the exclusion criteria summarized above is provided in Maurer et al. (2014, 2015b). Ground water table (GWT) depths were sourced from the robust, event-specific regional models of van Ballegooy et al. (2014b). Various geospatial parameters – to be identified subsequently in the *Methodology* – were computed at each case-

history location following the exact methods of Zhu et al. (2015; 2017) (i.e., the developers of the geospatial models to be evaluated).

GLOBAL EARTHQUAKE DATASET

To compare findings in Canterbury with regions worldwide, 275 case histories are compiled from 21 global earthquakes and assessed in parallel. These cases are sourced from the literature, to include observations of manifestation severity, CPT soundings, and estimation of GWT depth and ground-motion IMs, as generally reported by original investigators. When available, recent refinements are adopted from the literature. Whereas liquefaction was intensively cataloged via reconnaissance and remote sensing in Canterbury, the global cases are often documented in less detail, occasionally with scant information about the nature or severity of manifestation. Accordingly, while the Green et al. (2014) criteria were again used to binomially classify manifestations based on ejecta (while excluding cases with other expressions of liquefaction), there is inevitably some uncertainty. Of the 275 global cases, 58% are "Manifestation" and 42% are "No Manifestation." To properly recognize all sources of data used to compile this dataset, a separate reference list appears near the end of this paper, parsed by earthquake; a table of data for each case-history is also provided in an electronic supplement. In this regard, the case history assemblages of Moss et al. (2003) and Boulanger and Idriss (2014) are acknowledged for greatly assisting the present study. Lastly, various geospatial parameters (identified in the *Methodology*) – were computed at each case-history location per Zhu et al. (2015; 2017).

Methodology

The 23 liquefaction models to be evaluated in this study were identified in Table 1. Additional details are now presented, followed by methods that will be used to analyze model performance.

Geotechnical Model Methodology

CPTs were analyzed using six triggering models, all of which compute factor-of-safety against liquefaction (FS_{liq}) vs. depth. While the reader is referred to the model publications for complete details, nuances pertinent to this study are as follows. First, prior to using any of the six models, liquefaction-susceptible soils were inferred from the CPT soil-behavior-type index (I_c) (Robertson and Wride, 1998), such that soils with I_c < 2.50 were assumed susceptible. This criterion was developed specifically for Christchurch soils from extensive lab and field testing (Maurer et al., 2019). However, because an I_c threshold of 2.50 is within the range of generic values commonly used in practice (e.g., 2.4-2.6), this criterion is also used in all analyses of the global dataset. Ultimately, the results of these analyses were found to be insensitive to this decision. Second, for liquefaction-susceptible soils, the IB08, BI14, and Gea19 models compute liquefaction resistance as a function of fines-content (*FC*). Accordingly, *FC* was estimated for the Canterbury dataset using a Christchurch-specific I_c – *FC* correlation (Maurer et al., 2019), and for the global dataset using a general I_c – *FC* correlation (Boulanger and Idriss, 2014), with the former estimating *FC* to be approximately 10% higher for a given I_c.

Next, the outputs from triggering analysis were input to the *LPI*, *LSN*, and LPI_{ISH} manifestation models. The Liquefaction Potential Index (*LPI*) is defined as (Iwasaki et al., 1978):

$$LPI = \int_0^{20\,m} F(FS_{liq}) \cdot w(z)\,\mathrm{d}z \tag{1}$$

where $F(FS_{liq})$ and $w(z)$ are functions that weight the respective influences of FS_{liq} and depth, z, on surface manifestation. Specifically, $F(FS_{liq}) = 1 - FS_{liq}$ for $FS_{liq} \leq 1$ and $F(FS_{liq}) = 0$ otherwise; $w(z) = 10 - 0.5z$. Thus, *LPI* assumes that surface manifestation depends on the thickness of all liquefied strata in a profile's upper 20 m, their proximity to the ground surface, and the amount by which FS_{liq} in each stratum is less than 1.0. Given this definition, *LPI* can range from zero to 100.

A modified *LPI* was proposed by Maurer et al. (2015a) and inspired by Ishihara (1985), who proposed limit-state curves for predicting manifestations as a function of the "crust" thickness (H_1), among other factors. Using these curves, Maurer et al. (2015a) modified *LPI* to include the observed influence of H_1. Given its provenance, the result was termed LPI_{ISH} and is defined by:

$$LPI_{ISH} = \int_{H_1}^{20\ m} F(FS_{liq}) \cdot w(z)\ \mathrm{d}z \tag{2a}$$

where

$$F(FS_{liq}) = \begin{cases} 1 - FS_{liq} & if\ FS_{liq} \leq 1\ \cap\ H_1 \cdot m(FS_{liq}) \leq 3 \\ 0 & otherwise \end{cases} \tag{2b}$$

$$m(FS_{liq}) = \exp\left(\frac{5}{25.56(1-FS_{liq})}\right) - 1 \tag{2c}$$

In Eq. (2a), $F(FS_{liq})$ and $w(z)$ have the same objective as in *LPI*, but are functionally different, such that $F(FS_{liq})$ accounts for the crust thickness via the parameter H_1 and $w(z)$ is defined by $w(z) = 25.56 \cdot z^{-1}$. Maurer et al. (2015a) recommended a minimum H_1 of 0.4 m be used, even if liquefiable soils are present at shallower depths. Given this constraint, LPI_{ISH} can range from zero to 100.

The Liquefaction Severity Number (*LSN*) is an adaptation of methods for estimating post-liquefaction volumetric strain (e.g., to predict ground settlement), modified to include a power-law depth weighting function (van Ballegooy et al., 2014a):

$$LSN = \int_{0}^{20\ m} \varepsilon_v \cdot w(z)\ \mathrm{d}z \tag{3}$$

where ε_v is volumetric strain (%) and $w(z) = 10 \cdot z^{-1}$. For a given value of FS_{liq}, ε_v is inversely related to the soil's initial relative density (D_r). By corollary, surface manifestations should diminish as the D_r of liquefying soil increases. While there are several approaches to estimating ε_v (Geyin and Maurer, 2019), van Ballegooy et al. (2014a) used that of Zhang et al. (2002), which we thus adopt. *LSN* values can far surpass 100 when liquefiable soils are present at the surface, but

typically are between zero and 100. These values are not quantities of predicted settlement, but rather, are index values á la *LPI* and LPI_{ISH} that correlate to the likelihood of surface manifestation.

Geospatial Model Methodology

The five geospatial models have the general form $P(X) = (1 + e^{-X})^{-1}$ where X is a series of geospatial variables and model coefficients, and $P(X)$ is the likelihood of surface manifestation. In conjunction with this equation, the three region-specific and two global geospatial models are defined in Table 3. The variables are as follows: PGA_M = magnitude-weighted peak ground acceleration (g) using the weighting of Youd et al. (2001); *PGV* = peak ground velocity (cm/s); d_{r3} = distance (km) to a stream of order three or greater (Strahler, 1952); V_{s30} = shear-wave velocity of the upper 30-m (m/s), inferred from surface topography (Wald and Allen, 2007); d_r = shortest distance to a river (km) cataloged by Lehner et al. (2006); *CTI* = compound topographic index (Beaven and Kirkby, 1979); d_c = distance to coast (km); *precip* = mean annual precipitation (mm) (Fick and Hijmans, 2017); *ND* = d_c divided by the distance from the coast to the edge of the sedimentary basin; and *wtd* = water table depth (m). *GGM2* has variants for coastal and inland locations, the distinguishing threshold being d_c = 20 km. All variables were computed in accordance with Zhu et al. (2015; 2017), to which the reader is referred for additional information. Geospatial model predictions were generated at 30-m resolution, commensurate with the sampling area over which liquefaction manifestations were classified.

Table 3. Geospatial Liquefaction Model Equations

Model	Model Parameter X
RGM1	$2.053 + 1.267 \cdot \ln(PGA_M) - 0.239 \cdot d_{r3} - 9.191 \cdot ND$
RGM2	$0.316 + 1.225 \cdot \ln(PGA_M) + 0.145 \cdot CTI - 9.708 \cdot ND$
RGM3	$25.45 + 2.476 \cdot \ln(PGA_M) - 0.323 \cdot d_{r3} - 4.241 \cdot \ln(V_{s30})$
GGM1	$24.10 + 2.067 \cdot \ln(PGA_M) + 0.355 \cdot CTI - 0.4784 \cdot \ln(V_{s30})$
GGM2 (Coastal)	$12.435 + 0.301 \cdot \ln(PGV) - 2.615 \cdot \ln(V_{s30}) + 5.556 \times 10^{-4} \cdot precip - 0.0287 \cdot (d_c)^{0.5} + 0.0666 \cdot d_r - 0.0369 \cdot d_r \cdot (d_c)^{0.5}$
GGM2 (Inland)	$8.801 + 0.334 \cdot \ln(PGV) - 1.918 \cdot \ln(V_{s30}) + 5.408 \times 10^{-4} \cdot precip - 0.2054 \cdot d_w - 0.0333 \cdot wtd$

Performance-Evaluation Methodology

Standardized and objective methods are needed to analyze model efficacy (i.e., the ability to predict whether sites have liquefaction manifestations). Receiver-operating-characteristic (ROC) analyses, which are widely used in biostatistics and medical diagnostics (e.g., Fawcett, 2006; Zou, 2007) are adopted for this purpose. In any ROC analysis, the distributions of "positives" (e.g., manifestations are observed) and "negatives" (e.g., no manifestations are observed) overlap when the frequencies of the distributions are expressed as a function of a diagnostic test index (e.g., *LPI*, *LSN*, etc.). To demonstrate, two such distributions are shown in Figure 1a, plotted as a function of *LPI*. ROC curves plot the rates of true-positive predictions (R_{TP}) (i.e., manifestations are observed as predicted) and false-positive predictions (R_{FP}) (i.e., manifestations are predicted but not observed) as a function of classification "thresholds." These thresholds are used to predict outcomes, such that index values above and below a threshold respectively predict positives and negatives. Figure 1b illustrates the relationship among the positive and negative distributions, the threshold values, and the ROC curve. Low thresholds result in a high R_{FP}, while high thresholds

result in a low R_{TP}, equivalent to a high rate of false-negative predictions (R_{FN}), where R_{FN} = 1 - R_{TP}. In general, neither of these situations is desirable.

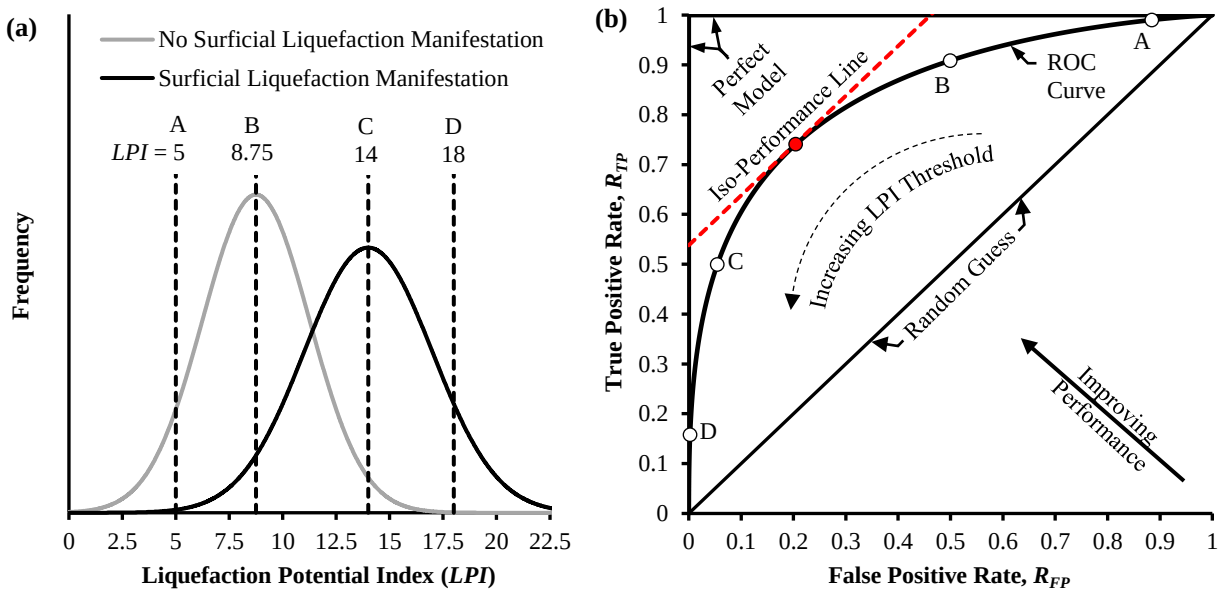


Figure 1. ROC analyses: (a) frequency distributions of liquefaction manifestation and no liquefaction manifestation as a function of *LPI*; (b) corresponding ROC curve, and illustration of how a ROC curve is used to assess the efficiency of a diagnostic test.

The optimal threshold may be defined as that which minimizes the prediction cost:

$$Cost = C_{FP} \text{ x } R_{FP} + C_{FN} \text{ x } R_{FN} \quad (4)$$

where C_{FP} and R_{FP} are respectively the cost and rate of false-positive predictions and C_{FN} and R_{FN} are respectively the cost and rate of false-negative predictions. Examples of false-positive costs include superfluous spending on design and construction (e.g., ground improvement costs), while false-negative costs are those resulting from liquefaction (e.g., property damage and lost productivity, among others). Normalizing by C_{FN}, Eq. 4 is alternatively expressed as:

$$\text{Cost'} = Cost/C_{FN} = R_{FP} \text{ x } CR + R_{FN} \quad (5)$$

where $CR = C_{FP}/C_{FN}$ and is the "cost ratio" representing different misprediction economies. Low-*CR* scenarios with are those where the costs of liquefaction far outweigh the costs of mitigation (e.g., a critical facility), while high-*CR* scenarios are those where the costs of mitigation far outweigh the costs of liquefaction (e.g., a car park). It follows from Eqs. 4-5 that two points in ROC space, (R_{FP1}, R_{TP1}) and (R_{FP2}, R_{TP2}), have equivalent performance if:

$$\frac{R_{TP1} - R_{TP2}}{R_{FP1} - R_{FP2}} = \frac{C_{FP}}{C_{FN}} = CR = m \quad (6)$$

Eq. (6) defines the slope, *m*, of an iso-performance line, such that all points defining the contour have equal *Cost'*. Thus, each *CR* corresponds to a unique contour in ROC space. One such line is shown in Figure 1b. With 1:1 slope, it corresponds to the case where false positives and false negatives have equal costs. Points tangent to this line on the ROC curve correspond to threshold values at which *Cost'* is minimized.

To evaluate model efficacy, two different ROC-based methods will be used. The first will quantify comprehensive performance via the area under a ROC curve (*AUC*). While no single parameter fully characterizes performance, *AUC* is commonly used for this purpose due to its statistical significance and objectivity (e.g., Fawcett, 2006). Here, *AUC* is the probability that sites with manifestations have higher computed index values (e.g., *LPI*) than sites without manifestations. Better prediction models thus have higher *AUC*. As shown in Figure 1b, random guessing is depicted in ROC space by a 1:1 line through the origin, for which *AUC* = 0.5. A perfect model, for which *AUC* = 1.0, plots as a point at (0,1), indicating the existence of a threshold value that perfectly separates the two distributions. To account for finite-sampling of case histories, bootstrap simulations will be performed to quantify ROC uncertainty, and in turn, to compute

confidence intervals on each model's *AUC*. This will illustrate the sensitivity of model performance to the case-histories compiled for analysis. To assess whether differences in *AUC* could arise from chance (i.e., due to finite sampling), tests of statistical significance will be performed per the method of DeLong et al. (1988), which is specific to ROC analyses. All models will be compared to one another to determine which, if any, are statistically better.

While *AUC* is widely used, it reflects overall performance across all misprediction economies. As a result, the model with highest *AUC,* the model with lowest misprediction rate ($R_{FN} + R_{FP}$), and the model most optimal when $C_{FN} \neq C_{FP}$ could conceivably all be different. This is shown in Figure 2a, where models A and B have identical *AUC*. If $CR = 1/5$, iso-performance lines have slope of 1/5 and define points with equal prediction cost. What that cost *is* depends on the lines' R_{TP}-intercepts and is given by Eq. 4 (the greater the R_{TP}-intercept, the lower the cost). Since an iso-performance line tangent to curve B has a greater R_{TP}-intercept than one tangent to curve A, model B is more optimal. That is, B is better in the "conservative" region, where models correctly classify most positives, but at the expense of high R_{FP}. Conversely, and by the same logic, A is optimal when $CR = 5$ and is better in the "liberal" region, where models correctly classify most negatives, but at the expense of low R_{TP}. Lastly, when $CR = 1$, A and B perform equally well. By similar logic, a model with higher *AUC* could be less efficient in a specific region of ROC space than a model with lower *AUC*. This is shown in Figure 2b: model A has higher *AUC* and is optimal when $CR > 0.27$, but model B is optimal for all other *CR*. Thus, *AUC* reflects overall performance, but not the nuances described above. Accordingly, a second ROC-based analysis will identify the optimal model as a function of *CR*. *Cost'* will be computed by Eq. 5 for each model over the domain $0 < CR < 2$; that with lowest *Cost'* at a given *CR* is most optimal, which is equivalent to

the graphical analysis shown in Figure 2. Together, these methods will assess model performance, both in a comprehensive sense, and for various misprediction economies of interest.

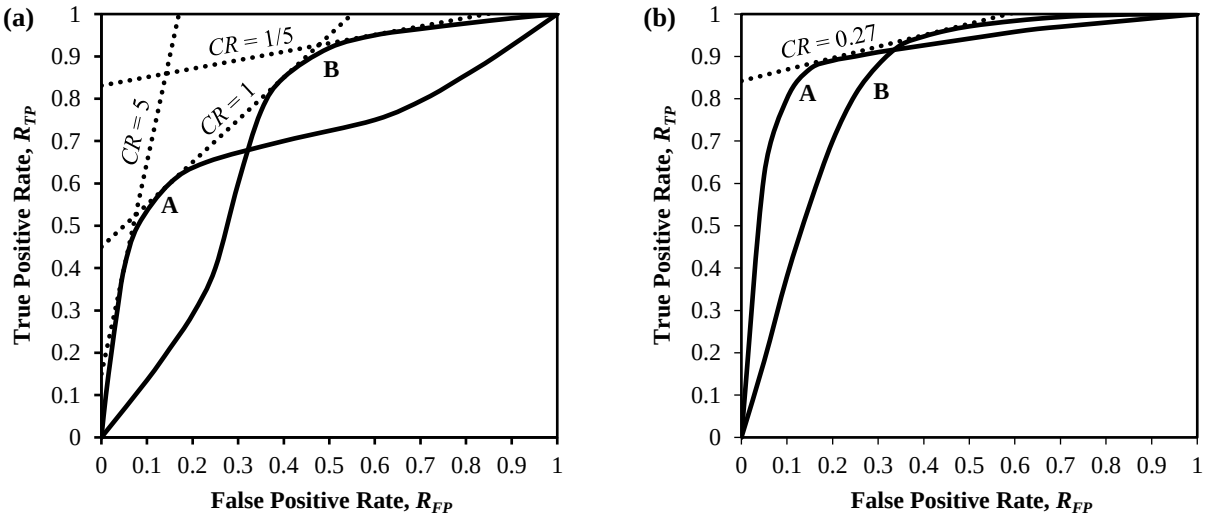


Figure 2. ROC analyses demonstrating that: (a) classifiers with equivalent *AUC* (i.e., equal overall efficiency) can perform very differently in specific regions of ROC space; (b) classifiers with higher *AUC* can, in specific regions of ROC space, perform worse than classifiers with lower *AUC*.

Results and Discussion

Utilizing the data and methodology above, 23 models were used to predict liquefaction manifestations for 15,223 case histories. To illustrate how ROC analyses and bootstrap simulations will be used to study model-performance, results for one model are first presented in detail; summary statistics from analyses of all 23 models will then be provided. In Figures 3a and 3b, ROC analyses of the BI14 – *LPI* geotechnical model (i.e., the BI14 triggering model used with the *LPI* manifestation model) are presented for the Canterbury and global datasets, respectively. In each case, a total of 10,000 bootstrap simulations were performed, from which 95%-confidence intervals (CIs) were computed. The 50th percentile ROC curve is equivalent to that resulting from

an analysis of all case-histories without resampling. Two observations are made from this figure: (1) the BI14-*LPI* model performs better on the Canterbury dataset than the global dataset, with respective median *AUC*s of 0.83 versus 0.77; and (2) finite-sample uncertainty is considerably larger for the global dataset. The 95% CI on *AUC* is 0.828 to 0.841 for the Canterbury dataset, and 0.709 to 0.826 for the global dataset. Each of these observations will be discussed further in the context all models.

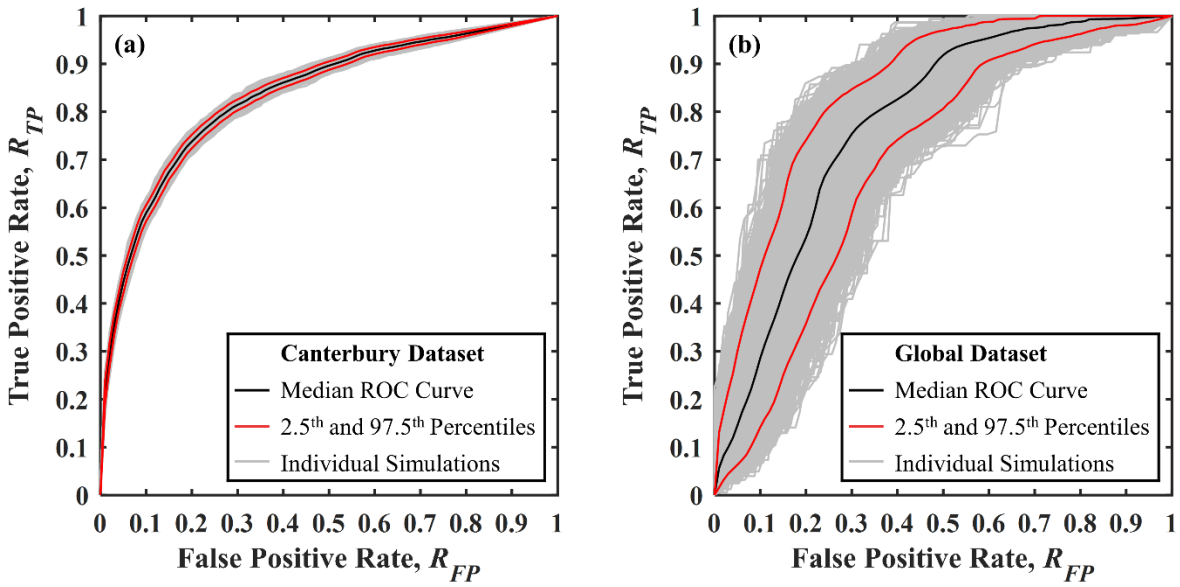


Figure 3. ROC analysis of BI14-*LPI* model performance in predicting liquefaction surface manifestation for the: (a) Canterbury dataset; and (b) global dataset.

From replicate ROC analyses of all 23 prediction models, *AUC* summary statistics are compiled and presented in Figures 4a and 4b for the Canterbury and global datasets, respectively. Shown are each model's median *AUC* and 95% CI, with results ordered by the year in which each model was proposed. In this regard, geotechnical models are dated in accordance with the triggering model but are grouped/symbolized by manifestation model. Thus, Figure 4 allows for assessments

to be made regarding the evolution of performance (i.e., prediction efficiency) through twenty years of model research and development.

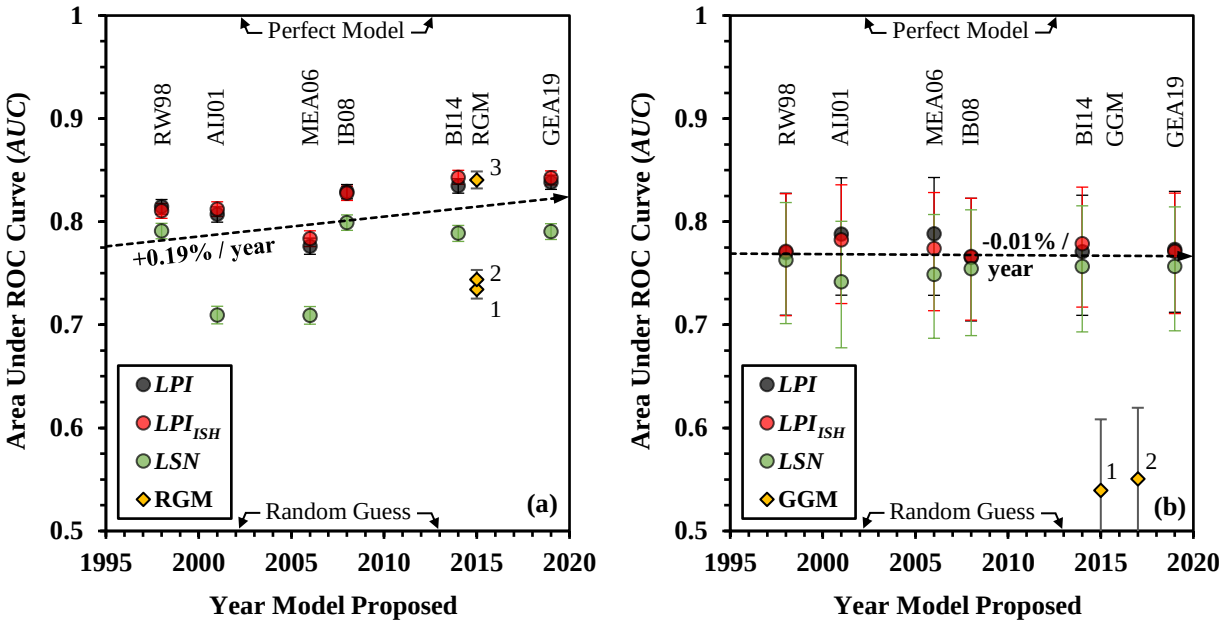


Figure 4. Summary of liquefaction-model performance – quantified by *AUC* – for 23 models, ordered by year proposed: (a) Canterbury dataset; and (b) global dataset. Markers denote median *AUC*; bars are 95% confidence intervals on *AUC*; all model acronyms are identified in Table 1. Trendlines are developed from linear regression and do not include RGM or GGM data points.

With respect to Figure 4, several observations are made as follows. *First*, the two trends identified in Figure 3 for BI14-*LPI* are true of most geotechnical models. That is, their performance on the Canterbury dataset tends to be marginally better and much less uncertain, relative to the global dataset. Considering all 18 geotechnical models, the average *AUC*s are respectively 0.80 and 0.77 for the Canterbury and global datasets, indicating also that the models are nearer to perfection (*AUC* = 1.0) than to random guessing (*AUC* = 0.5). The average 95% CIs are respectively 0.015 and 0.118 for the Canterbury and global datasets. These differences may be

attributable to the global dataset's greater seismologic, geologic, and geomorphic diversity and/or because the global field-data (e.g., ground-motion IMs, CPTs) were collected over many decades using different means and methods. There are also far fewer global case-histories; all else being equal, greater finite-sample uncertainty is thus expected. *Second*, as shown in Figure 4, trendlines were fit to the *AUC* values for all geotechnical models; these regressions did not include the regional or global geospatial models, which will be discussed momentarily. Through 20-years of model research and development, the trendlines suggest improvements to *AUC* of 0.19% per year and -0.01% per year for the Canterbury and global datasets, respectively. When considering these trends, it should be noted that: (i) the BI14 and Gea19 triggering models, when developed, were trained on a case-history dataset in which 20% of data was from Canterbury, so their evaluation in Figure 4a is not completely unbiased, unlike the other models; and (ii) some of the global case-histories compiled to test performance in Figure 4b were also used to train triggering models during their respective developments. Specifically, of the 275 global test cases, the portion used in training ranges from 0% (RW98) to 75% (e.g., BI14). Nevertheless, successive models proposed over the last twenty years show little or no demonstrable improvement for the two datasets analyzed, as shown in Figure 4. This is despite: (i) ever-increasing data that can be used to train and validate models; and (ii) greater knowledge of liquefaction mechanics, embodied by new and revised model components (e.g., magnitude scaling; depth-stress reduction; overburden correction). Of course, greater performance-variation might result if the models were tested on data outside the parameter space of that used to train them (e.g., liquefaction-susceptible soils with atypical density, depth, fines-content, age, etc.). However, liquefaction case-histories tend to have much in common, so such data is not easily obtained.

Third, it can be seen in Figure 4a that the three regional geospatial models perform remarkably well for the Canterbury dataset, with average *AUC* of 0.77 (versus 0.80 for all geotechnical models). One geospatial model – RGM3 – outperforms 16 of the 18 geotechnical models, a provocative result given the relative costs and complexities of the required model inputs. While differences in model specificity must be acknowledged (RGM3 is region-specific, while the geotechnical models are not), it is nonetheless surprising that an empirical model based on surface parameters could outperform a semi-mechanistic model based on subsurface measurements. However, as shown in Figure 4b, geospatial models do not perform as well on the global dataset, with global models GGM1 and GGM2 having median *AUC*s of 0.54 and 0.55, respectively (the 95% CIs are 0.47 to 0.61 and 0.48 to 0.62, respectively). Thus, while the geospatial models *do* tend to be useful (e.g., performing better than random guessing for the compiled test cases), the geotechnical models are significantly more efficient. This is unsurprising when considering: (i) the geomorphic, topographic, and climatic diversity of the global dataset; and (ii) the difficulty, given this diversity, of accurately inferring below-ground conditions from above-ground parameters. Inherently, the geotechnical models - being based on direct measurements of the subsurface - should have better portability across environs. Nonetheless, the results from Canterbury in Figure 4a demonstrate the provocative potential of geospatial modeling. As more test data becomes available, and as better geospatial predictors are identified, improved global performance should result. In contrast, considering the 20-year trendlines in Figure 4, it appears less likely that geotechnical models will soon improve greatly, should the status quo continue. Neither fine-tuning of model parameters, nor incremental grown in training data, are likely to inflect these trends upwards. Arguably, this would occur only with disruptive innovation (e.g., to the in-situ characterization method, or to the fundamental modeling approach). *Fourth*, the top-performing

models, as quantified via median *AUC*, are: (i) Canterbury dataset: BI14-LPI_{ISH} (*AUC* = 0.843), Gea19-LPI_{ISH} (*AUC* = 0.843), and RGM3 (*AUC* = 0.841); and (ii) global dataset: Mea06-*LPI* (*AUC* = 0.788), AIJ01-*LPI* (*AUC* = 0.788), and AIJ01- LPI_{ISH} (*AUC* = 0.782). Thus, the top-performing models appear to be different for the two datasets. However, given the uncertainties of *AUC* values – particularly for the global dataset – it should be determined whether the measured differences in performance are statistically significant.

Using the method of DeLong et al. (1988), P-values were computed to compare each of the 23 models to all others. These values are presented in Tables 3 and 4 for the Canterbury and global datasets, respectively, and are the probabilities that *AUC* samples for two models could have come from the same distribution (i.e., that the observed difference in performance arose from chance and not because one model is better). Large P-values can be expected when (i) differences between two *AUC* values are small; and (ii) the uncertainty of one or more of the *AUC* values is large. The popular significance level of 0.05 of adopted herein, such that P-values below 0.05 denote that models have significantly different performance. Using this criterion, Tables 4 and 5 compare all model pairs and identify which model, if any, is statistically better.

As seen in Table 4, instances of large P-values are rare for the Canterbury dataset, which can be partly attributed to the large number of case histories. In fact, nearly every model performs statistically differently than all others. Notable observations from Table 4 are: (i) among all 23 models evaluated, no model is statistically better than all others; but (ii) three models are either statistically better, or not statistically different, than all others. In other words, three models are never bested. These are BI14-LPI_{ISH}, Gea19-LPI_{ISH}, and RGM3. Most striking is that no geotechnical model is statistically better than geospatial model RGM3.

Table 4. P-Value Matrix to Compare Model Performance for the Canterbury Dataset

Statistically Better*	↑	LPI						LPI_{ISH}						LSN						RGM		
←		RW98	AIJ01	Mea06	IB08	BI14	Gea19	RW98	AIJ01	Mea06	IB08	BI14	Gea19	RW98	AIJ01	Mea06	IB08	BI14	Gea19	1	2	3
LPI	RW98		0.000	0.000	0.000	0.000	0.000	0.000	0.018	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	AIJ01			0.000	0.000	0.000	0.000	0.177	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Mea06				0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	IB08					0.000	0.000	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
	BI14						0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.082
	Gea19							0.000	0.000	0.000	0.000	0.004	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.362
LPI_{ISH}	RW98								0.530	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	AIJ01									0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Mea06										0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.002	0.000	0.000	0.000
	IB08											0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	BI14												0.271	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.943
	Gea19													0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.823
LSN	RW98														0.000	0.000	0.000	0.043	0.399	0.000	0.000	0.000
	AIJ01															0.760	0.000	0.000	0.000	0.000	0.000	0.000
	Mea06																0.000	0.000	0.000	0.000	0.000	0.000
	IB08																	0.000	0.000	0.000	0.000	0.000
	BI14																		0.000	0.000	0.000	0.000
	Gea19																			0.000	0.000	0.000
RGM	1																				0.026	0.000
	2																					0.000
	3																					

*Cell values are the probabilities that *AUC* samples for two models could have come from the same distribution. Values less than 0.05 are deemed "significant", in which case the model with significantly better performance is indicated via the cell shading.

In contrast to the Canterbury dataset, few P-values are below 0.05 for the global dataset, indicating that performance-differences between models are typically not statistically significant. Notable observations from Table 5 are: (i) all geotechnical models are statistically better than both global geospatial models; and (ii) no model is statistically superior. The previously identified top-performing model globally – Mea06-*LPI* – is statistically better than nine other models, but is statistically indifferent from eight others, including the top-performing models in Canterbury (BI14-LPI_{ISH} and Gea19-LPI_{ISH}). Thus, once statistical significance is considered, the global

results are largely inconclusive due to the large finite-sample uncertainties of *AUC* values. Consequently, the top-performing geotechnical models in Canterbury may also be the top-performing models globally, but more global data would be needed to confirm this, or to draw other conclusions about model performance.

Table 5. P-Value Matrix to Compare Model Performance for the Global Dataset

Statistically Better*	↑	LPI						LPI_{ISH}						LSN						GGM	
←		RW98	AIJ01	Mea06	IB08	BI14	Gea19	RW98	AIJ01	Mea06	IB08	BI14	Gea19	RW98	AIJ01	Mea06	IB08	BI14	Gea19	1	2
LPI	RW98		0.166	0.046	0.446	0.933	0.682	0.898	0.467	0.759	0.696	0.435	0.894	0.571	0.224	0.263	0.244	0.377	0.369	0.000	0.000
	AIJ01			0.978	0.105	0.211	0.255	0.245	0.439	0.368	0.137	0.500	0.248	0.092	0.004	0.031	0.032	0.051	0.052	0.000	0.000
	Mea06				0.040	0.135	0.162	0.120	0.657	0.135	0.066	0.441	0.186	0.026	0.020	0.004	0.009	0.021	0.021	0.000	0.000
	IB08					0.355	0.101	0.582	0.338	0.565	0.887	0.144	0.521	0.857	0.346	0.420	0.400	0.569	0.555	0.000	0.000
	BI14						0.358	0.975	0.525	0.824	0.678	0.379	0.944	0.596	0.259	0.290	0.231	0.328	0.329	0.000	0.000
	Gea19							0.870	0.598	0.931	0.466	0.511	0.855	0.461	0.205	0.220	0.140	0.224	0.224	0.000	0.000
LPI_{ISH}	RW98								0.458	0.759	0.490	0.472	0.973	0.544	0.181	0.240	0.245	0.380	0.367	0.000	0.000
	AIJ01									0.582	0.301	0.815	0.475	0.247	0.005	0.074	0.113	0.163	0.157	0.000	0.000
	Mea06										0.527	0.787	0.819	0.361	0.077	0.070	0.165	0.250	0.240	0.000	0.000
	IB08											0.053	0.353	0.790	0.285	0.368	0.354	0.521	0.504	0.000	0.000
	BI14												0.076	0.319	0.134	0.154	0.097	0.158	0.154	0.000	0.000
	Gea19													0.557	0.206	0.256	0.215	0.325	0.313	0.000	0.000
LSN	RW98														0.258	0.154	0.142	0.362	0.321	0.000	0.000
	AIJ01															0.657	0.551	0.461	0.470	0.000	0.000
	Mea06																0.678	0.493	0.510	0.000	0.000
	IB08																	0.451	0.498	0.000	0.000
	BI14																		0.792	0.000	0.000
	Gea19																			0.000	0.000
GGM	1																				0.779
	2																				

*Cell values are the probabilities that AUC samples for two models could have come from the same distribution. Values less than 0.05 are deemed "significant", in which case the model with significantly better performance is indicated via the cell shading.

As discussed, *AUC* is widely used to study model behavior, but it reflects overall performance across all misprediction economies. To identify the most efficient model for particular economies, *Cost'* is computed for each model at different *CR* values. The model with lowest *Cost'* is identified

in Figure 5 for *CR*s ranging from 0.01 to 2; results are presented for the Canterbury and global datasets in Figures 5a and 5c, respectively. Also, because multiple models could have nearly equivalent *Cost'*, and thus, be equally optimal, any model whose *Cost'* is within 1% of minimum is identified as "optimal." To assess the sensitivity of results to this criterion, analyses are repeated by relaxing the threshold to 10%, as shown in Figures 5b and 5d.

While the models with highest *AUC* (see Figure 4) tend to appear in Figure 5 (i.e., they are optimal for some range of *CR*), the results convey more nuanced information than *AUC* alone. For example, as seen in Figure 5a for the Canterbury dataset, geospatial model *RGM3* is optimal in both the conservative, or "northeast", region of ROC space (low *CR*), and in the liberal, or "southwest", portion of ROC space (high *CR*). Thus, for scenarios in which either false negatives or false positives are relatively inconsequential, model *RGM3* outperforms all others. While this may sound abstract, it is exactly the misprediction economy for critical facilities (e.g., powerplants, hospitals, etc.) in which C_{FN} tends to far exceed C_{FP}, making avoidance of the former paramount.

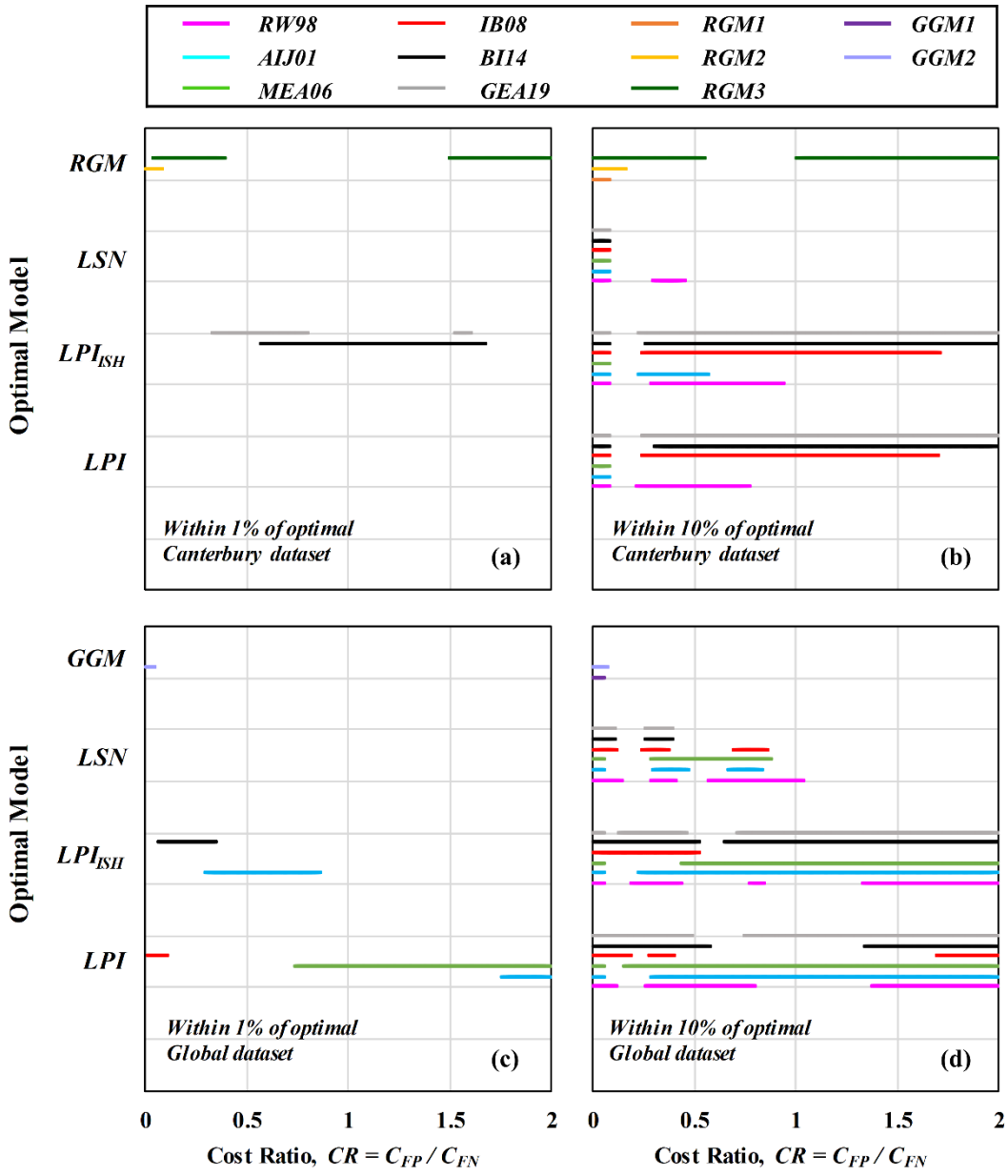


Figure 5. [color] Optimal liquefaction model as a function of *CR*, as determined from ROC analyses of the Canterbury dataset, where "optimal" models are those within (a) 1% and (b) 10% of optimal; analogous analyses are presented in panels (c) and (d) for the global dataset.

Conversely, geotechnical models BI14-LPI_{ISH} and Gea19-LPI_{ISH} are optimal in the range of $0.25 < CR < 1.75$, or scenarios in which C_{FN} and C_{FP} are relatively more similar. It can also be

seen that models with lower *AUC* are sometimes optimal in specific regions of ROC space. For example, most striking is that geospatial model *GGM2* is the most optimal model for the global dataset at low *CR*. This is despite the fact that its *AUC*, or overall prediction efficiency, was significantly lower than all geotechnical models. Extending the threshold for identifying "optimal" models to 10%, it can be seen that each of the 23 models evaluated is "optimal" for some range of *CR*. Thus, Figure 5 provides additional performance insights and demonstrates how misprediction economies may be considered when assessing or selecting prediction models.

Lessons For Geotechnical and Geospatial Models

Informed by the preceding analyses, four lessons for improving model performance are next highlighted. First, while obvious, it is worth contextualizing the importance of inter-event model stability. Given imperfect models "A" and "B", A may outperform B on all events individually, but B may outperform A when the data from all events is combined. This is true of models tested here (e.g., BI14-LPI vs. RGM3). For the 2010 Darfield and 2011 Christchurch earthquakes, BI14-LPI has event-specific AUCs of 0.61 and 0.73, respectively. The threshold LPI values that minimize mispredictions in these events are 2.82 and 8.67, respectively. In contrast, the RGM3 event specific AUCs and threshold values for these respective earthquakes are 0.53 and 0.68, and 11.7 and 11.8. While BI14-LPI has better AUC in each individual event (i.e., better segregating sites with and without manifestations), the inter-event instability of its scaling, or the LPI value at which segregation is maximum, leads to it performing no better than RGM3 when the events are analyzed together. The relative success of geospatial models is due at least partly to this phenomenon. Thus, as demonstrated by BI14-LPI, identifying and addressing factors which cause inter-event instability is central to improving model performance.

Second, geospatial models have a difficult task – to accurately infer below-ground conditions from above-ground parameters. In this respect, the inability to predict soil type (and by corollary, liquefaction susceptibility) was a common cause of mispredictions for the models and cases tested here. As seen in Table 3, these models generally expect ground to be susceptible to liquefaction if it is flat, saturated, and near rivers, streams, or coasts. However, as identified from global testing, such sites may consist predominately of soils less- or un-susceptible to liquefaction (e.g., clays, peats, or gravels). Moreover, we find in many of these cases that surface-geology maps accurately predict the predominance of such soils. Geospatial models could thus benefit from including mapped geologic data, when available, as well as other, yet unidentified proxies of soil type. In addition, the coarse resolution of some geospatial inputs (e.g., mapped rivers) hinders performance at the site-specific scale. Being central to their performance, more efficient and higher resolution predictors of subsurface properties must be identified for geospatial models to improve.

Third, geospatial models inherently merge liquefaction triggering and manifestation. This calls attention to the "simplified" framework for developing geotechnical triggering models (Seed & Idriss, 1971; Whitman, 1971) – specifically to the interdependent relationship between triggering and manifestation. While the six triggering models tested herein have several post-1971 improvements (e.g., new or modified terms that account for influential factors), the fundamental approach to developing these models has never changed. Using this approach, an in-situ test is performed where liquefaction was observed to have triggered or inferred to have not triggered. An analyst interprets the test data and ties the observed response of the profile to a "critical" stratum. For "no" cases - where no manifestation is observed - the "critical" stratum is the most susceptible stratum which, if liquefied, is conjectured to manifest at the surface. This is repeated for many case histories, and for each, the cyclic-stress demand in the critical stratum is plotted vs. a measured

proxy of soil density (e.g., CPT resistance). Lastly, a "triggering curve" is regressed to define the relationship between the cyclic demand imposed on a soil and its capacity to resist liquefaction. Within this process, the use of surface observations is problematic, since such observations are a function both of factors influencing triggering and factors influencing manifestation. To select the critical stratum, the analyst must therefore use a manifestation model in reverse (i.e., a model relating surface observations to triggering at-depth). This selection must be made such that the critical stratum's thickness, depth, density, fines content, plasticity, and strain-potential, considering also all properties of all overlying strata, is consistent with the surface observation. If this is not achieved, then embedded in the derivative triggering curve will be factors which relate not only to liquefaction triggering, but also to the post-triggering manifestation of liquefaction. Problematic, then, is that the manifestation model used by all analysts to-date when developing triggering models has been personal judgement, rather than a defined analytical model. Meanwhile, and predicated on the accuracy and "purity" of triggering models, a different set of researchers has separately proposed manifestation models that predict surface expressions from triggering (e.g., LPI, LSN) by considering the cumulative response of a profile. By inputting the results from triggering analysis (i.e., FSliq vs depth) into manifestation models, the latter may be trained using case-history data (note that such data was also used to train triggering models, albeit using different, undefined manifestation models). Given that: (i) an effective manifestation model is required in order to formulate an effective triggering model, and (ii) an effective triggering model is required in order to formulate an effective manifestation model, it follows that (iii) these models should be developed harmoniously and iteratively within a consistent framework, just as any inverse problem requires. Failure to do so potentially leads to the omission, double counting, and general obscuration of the mechanics controlling triggering and manifestation. While the effect

on performance is unknown, the development of geotechnical triggering and manifestations models has historically been, and presently is, less than completely rational.

Fourth, within the broader shortcoming just described, the multi-scale interactions which influence surface manifestation must be understood and modeled more completely. For example, it is known that low-permeability soils within or atop a profile can influence the morphology of manifestations (e.g., by effecting pore-pressure development and transmission). This has been documented in lab, numerical, and field research (e.g., Fiegel and Kutter, 1994; Ozutsumi et al., 2002; Brennan and Madabhushi, 2005; Juang et al., 2005; Özener, et al., 2008; Maurer et al., 2015b; Cubrinovski et al., 2019). Despite this, only one of the manifestation models tested, LPIISH, considers the influence of soils expected not to liquefy, and it does so in a way that is insufficient for capturing their total influence. Other factors likely pertinent to surface manifestation (e.g., strata thickness, depth, sequencing, and permeability, among others) are either modeled in a generally heuristic manner or are not considered by LPI, LSN, and LPIISH. Accordingly, the performance of geotechnical models would surely benefit from a more rigorous understanding of the mechanics and interactions which drive a soil profile's cumulative response.

Summary and Conclusions

Using 15,223 case-histories from 24 earthquakes, this study evaluated the performance of 23 liquefaction models based on geotechnical or geospatial data. Case histories were parsed into the "Canterbury" and "global" datasets and finite-sample uncertainty was used to test for significance. The most salient findings are as follows. (i) For the Canterbury dataset, three models were either statistically better, or not statistically different, than all others: BI14-LPI_{ISH}, Gea19-LPI_{ISH}, and RGM3. (ii) For the global dataset, the top-performing models were Mea06-*LPI* and AIJ01-*LPI*,

yet finite-sample uncertainties yield these results inconclusive. The top-performing models in Canterbury may also be the top-performing models globally, but more global case history data is needed to confirm this, or to draw other conclusions about performance. (iii) Successive "simplified" geotechnical models proposed over the last 20-years show little or no demonstrable improvement. It is unlikely that fine-tuning of model parameters or incremental grown in training data will inflect this trend upwards. More likely, this would occur only with disruptive innovation to the in-situ test method or fundamental modeling approach. (iv) Geospatial models performed remarkably well for the Canterbury dataset, with one such model performing significantly better than most geotechnical models – a provocative result given the relative costs of model inputs. (v) However, for the global dataset, all geotechnical models performed significantly better than all geospatial models, with the latter performing only marginally better than random guessing. This highlights the difficulty of inferring below-ground conditions from above-ground parameters, especially across diverse geomorphic, topographic, and climatic environs. (vi) Informed by the presented analyses, four paths to improving the performance of geotechnical or geospatial models were highlighted. These include: (a) identifying and addressing sources of inter-event instability; (b) identifying better geospatial proxies of soil type, and by corollary, liquefaction susceptibility; (c) developing geotechnical triggering and manifestation models within a more rational and internally consistent framework; and (d) more rigorous investigation and modelling of the multi-scale interactions which drive a soil profile's cumulative response.

The findings of this study are tied to the data analyzed, which in effect is the present sum of CPT case histories. The applicability of these findings to other case-history data – particularly that with different parameter space (e.g., soils with unusual minerology, composition, age, etc.; earthquakes of large or small magnitude) – or to other models and methodologies, is unknown. In

addition, the presented findings should be considered in the context of model regionality and potential bias. Ultimately, additional data will confirm or update the conclusions drawn above.

Acknowledgements

The presented study is based upon work supported by the National Science Foundation (NSF), US Geological Survey (USGS), and Pacific Earthquake Engineering Research Center (PEER) under Grant Nos. CMMI-1751216, G18AP-00006, and 1132-NCTRBM, respectively. The authors also wish to acknowledge the numerous researchers who contributed to the liquefaction case-history data studied herein. Much of this data was collected under the sponsorship of either the New Zealand Earthquake Commission (EQC) or one of the three aforementioned organizations. 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, USGS, PEER, or EQC.

Data

The curated Canterbury dataset is available in digital format from the authors upon request. The global dataset was compiled from the following sources, parsed by event, and may soon be available in full or part from the Next-Generation Liquefaction project (e.g., Zimmaro et al., 2019): **1964 M_w7.6 Niigata, JPN** - Ishihara & Koga (1981), Farrar (1990), Moss et al. (2003); **1971 M_w6.6 San Fernando, USA** - Bennett et al., 1998, Toprak and Holzer (2003); **1975 M_w7.0 Haicheng, CHN** - Arulandan et al. (1986), Shengcong & Tatsuoka (1984); **1976 M_w7.6 Tangshan, CHN** - Shibata & Teparaska (1988), Moss et al. (2009; 2011); **1979 M_w6.53 Imperial Valley, USA** - Diaz-Rodriguez (1984), Diaz-Rodriguez and Armijo-Palaio (1991), Moss et al. (2003); **1981 M_w5.9 Westmoreland, USA** - Bennett et al. (1984), Seed et al. (1984), Cetin et al. (2000), Moss et al. (2005); **1983 M_w7.7 Nihonkai-Chubu, JPN** - Farrar (1990); **1983 M_w6.88**

Borah Peak, USA - Andrus (1986), Andrus & Youd (1987), Moss et al. (2003); **1987 M_w6.6 Edgecumbe, NZ** - Christensen (1995), Moss et al. (2003); **1987 M_w6.54 Superstition Hills, USA** - Bennett et al. (1984), Cetin et al. (2000), Toprak & Holzer (2003), Moss et al. (2005), Holzer & Youd (2007); **1989 M_w6.93 Loma Prieta, USA** - Mitchell et al. (1994), Pass (1994), Bennett & Tinsely (1995), Boulanger et al. (1995; 1997), Kayen et al. (1998), Toprak & Holzer (2003), Youd & Carter (2005); **1994 M_w6.69 Northridge, USA** - Abdel-Haq & Hryciw (1998), Bennett et al., 1998, Holzer et al. (1999), Moss et al. (2003); **1995 M_w6.9 Hyogoken-Nambu, JPN** - Suzuki et al. (2003); **1999 M_w7.51 Kocaeli, TUR** - PEER (2000a), Youd et al. (2009); **1999 M_w7.62 Chi-Chi, TWN** - Lee et al. (2000), PEER (2000b); **2008 M_w6.4 Achaia-Ilia, GRC** - Batilas et al. (2014); **2008 M_w7.2 El Mayor-Cucapah, MEX** - Moss et al. (2005); CESMD (2016), Turner et al. (2016); **2011 M_w9 Tohoku, JPN** - Cox et al. (2013), Boulanger & Idriss (2014); **2012 M_w6.1 Emilia, ITA** - Papathanassiou et al. (2015), Facciorusso et al. (2015), Servizio Geologico (2016).

References

Abdel-Haq, A., and Hryciw, R. D., 1998. Ground settlement in Simi Valley following the Northridge earthquake. *JGGE* **124**(1), 80-89.

Allstadt, K.E., Thompson, E.M., Hearne, M., Wald, D.J., Fee, J.M., Martinez, E.M., Hunter, E., and Brown, J.D., 2019. The new USGS near-real-time ground failure product and its performance for recent earthquakes. *Seismological Society of America Annual Meeting*, Seattle, WA.

Andrus, R.D., 1986. Subsurface investigations on a liquefaction induced lateral spread, Thousand Springs Valley, Idaho: liquefaction recurrence and a case history in gravel. M.S. Thesis, Dept. of Civil Engineering, Brigham Young University, Provo, Utah.

Andrus, R. D. and Youd, T. L. 1987. Subsurface investigation of a liquefaction-induced lateral spread Thousand Springs Valley, Idaho. Misc. paper GL-87-8, U.S. Army Corps of Engineers.

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

Arulanandan, K., Yogachandran, C., Meegoda, N.J., Ying, L., and Zhauji, S., 1986. Comparison of the SPT, CPT, SV and electrical methods of evaluating earthquake induced liquefaction susceptibility in Ying Kou City during the Haicheng Earthquake. *Geotechnical Special Publication 6*, 389–415, ASCE.

Batilas, A., Pelekis, P., Vlachakis, V., and Athanasopoulos, G., 2014. Soil Liquefaction/ Nonliquefaction in the Achaia-Ilia (Greece) 2008 Earthquake: field evidence, site characterization and ground motion assessment. *International Journal of Geoengineering Case histories* **2**(4), 270-287.

Beven, K.J. and Kirkby, M.J. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* 24, 43–69.

Bennett, M.J., McLaughlin, P.V., Sarmiento, J.S., and Youd, T. L., 1984. Geotechnical investigation of liquefaction sites, Imperial Valley, CA. U.S. Geological Survey, Open-File Report 84-252.

Bennett, M.J., Ponti, D.J., Tinsley, J.C., III, Holzer, T.L., and Conaway, C.H., 1998. Subsurface geotechnical investigations near sites of ground deformations caused by the January 17, 1994, Northridge, California, earthquake. U.S. Geological Survey, Open-File Report 98-373.

Bennett, M.J., and Tinsley, J.C. III, 1995. Geotechnical data from surface and subsurface samples outside of and within liquefaction-related ground failures caused by the October 17, 1989, Loma Prieta earthquake, Santa Cruz and Monterey Counties, CA. U.S. Geological Survey Open-File Report 95-663.

Boulanger, R.W., Idriss, I.M., and Mejia, L.H., 1995. Investigation and evaluation of liquefaction related ground displacements at Moss Landing during the 1989 Loma Prieta earthquake. Report No. UCD/CGM-95/02, Center for Geotechnical Modeling, University of California, Davis, 231 pp.

Boulanger, R.W., Mejia, L.H., and Idriss, I.M., 1997. Liquefaction at Moss Landing during Loma Prieta earthquake. *JGGE* **123**(5), 453–67.

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, University of California, Davis, CA.

Bradley, B.A., 2013. Site-specific and spatially-distributed ground motion intensity estimation in the 2010-2011 Christchurch earthquakes. *SDEE* **48**, 35-47.

Bradley, B.A., Razafindrakoto, H., Maurer, B.W., Motha, J., Tarbali, L., and Lee, R., 2018. Simulation-based ground motion prediction of historical and future New Zealand earthquakes and consequent geohazard impacts. *Geotechnical Special Publication 291*, 29-42, ASCE.

Bray, J.D. and Macedo, J., 2017. Simplified procedure for estimating liquefaction-induced building settlement, *SDEE* **102**, 215-231.

Brennan, A.J., and Madabhushi, S.P., 2005. Liquefaction and drainage in stratified soil. *JGGE* **131**(7), 876-885.

Buck, J.R., Daniel, M.M, and Singer, A.C, 2002. *Computer Explorations in Signals and Systems Using MATLAB®*, 2nd Edition. Upper Saddle River, NJ: Prentice Hall.

Bullock, Z., Karimi, Z., Dashti, S., Porter, K., Liel, A.B., and Franke, K.W., 2018. A physics-informed semi-empirical probabilistic model for the settlement of shallow-founded structures on liquefiable ground. *Géotechnique,* doi.org/10.1680/jgeot.17.P.174.

Byrne, P.M., Park, S.S., Beaty, M., Sharp, M.K., Gonzalez, L., Abdoun, T., 2004. Numerical modeling of liquefaction and comparison with centrifuge tests. *Canadian Geotechnical Journal* **41**(2), 193-211.

CERA (Canterbury Earthquake Recovery Authority), 2012. Purpose and scope of the Canterbury geotechnical database (now the New Zealand Geotechnical Database); www.nzgd.org.nz.

CESMD (Center for Engineering Strong Motion Data), 2016. U.S. structural and ground response data; www.strongmotioncenter.org/

Cetin, K.O., Seed, R.B., Moss, R.E.S., Der Kiureghian, A.K., Tokimatsu, K., Harder, L.F., and Kayen, R.E., 2000. Field Performance Case Histories for SPT-Based Evaluation of Soil Liquefaction Triggering Hazard, Geotechnical Engineering Research Report No. UCB/GT-2000/09, UC Berkeley.

Cetin, K.O., Bilge, H.T., Wu, J, Kammerer, A.M., and Seed, R.B., 2009. Probabilistic model for the assessment of cyclically induced reconsolidation (Volumetric) Settlements. *JGGE* **135**(3), 387-398.

Christensen, S.A., 1995. Liquefaction of cohesionless soils in the March 2, 1987 Edgecumbe earthquake, Bay of Plenty, New Zealand, and other earthquakes, Master of Engineering Thesis, Department of Civil Engineering, University of Canterbury, Christchurch, New Zealand.

Cox, B.R., Boulanger, R.W., Tokimatsu, K., Wood, C., Abe, A., Ashford, S., Donahue, J., Ishihara, K., Kayen, R., Katsumata, K., Kishida, T., Kokusho, T., Mason, B., Moss, R., Stewart, J., Tohyama, K., and Zekkos, D., 2013. Liquefaction at strong motion stations in Urayasu City during the 2011 Tohoku-Oki earthquake. *Earthquake Spectra* **29**(S1), S55-S80.

Cubrinovski, M. and Ishihara, K., 1998. State concept and modified elastoplasticity for sand modeling. *Soils and Foundations* **38**(4), 213-225.

Cubrinovski, M., Rhodes, A., Ntritsos, N., and van Ballegooy, S., 2019. System response of liquefiable deposits. SDEE 124, 212-229.

DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837-845.

Diaz-Rodriquez, J.A., 1984. Liquefaction in the Mexicali Valley during the earthquake of June 9, 1980. *Eighth World Conference on Earthquake Engineering*, San Francisco, 223-230.

Diaz-Rodriquez, J.A., and Armijo-Palacio, G., 1991. Liquefaction potential of fine cohesionless soils using the CPT. *Soils and Foundations* **31**(3), 111-119.

Facciorusso, J., Madiai, C., and Vannucchi, G., 2015. CPT-Based Liquefaction Case History from the 2012 Emilia Earthquake in Italy. *JGGE* **141**(12), 05015002.

Farrar, J.A., 1990. Study of in situ testing for evaluation of liquefaction resistance, R-90-06, U.S. Department of the Interior, Bureau of Reclamation, Geotechnical Services Branch, Denver Office.

Fawcett, T., 2006. An introduction to ROC analysis, *Pattern Recognition Letters* **27**(8), 861–874.

Federal Emergency Management Agency (FEMA), 2013. *Earthquake Model HAZUS-MH 2.1 Technical Manual*, Washington, D.C.

Fick, S. E., and Hijmans, R. J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**(12), 4302-4315.

Fiegel, G. L., and Kutter, B. L., 1994. Liquefaction mechanism for layered soils. *JGGE* **120**(4), 737-755.

Geyin, M., and Maurer, B.W., 2019. An analysis of liquefaction-induced free-field ground settlement using 1,000+ case-histories: observations vs. state-of-practice predictions. *Geotechnical Special Publication* 308, 489-498, ASCE.

Green, R.A., Cubrinovski, M., Cox, B., Wood, C., Wotherspoon, L., Bradley, B, and Maurer, B., 2014. Select Liquefaction Case Histories from the 2010-2011 Canterbury Earthquake Sequence. *Earthquake Spectra* **30**(1), 131-153.

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.

Holzer, T.L., Bennett, M.J., Ponti, D.J., and Tinsley, J.C., 1999. Liquefaction and soil failure during 1994 Northridge earthquake. *JGGE* **125**(6), 438-452.

Holzer, T.L., and Youd, T.L., 2007. Liquefaction, ground oscillation, and soil deformation at the wildlife array, California. *BSSA* **97**(3), 961-976.

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. *11th International Conference on Soil Mechanics and Foundation Engineering*, San Francisco, CA, USA, **1**: 321-376.

Ishihara, K., and Koga, Y., 1981. Case studies of liquefaction in the 1964 Niigata earthquake. *Soils and Foundations* **21**(3), 35-52.

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

Juang, C.H., Yang, S.H., Yuan, H., and Fang, S.Y., 2005. "Liquefaction in the Chi-Chi earthquake-effect of fines and capping non-liquefiable layers." *Soils and Foundations* **45**(6), 89-101.

Kayen, R.E., Mitchell, J.K., Seed, R.B., and Nishio, S., 1998. Soil liquefaction in the east bay during the earthquake, *The Loma Prieta, California, Earthquake of October 17, 1989 – Liquefaction*. Thomas L. Holzer, editor, USGS Professional Paper 1551-B, B61-B86.

Kayen, R.E., Moss, R.E.S., Thompson, E.M., Seed, R.B., Cetin, O.K., Der Kiureghian, A., Tanaka, Y., and Tokimatsu, T., 2013. Shear wave velocity-based probabilistic and deterministic assessment of seismic soil liquefaction potential. *JGGE* **139**(3), 407-419.

Lee, D.H., Ku, C.S., and Juang, C.H., 2000. Preliminary investigation of soil liquefaction in the 1999 Chi-Chi, Taiwan, earthquake, *International Workshop on Annual Commemoration of Chi-Chi Earthquake, Vol. III – Geotechnical Aspect*, C.H. Loh and W.I. Liao, eds., National Center for Research on Earthquake Engineering, Taipei, Taiwan, pp. 140-151.

Lehner, B., Verdin, K, and Jarvis A., 2006. *HydroSHEDS Technical Documentation*. World Wildlife Fund US, Washington, D.C

Kramer, S.L., 2008. *Evaluation of Liquefaction Hazards in Washington State, Report No. WA-RD 668.1*, Washington State Transportation Center, Seattle, Washington.

Maurer, B.W., Green, R.A., Cubrinovski, M., and Bradley, B.A., 2014. Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand. *JGGE* **140**(7), 04014032.

Maurer, B.W., Green, R.A., and Taylor, O.D.S., 2015a. 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., Bradley, B.A., 2015b. Fines-content effects on liquefaction hazard evaluation for infrastructure during the 2010-2011 Canterbury, New Zealand earthquake sequence. *SDEE* **76**, 58-68.

Maurer, B.W., van Ballegooy, S., and Bradley, B.A., 2017. Fragility functions for performance-based damage assessment of soil liquefaction, *3rd International Conference on Performance-based Design in Earthquake Geotechnical Engineering*, July 16-19; Vancouver, Canada.

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. *SDEE* **117**, 96-105.

Mitchell, J.K., Lodge, A.L., Coutinho, R.Q., Kayen, R.E., Seed, R.B., Nishio, S., and Stokoe, K. H., II 1994. Insitu test results from four Loma Prieta earthquake liquefaction sites: SPT, CPT, DMT and shear wave velocity. Report No. UCB/EERC-94/04. Earthquake Engineering Research Center, UC Berkeley.

Moss, R.E.S., Seed, R.B., Kayen, R.E., Stewart, J. P., Youd, T.L., and Tokimatsu, K. 2003. Field case histories for CPT-based in situ liquefaction potential evaluation. Geoengineering Research Rep. UCB/GE-2003/04.

Moss, R.E.S., Collins, B.D., and Whang, D.H., 2005. Retesting of liquefaction/nonliquefaction case histories in the imperial valley. *Earthquake Spectra* **21**(1), 179–196.

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

Moss, R.E.S., Kayen, R.E., Tong, L.-Y., Liu, S.-Y., Cai, G.-J., and Wu, J., 2009. Re-investigation of liquefaction and nonliquefaction case histories from the 1976 Tangshan earthquake. Rep. No.

209/102, Pacific Earthquake Engineering Research (PEER) Center, Berkeley, CA.

Moss, R.E.S., Kayen, R.E., Tong, L.-Y., Liu, S.-Y., Cai, G.-J., and Wu, J. 2011. Retesting of liquefaction and nonliquefaction case histories from the 1976 Tangshan earthquake. *JGGE* **137**(4), 334-343.

Nakata, T. and Shimazaki, K., 1997. Geo-slicer, a newly invented soil sampler, for high-resolution active fault studies. *Journal of Geography* **106**(1), 59-69 (in Japanese).

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.

Özener, P., Özaydin, K., and Berilgen, M., 2008. Numerical and Physical Modeling of Liquefaction Mechanisms in Layered Sands. *Geotechnical Earthquake Engineering and Soil Dynamics* IV, 1-12.

Ozutsumi, O., Sawada, S., Iai, S., Takeshima, Y., Sugiyama, W., & Shimazu, T., 2002. Effective stress analyses of liquefaction-induced deformation in river dikes. *SDEE* **22**(9), 1075-1082.

Papathanassiou, G., Mantovani, A., Tarabusi, G., Rapti, D., and Caputo, R., 2015. Assessment of liquefaction potential for two liquefaction prone areas considering the May 20, 2012 Emilia (Italy) earthquake. *Engineering Geology*, **189**, 1-16.

Pass, D.G., 1994. Soil characterization of the deep accelerometer site at Treasure Island, San Francisco, California. MS thesis, University of New Hampshire.

PEER (Pacific Earthquake Engineering Research Center), 2000a. Documenting incidents of ground failure resulting from the 1999 Kocaeli, Turkey Earthquake. < https://tinyurl.com/ujtpwrk>

PEER (Pacific Earthquake Engineering Research Center), 2000b. Documentation of soil conditions at liquefaction sites from 1999 Chi-Chi, Taiwan Earthquake. < https://tinyurl.com/sbh9cpg>

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

Raschke, S.A., and Hryciw, R.D., 1997. Vision cone penetrometer for direct subsurface soil observation. *JGGE* **123**(11), 1074-1076.

Robertson, P.K., and Wride, C.E., 1998. Evaluating cyclic liquefaction potential using cone penetration test. *Canadian Geotechnical Journal* **35**(3), 442-459.

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.

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

Servizio Geologico, 2016. Database of the Emilia-Romagna region: geological, seismic, and soil Survey.” http://ambiente.regione.emiliaromagna.it/geologia/cartografia/webgis-banchedati/.

Shengcong, F. and Tatsuoka, F., 1984. Soil liquefaction during Haicheng and Tangshan earthquake in China; a review. *Soils and Foundations* **24**(4), 11-29.

Shibata, T., and Teparaska, W., 1988. Evaluation of Liquefaction Potential of Soils Using Cone Penetration Testing. *Soils and Foundations* **28**(2), 49-60.

Strahler, A.N., 1952. Hypsometric (area-altitude) analysis of erosional topology. *Geological Society of America Bulletin* **63**, 1117–1142

Suzuki, Y., Tokimatsu, K., Moss, R.E.S., Seed, R.B., and Kayen, R.E., 2003. CPT-based liquefaction case histories from the 1995 Hyogoken-Nambu (Kobe) earthquake, Japan. Geotechnical Engineering Research Report No. UCB/GE-2003/03.

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**(2), 550-575.

Toprak, S., and Holzer, T.L., 2003. Liquefaction potential index: field assessment. *JGGE* **129**(4), 315-322.

Turner, B., Brandenberg, S., and Stewart, J. (2016). Case Study of Parallel Bridges Affected by Liquefaction and Lateral Spreading. *JGGE* **142**(7), 05016001.

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., 2014a. Assessment of liquefaction-induced land damage for residential Christchurch. *Earthquake Spectra* **30**(1), 31-55.

van Ballegooy, S., Cox S.C., Thurlow C., Rutter H.K., Reynolds, T., Harrington, G., Fraser, J., and Smith, T., 2014b. Median water table elevation in Christchurch and surrounding area after the 4 September 2010 Darfield earthquake: Version 2. *GNS Science Report 2014/18*, 2014b.

van Ballegooy, S., Green, R.A., Lees, J., Wentz, F., and Maurer, B.W., 2015. Assessment of various CPT based liquefaction severity index frameworks relative to the Ishihara (1985) H_1-H_2 boundary curves. *SDEE* **79**, 347-364.

Wald, D. J., and Allen, T. I., 2007. Topographic slope as a proxy for seismic site conditions and amplification. *BSSA* **97**, 1379–1395.

Worden, C.B., and D.J. Wald, 2016. ShakeMap: technical manual, user's guide, and software guide, US Geological Survey, usgs.github.io/shakemap. DOI: 10.5066/F7D21VPQ.

Whitman, R.V., 1971. Resistance of soil to liquefaction and settlement, *Soils & Foundations* **11**(4), 59–68.

Youd, T.L., and Hoose, S.N., 1977. Liquefaction Susceptibility and Geologic Setting. *Proceedings of the Sixth World Conference on Earthquake Engineering* **3**, 2189-2194.

Youd, T.L., Idriss, I.M., Andrus, R.D., Arango, I., Castro, G., Christian, J.T., Dobry, R., Finn, W.D.L., Harder, L.F., Hynes, M.E. et al. 2001. Liquefaction resistance of soils: summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils. *JGGE* **127**, 817–833

Youd, T.L., and Carter, B.L., 2005. Influence of soil softening and liquefaction on spectral acceleration, *JGGE* **131**(7), 811-825.

Youd, T.L., DeDen, D.W., Bray, J.D., Sancio, R., Cetin, K.O., and Gerber, T.M., 2009. Zero displacement lateral spreads, 1999 Kocaeli, Turkey, earthquake. *JGGE* **135**(1), 46-61.

Zhang, G., Robertson, P. K., and Brachman, R.W.I., 2002. Estimating liquefaction-induced ground settlements from CPT for level ground. *Canadian Geotechnical Journal* **39**(5), 1168-1180.

Zhang, G., Robertson, P.K., and Brachman, R.W.I, 2004. Estimating liquefaction-induced lateral displacements using the standard penetration test or cone penetration test. *JGGE* **130**(8), 861-871.

Zhu, J., Daley, D., Baise, L.G., Thompson, E.M., Wald, D.J., and Knudsen, K.L., 2015. A geospatial liquefaction model for rapid response and loss estimation. *Earthquake Spectra* **31**(3), 1813-1837.

Zhu, J., Baise, L.G., and Thompson, E.M., 2017. An updated geospatial liquefaction model for global application. *BSSA* **107**(3), 1365-1385.

Zimmaro, P., Brandenberg, S.J., Bozorgnia, Y., Stewart, J., Kwak, D.Y., Cetin, K.O., Can, G., Ilgac, M., Franke, K.W., Moss, R.E.S., Kramer, K.L., Stamatakos, J., Juckett, M., and Weaver, T., 2019. Quality control for next-generation liquefaction case histories. *7th Int. Conference on Earthquake Geotechnical Engineering (VII ICEGE)*, 5905-5912, Rome, Italy.

Ziotopoulou, K., and Boulanger, R.W., 2016. Plasticity modeling of liquefaction effects under sloping ground and irregular cyclic loading conditions. *SDEE* **84**, 269-283.

Zou, K.H., 2007. Receiver operating characteristic (ROC) literature research. *On-line bibliography*, http://www.spl.harvard.edu/archive/spl-pre2007/pages/ppl/zou/roc.html

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	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology*, In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

FRAGILITY FUNCTIONS FOR LIQUEFACTION-INDUCED GROUND FAILURE

Mertcan Geyin, S.M.ASCE[1] and Brett W. Maurer, M.ASCE[2]

Abstract

The predicted severity of liquefaction manifested at the ground surface is a popular and pragmatic proxy of damage potential for infrastructure. Towards this end, the Liquefaction Potential Index (*LPI*) and similar models are commonly used, and often codified, to predict surface manifestations on level ground. These predictions typically use deterministic thresholds from the literature - obtained via calibration on case-history data - to classify the expected manifestation. While widely adopted, such thresholds obscure the uncertainty of expected outcomes and are incompatible with probabilistic frameworks. Proposed thresholds are also intimately tied to the liquefaction analytics used to compute them and to the methodology used to select them, each of which can conflict with forward applications, leading to erroneous predictions. Accordingly, using 15,223 case histories from 24 earthquakes, this study develops fragility functions that probabilistically predict surficial manifestations of liquefaction using triggering and manifestation models popular in practice. Deterministic workflows are easily extended by selecting appropriate fragility coefficients; options are provided for six CPT-based triggering models, one CPT-inversion filter, three manifestation models, and three manifestation severities. Model application is demonstrated by predicting: (i) liquefaction manifestations in Christchurch, NZ, resulting from an Alpine Fault earthquake, wherein a logic-tree is used to ensemble predictions from 18 models; and (ii) the return period of liquefaction manifestations in the SODO district of Seattle, USA, wherein predictions are compared to historical observations.

[1]Graduate Research Asst., Civil & Environmental Eng, University of Washington, Seattle, WA (mgeyin@uw.edu)
[2]Asst. Professor, Civil & Environmental Eng, University of Washington, Seattle, WA (bwmaurer@uw.edu)

Introduction

Proposed herein are fragility functions for probabilistically predicting the occurrence and severity of liquefaction-induced land damage on level ground. These functions are developed as extensions to deterministic liquefaction triggering and manifestation models popular in practice, such that users need only select fragility coefficients, presented in Tables 4 and 5, corresponding to the models of choice. Following development and analysis of the fragility functions, application is demonstrated via (i) a scenario earthquake simulation, and (ii) within a probabilistic seismic-hazard analysis.

The severity of soil liquefaction manifested at the ground surface is a pragmatic proxy of damage potential for many types of infrastructure (e.g., shallow foundations and lifelines). The greater the severity of surface manifestation, the greater the likelihood of damage. By way of this proxy, manifestation models have been proposed to link the computed factor of safety against liquefaction triggering (FS_{liq}) at-depth within a profile to damage potential at the surface. Iwasaki et al. (1978) proposed what may be the first manifestation model – the Liquefaction Potential Index (*LPI*) – which has been used in countless studies worldwide (e.g., among many, Luna and Frost 2000; Rix and Romero 2001; Holzer et al. 2006; Lenz and Baise 2007; Hayati and Andrus 2008, Chen et al. 2016; Boulanger et al. 2018). Other manifestation models also in current use include, but are not limited to, the Liquefaction Severity Number (*LSN*) (van Ballegooy et al. 2014a) and a modified version of *LPI* termed LPI_{ISH} (Maurer et al. 2015a).

Central to the typical use of these and similar models, which will be defined subsequently, are proposed index thresholds for classifying the expected risk or severity of manifestation. For example, Iwasaki et al. (1984) proposed that risk of manifestation is "low" at sites where $LPI \leq 5$, "high" where $5 < LPI \leq 15$, and "very high" where $LPI > 15$. Some, but not all, evaluations of *LPI*

following more recent earthquakes (e.g., Toprak and Holzer 2003; Holzer et al. 2005; Maurer et al. 2014a) have supported the Iwasaki et al. (1984) classification thresholds. In similar fashion, Tokin and Taylor (2013) proposed that manifestations should be "little to none" at sites where $LSN < 10$; "minor" where $10 < LSN < 20$; "moderate to severe" where $20 < LSN < 40$; and "major" where $LSN > 40$. Thus, an *LPI* of 5 and an *LSN* of 10 to 20 may correspond to similar expected outcomes. While manifestation models are widely employed in this manner – using deterministic classification thresholds - it is often unappreciated that such thresholds have innate limitations. Namely, they (i) are unique to the liquefaction analytics used to compute them; (ii) are tied to the method used to select them, and by corollary, to the relative consequences of misprediction assumed therein; and (iii) inherently conceal the probabilities of possible outcomes. These limitations are elaborated as follows.

First, many liquefaction triggering models are available to predict FS_{liq}, an input to *LPI*, LPI_{ISH}, *LSN*, and other manifestation models. These models typically yield different FS_{liq} values for the same soil profile and seismic loading, and thus different manifestation model predictions (Lee et al. 2003; Maurer et al. 2015b). For example, using data from the 2010-2011 Canterbury, New Zealand, earthquakes, Maurer et al. (2015b) found that the relationship between expected manifestation severity and computed *LPI* was unique to the adopted triggering model. That is, each had different optimal classification thresholds. Other procedural variants, such as differences in CPT data processing, correction, or filtering (e.g., Boulanger and DeJong 2018) could systematically alter FS_{liq} values, further biasing proposed thresholds.

Second, researchers have used different methods and justifications – often unstated or nonstandard - to select classification thresholds. For example, Iwasaki et al. (1984) found that among 87 study sites, 80% of sites with manifestations had LPI > 5 and 70% of sites without

manifestations had $LPI < 5$; this led to their proposal of $LPI = 5$ as a classification threshold for predicting manifestations. Implicit to this, and any such threshold, is an assumed economy of misprediction. Iwasaki et al. (1984) implicitly treated the costs of false positives (manifestations are expected but not observed) and false negatives (manifestations are observed but not expected) as similar. If Iwasaki et al. (1984) instead assumed that false negatives were significantly more costly than false positives – which is true of many engineering projects – then their proposed threshold would presumably be less than $LPI = 5$. In other words, the threshold would be not that which minimizes the rate of mispredictions, but rather, that which minimizes the cost of mispredictions. Thus, researchers assuming different misprediction economies will invariably propose different classification thresholds. This presents a problem for forward use when the misprediction economy implicit to a proposed threshold either is unknown or differs from that desired in forward analysis.

Due at least partly to the combined effects of the above, proposed thresholds vary significantly for the same manifestation model and what appear to be equivalent expected outcomes. For example, Iwasaki et al. (1984), Toprak and Holzer (2003), Lee et al. (2003), Kang et al. (2014), Papathanassiou et al. (2015), and Maurer et al. (2015b), each calibrating *LPI* on case-history data, proposed *LPI* thresholds of 5, 5, 13, 14, 13.5, and 5, respectively, for binomially classifying the expected occurrence and non-occurrence of surface manifestation. It follows that classification thresholds proposed by one study could be far from optimal (i.e., result in erroneous predictions) when applied to the methods of another. It is reasonable to assume this is true of other manifestation models, such as LPI_{ISH} and *LSN*, though analogous suites of studies are unavailable.

Third, deterministic classifications mask the uncertainty of ground failure and are incompatible with fully probabilistic frameworks, leading to incomplete characterizations of

hazard and risk. All existing liquefaction models are imperfect. As a result, and counter to how classification thresholds may be interpreted, multiple outcomes are always possible at a given classifier index value. For example, considering the popular threshold of $LPI = 5$ for predicting manifestations, two sites with respective *LPI* values of 5.01 and 4.99 could be classified differently (e.g., "hazardous" vs. "non-hazardous", "damage likely" vs. "damage unlikely", etc.) even though the probability of liquefaction manifestation is identical. Moreover, because this probability is unknown, many "probabilistic" studies have resigned to compute the probability that some classification threshold will be exceeded – commonly $LPI = 5$ – in lieu of computing the probability of some physical outcome. Various liquefaction hazard assessments in North America have employed *LPI* in this way, computing either the probability or return period of *LPI* exceeding a threshold value (e.g., Cramer et al. 2008, 2017, 2018; Holzer 2008; Gathro et al. 2018; Goda et al. 2011). However, because the probabilistic relationship between *LPI* and liquefaction manifestation is not well defined, the results of these assessments cannot be properly interpreted in the context of hazard and risk. The probability of *LPI* exceeding a classification threshold could be 50% at a site of interest, but what is the probability of liquefaction manifestation? This latter, more meaningful probability is unknown and could be nearly any value between 0 and 100%, depending on the means and methods used to select the threshold.

Accordingly, motivated by these uncertainties and limitations, the objective of this study is to formulate fragility functions that probabilistically predict the occurrence and severity of liquefaction manifestations on free-field level ground. Analyzing 15,223 case histories compiled from 24 earthquakes in 9 countries, fragility functions will be conditioned on the *LPI*, LPI_{ISH}, and *LSN* manifestation models. Given the observed dependence of these models on the procedures used to compute their inputs, functions will be separately proposed using six CPT-based triggering

models, each implemented with and without CPT inversion via the Boulanger and DeJong (2018) procedure. This will result in fragility functions for 36 distinct liquefaction models (i.e., 3 manifestation models x 6 triggering models x 2 CPT processing possibilities), allowing users to select fragility coefficients consistent with the models they utilize.

Data

This study analyzes 15,223 case histories, as summarized in Table 1. However, since the majority were compiled from three earthquakes in the Canterbury region of New Zealand, fragility functions will be separately developed and compared for these and the other 21 events, henceforth respectively referred to as the "Canterbury dataset" and "Global dataset."

Canterbury Earthquake Dataset

Earthquakes occurring over the last decade in the Canterbury region of New Zealand have resulted in case-history data of unprecedented quantity and quality. The present study analyzes data compiled from the M_w7.1, 4 Sept. 2010 Darfield earthquake, the M_w6.2, 22 Feb. 2011 Christchurch earthquake, and the M_w5.7, 14 Feb. 2016 Christchurch earthquake. This effort built on a series of successive compilations (Maurer et al. 2014a, 2015c, 2019), augmenting the largest by more than 50% and resulting in a total of 14,948 case histories (following exclusions discussed subsequently). The case histories consist of classifications of liquefaction manifestations, geotechnical and hydrological data, and ground-motion intensity measures.

Table 2. Summary of Liquefaction Case Histories Analyzed

Date	Earthquake	Country	Magnitude (M_w)	Number of Case Histories
16/6/1964	Niigata	Japan	7.60	3
9/2/1971	San Fernando	USA	6.60	2
4/2/1975	Haicheng	China	7.00	2
27/7/1976	Tangshan	China	7.60	10
15/10/1979	Imperial Valley	USA	6.53	7
9/6/1980	Victora (Mexicali)	Mexico	6.33	5
26/4/1981	Westmoreland	USA	5.90	9
26/5/1983	Nihonkai-Chubu	Japan	7.70	2
28/10/1983	Borah Peak	USA	6.88	3
2/3/1987	Edgecumbe	New Zealand	6.60	23
24/11/1987	Elmore Ranch	USA	6.22	2
24/11/1987	Superstition Hills	USA	6.54	8
18/10/1989	Loma Prieta	USA	6.93	67
17/1/1994	Northridge	USA	6.69	3
16/1/1995	Hyogoken-Nambu	Japan	6.90	21
17/8/1999	Kocaeli	Turkey	7.51	16
20/9/1999	Chi-Chi	Taiwan	7.62	34
8/6/2008	Achaia-Ilia	Greece	6.40	2
4/4/2010	El Mayor-Cucapah	Mexico	7.20	3
4/10/2010	Darfield	New Zealand	7.10	5371
22/2/2011	Christchurch	New Zealand	6.20	4806
11/3/2011	Tohoku	Japan	9.00	7
20/5/2012	Emilia	Italy	6.10	46
14/2/2016	Christchurch	New Zealand	5.70	4771
			Total	**15,223**

The fragility functions to be developed herein will predict manifestations of liquefaction on free-field level ground – specifically the occurrence and severity of liquefaction ejecta – rather than any other metric of land damage. Sites with lateral spreading were expressly removed from the dataset because the *LPI*, LPI_{ISH}, and *LSN* manifestation models do not fully account for the factors which cause lateral spreading and thus may predict it poorly (e.g., Maurer et al. 2015b; Rashidian and Gillins 2018). Observations of liquefaction ejecta were compiled by the authors and classified as "none," "minor," "moderate," and "severe" using criteria modified from Green et al. (2014) and given in Table 2. This was accomplished using high-resolution satellite imagery and reconnaissance reports available in the New Zealand Geotechnical Database (NZGD, 2019). Cases for which surface manifestations could not be reliably classified are not included in the dataset. Of

the resulting 14,948 cases compiled from Canterbury, 65% are classified as "none" and 35% are cases in which manifestations were observed and classified in accordance with Table 2.

Table 3. Criteria to Classify the Severity of Liquefaction Manifestation

Classification	**Criteria**
None	No observed liquefaction ejecta or lateral spreading
Minor	Less than 5% of ground surface is covered by ejecta; liquefaction features are generally isolated and smaller than a vehicle width; no lateral spreading.
Moderate	Between 5% and 40% of ground surface is covered by ejecta; liquefaction features are frequently adjoining and larger than a vehicle width; streets are generally passable; no lateral spreading.
Severe	More than 40% of ground surface is covered by ejecta; contiguous masses of liquefaction that are larger than a vehicle width; streets are generally impassable; no lateral spreading.

Source: Data from Green et al. (2014)

CPT data was compiled from the New Zealand Geotechnical Database at sites where liquefaction manifestations were classified as described above. In compiling case-histories, CPTs were rejected: (1) if the depth of "pre-drill" significantly exceeded the depth to ground water; and (2) if inferred from geospatial autocorrelation analysis (Anselin 1995) to have prematurely terminated (e.g., due to impedance from gravel) at a depth beneath which liquefiable soil could be present. Additional coverage of the CPT data and autocorrelation analyses is provided in Maurer et al. (2014a, 2015b); further processing and use of the CPT data are discussed subsequently. Ground water depths at CPT locations were obtained from the time-dependent regional models of van Ballegooy et al. (2014b), which were derived, in part, using monitoring data from ~1000 piezometers. Peak Ground Accelerations (*PGA*s) were estimated via the Bradley (2013) method, which has previously been used in Canterbury research (e.g., Maurer et al. 2014b; van Ballegooy

et al. 2015) and which geostatistically coalesces instrumentally-recorded *PGA*s with *PGA*s from ground-motion prediction equations.

Global Dataset

To compare fragility functions in Canterbury with regions worldwide, 274 case histories were compiled from 21 global earthquakes in nine countries. These cases were obtained from the existing literature, including CPT soundings, observations of liquefaction, and estimations of ground water depth and *PGA*, as generally reported by original investigators. When available, refinements were adopted from more recent literature. Whereas liquefaction in Canterbury was intensively cataloged via reconnaissance and remote sensing, the case histories in the global dataset are typically documented in less detail, often with scant information about the nature or severity of manifestation. Since most global cases do not support use of the Green et al. (2014) classification criteria, a binomial "Manifestation" or "No Manifestation" classification was adopted instead. The implications of using two classification schemes will be discussed later. Of the 274 cases compiled, 58% are "Manifestation" and 42% are "No Manifestation." To properly recognize all sources of data used to compile the global dataset, data and references are provided in Table S1 (electronic supplement) for each case history. In this regard, the data assemblages of Moss (2003) and Boulanger and Idriss (2014) greatly assisted the present effort.

Table 3. Summary of Liquefaction Triggering and Manifestation Models used in this Study

Model	Type	Symbol
Robertson & Wride (1998)	Triggering	RW98
Arch. Institute Japan (2001)	Triggering	AIJ01
Moss et al. (2006)	Triggering	Mea06
Idriss & Boulanger (2008)	Triggering	IB08
Boulanger & Idriss (2014)	Triggering	BI14
Green et al. (2019)	Triggering	Gea19
Iwasaki et al. (1978)	Manifestation	*LPI*
van Ballegooy et al. (2014)	Manifestation	*LSN*
Maurer et al. (2015a)	Manifestation	LPI_{ISH}

Methodology

Fragility functions will be conditioned on three liquefaction manifestation models, each computed using six CPT-based liquefaction triggering models, implemented with and without CPT inverse filtering. These models, and the symbology henceforth used to identify them, are summarized in Table 3. The methodologies underlying the CPT processing, liquefaction modeling, and fragility function development are detailed as follows. In addition, all calculations performed herein can be carried out via the open-source software program *Horizon* (Geyin and Maurer, 2020).

CPT Processing Methodology

The CPT offers advantages among in-situ tests used to predict liquefaction (NRC 2016). Yet, as a penetration test, the CPT is still potentially limited by the volume of soil mobilized around the cone, which acts as a physical low-pass filter on the true soil stratigraphy. This filter removes information from the high spatial frequencies, such as the data defining a thin soil stratum or the interface between two disparate soils. These smoothing effects, which are commonly referred to

as “thin layer” and “transition” effects, have long been recognized and studied (e.g., Treadwell 1976; Lunne et al. 1997; Ahmadi and Robertson 2005; Robertson 2011; van der Linden 2016). While chart-based methods exist for manually correcting these effects on CPT data, Boulanger and DeJong (2018) proposed the first programmable procedure. This methodology, referred to as an “inverse filtering and interface detection” procedure, predicts the “true” CPT profile from measured CPT values. Since these measured values reflect a filtered view of reality, their correction would improve subsurface characterization. As a demonstration of the methodology, CPT data is shown in Fig. 1, both with and without correction.

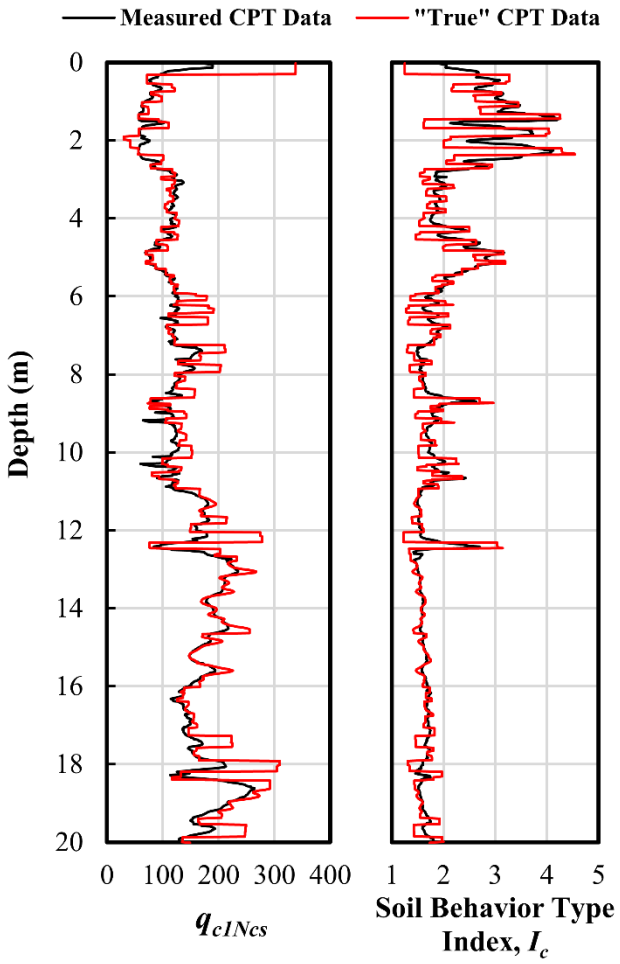


Fig. 1. Example CPT data with (i.e. “true”) and without (i.e. “measured”) inverse-filtering and interface correction via the Boulanger and DeJong (2018) procedure.

While the performance of Boulanger and DeJong's (2018) procedure has not yet been evaluated in the literature, its use can change a site's perceived liquefaction hazard, with the direction and magnitude of change dependent on numerous factors. Considering this potential influence, and that the Boulanger and DeJong (2018) procedure could soon become popular, both measured and "true" CPT data will be used to develop fragility functions, thereby providing users with the same option. While the reader is referred to Boulanger and DeJong (2018) for complete details, the procedure's "baseline" parameters were used to compute "true" CPT data. This was the case both for the methods which invert tip resistance and sleeve friction, and that which detects and corrects stratigraphic interfaces. These defaults can conceivably be calibrated via site-specific study (e.g., from borings adjacent to a CPT), but the information compiled for this study either was insufficient to attempt calibration or provided insufficient statistical support to justify it. However, the sensitivity of results to these parameters will be investigated and discussed later in the paper. As part of the processing methodology, CPT tip- and sleeve-measurements were aligned using cross-correlation (Buck et al. 2002), both for measured and "true" CPT data.

Liquefaction Triggering and Manifestation Model Methodology

Six triggering models, as summarized in Table 3, were used to compute the factor-of-safety against liquefaction (FS_{liq}) vs. depth for each CPT. While the reader is referred to these publications for complete details, two nuances pertinent to this study are as follows. First, prior to using any of the six models, liquefaction susceptibility was inferred using the CPT soil-behavior-type index (I_c) (Robertson and Wride 1998), such that soils with $I_c > 2.50$ were assumed not susceptible. This criterion was developed specifically for the Canterbury dataset using lab and field test data (Maurer et al. 2019). However, because an I_c threshold of 2.50 is within the range of

general, commonly used values (e.g., 2.4-2.6) (Youd et al. 2001), this criterion is also adopted in analyses of the global dataset. Ultimately, the results of this study were insensitive to this decision. Second, for liquefaction-susceptible soils, the IB08, BI14, and Gea19 models compute liquefaction resistance as a function of fines-content (*FC*). Accordingly, *FC* was estimated for the Canterbury dataset using a Canterbury-specific $I_c - FC$ correlation (Maurer et al. 2019), and for the global dataset using a global $I_c - FC$ correlation (Boulanger and Idriss 2014), with the latter estimating *FC* to be ~10% less for a given I_c.

Next, the results from triggering analysis were input to the *LPI*, *LSN*, and LPI_{ISH} manifestation models, which have been given other general names in the literature, including liquefaction hazard frameworks, vulnerability parameters, and damage indices. Nomenclature aside, these models have the same basic objective - to characterize the system-response of a liquefiable soil profile, thereby linking seismic demand to ground failure.

The Liquefaction Potential Index (*LPI*) is defined as (Iwasaki et al. 1978):

$$LPI = \int_0^{20\,m} F(FS_{liq}) \cdot w(z)\,\mathrm{d}z \tag{1}$$

where $F(FS_{liq})$ and $w(z)$ are functions that weight the respective influences of FS_{liq} and depth, z, on surface manifestation. Specifically, $F(FS_{liq}) = 1 - FS_{liq}$ for $FS_{liq} \leq 1$ and $F(FS_{liq}) = 0$ otherwise; $w(z) = 10 - 0.5z$. *LPI* thus assumes that surface manifestation depends on the thickness of all liquefied strata in a profile's upper 20 m, their proximity to the ground surface, and the amount by which FS_{liq} in each stratum is less than 1.0. Given this definition, *LPI* can range from zero to 100.

A modified *LPI* was proposed by Maurer et al. (2015a) and inspired by Ishihara (1985), who proposed limit-state curves for predicting manifestations as a function of the "crust" thickness (H_1), among other factors. Using these curves, Maurer et al. (2015a) modified *LPI* to include the observed influence of H_1. Given its provenance, the result was termed LPI_{ISH} and is defined by:

$$LPI_{ISH} = \int_{H_1}^{20\ m} F(FS_{liq}) \cdot w(z)\ \mathrm{d}z \tag{2a}$$

where

$$F\left(FS_{liq}\right) = \begin{cases} 1 - FS_{liq} & if\ FS_{liq} \leq 1\ \cap\ H_1 \cdot m\left(FS_{liq}\right) \leq 3 \\ 0 & otherwise \end{cases} \tag{2b}$$

$$m\left(FS_{liq}\right) = \exp\left(\frac{5}{25.56\left(1-FS_{liq}\right)}\right) - 1 \tag{2c}$$

In Eq. (2a), $F(FS_{liq})$ and $w(z)$ have the same objective as in *LPI*, but are functionally different, such that $F(FS_{liq})$ accounts for the crust thickness through parameter H_1 and $w(z)$ is defined by $w(z) = 25.56 \cdot z^{-1}$. Maurer et al. (2015a) recommended a minimum H_1 of 0.4 m, even if liquefiable soils are present at shallower depths. Provided this constraint, LPI_{ISH} can range from zero to 100.

The Liquefaction Severity Number (*LSN*) is adapted from methods for estimating post-liquefaction volumetric strain (e.g., to predict ground settlement), modified to include a power-law depth weighting function (van Ballegooy et al. 2014a):

$$LSN = \int_{0}^{20\ m} \varepsilon_v \cdot w(z)\ \mathrm{d}z \tag{3}$$

where ε_v is volumetric strain (%) and $w(z) = 10 \cdot z^{-1}$. While there are many methods to estimate ε_v (e.g., Geyin and Maurer 2019), van Ballegooy et al. (2014a) used that of Zhang et al. (2002), which we also adopt. *LSN* values can surpass 100 if liquefiable soils are near the surface, but typically are between zero and 100. These values are not quantities of predicted settlement, but rather, are index values á la *LPI* and LPI_{ISH} that correlate to the probability of surface manifestation.

Fragility Function Methodology

The probability of surface manifestations reaching or exceeding a defined manifestation severity (MS_i), given a computed liquefaction manifestation model (*LMM*) value, is herein denoted

$F_{MS_i}(LMM)$, where *i* corresponds to the severity of manifestation (e.g., 1 = minor, 2 = moderate, etc.). Adopting the lognormal cumulative distribution function, as is common for fragility functions (e.g., Bradley 2010; Kwak et al. 2016), and which best fit the data relative to other distributions (e.g., beta, chi-squared), $F_{MS_i}(LMM)$ is defined by:

$$F_{MS_i}(LMM) = \Phi\left(\frac{ln(LMM) - ln(\theta)}{\beta}\right) \quad (4)$$

where Φ is the Gaussian cumulative distribution function and θ and β are the distribution's median and logarithmic standard deviation, respectively. In this context, θ is the value of the *LMM* (i.e., *LPI*, LPI_{ISH}, or *LSN*) corresponding to a 50% probability of exceeding a given MS_i .

Several methods exist for fitting fragility functions to empirical data (e.g., Baker 2015; Porter 2019). These include: (1) maximum likelihood estimation; (2) logistic regression; or (3) as utilized in this work, minimizing the squared error term, ε^2 (θ, β), defined as (Porter 2019):

$$\varepsilon^2(\theta, \beta) = \sum_{j=1}^{m} n_j \cdot \left(F_{MS_i} - \frac{f_j}{n_j}\right) \quad (5)$$

where *m* is the number of bins into which similar *LMM* values are grouped; *j* is the bin index; and n_j is the total number of cases in each bin, of which f_j are cases for which observed manifestations reached or exceeded a given MS_i. While the adopted approach is attractive considering the large number of compiled case histories (Porter 2019), fragility functions were found to be insensitive to this choice (i.e., very similar model parameters were obtained using each of the methods above). To investigate the uncertainty of fragility functions due to finite case-history data, non-parametric bootstrap sampling (e.g., Diaconis and Efron 1983) was used to generate 10,000 realizations of both the Canterbury and global datasets, with subsamples (i.e., realizations) randomly selected and equal in size to that of the respective, original datasets. By fitting a fragility function to each realization, distributions of possible function parameters are produced, thus quantifying finite-

sample uncertainty. All else being equal, this uncertainty should diminish as more case history data is compiled and analyzed. For this study, the 16th, 50th, and 84th percentile fragility-functions will be reported and discussed, though only the 50th percentile (or median) functions will be of interest to most users.

Results and Discussion

Analyzing 15,223 liquefaction case histories, empirical fragility functions conditioned on *LPI*, LPI_{ISH}, and *LSN* were developed, each separately formulated using six triggering models, implemented with and without CPT inverse filtering. All functions are defined by Eq. (4), such that users can easily select fragility coefficients consistent with the analytics of their choosing. In this regard, θ and β values for functions developed using the measured and "true" CPT data (i.e., inverse filtered) are presented in Tables 4 and 5, respectively. As discussed, the Canterbury and global datasets deviate in the classification of manifestations. As a result, analyses of the global dataset result in one function for "any manifestation", whereas analyses of the Canterbury dataset result in three functions for "minor," "moderate," and "severe manifestations." In total, 144 functions are defined in Tables 4 and 5, from which select results will be plotted and discussed.

As demonstrated in Fig. 2a, the functions relate the probability of reaching or exceeding a defined manifestation severity (MS_i) to a computed *LMM* value; in this case, BI14-*LPI* using measured CPT data from Canterbury. As an example, the probabilities of manifestations being *at least* minor, moderate, and severe at *LPI* = 5 are approximately 47%, 16%, and 1%, respectively, per the median functions. Three observations from Fig. 2a, generally true of all *LMMs*, are: (1) BI14-*LPI* is more effective at predicting the presence of *some* manifestation than it is at distinguishing manifestation severity, as evidenced by the flatter functions as MS_i increases; (2) finite-sample uncertainty increases as MS_i increases, as evidenced by the dispersion of bootstrap

simulations; and (3) uncertainty also increases as *LPI* increases, principally due to less data at very large *LPI*. Whereas uncertainties in the functions for minor manifestations are relatively inconsequential with respect to computed hazard and risk, those in the functions for severe manifestations should potentially be considered, especially at large *LPI*. With simple arithmetic, the fragility functions can also assess the probability that manifestations will be *in* a severity class (i.e., F_{MS_i} - $F_{MS_{i+1}}$). This is demonstrated in Fig. 2b using the functions from Fig. 2a; for clarity, only median functions are shown. Again using *LPI* = 5 as an example, the probabilities of manifestations being none, minor, moderate, and severe are 53%, 31%, 15%, and 1%, respectively.

In Fig 3., median functions resulting from Canterbury (and shown in Fig. 2a) are compared to results from the Global dataset, for which bootstrap simulations are also plotted. All functions shown are based on BI14-*LPI* and measured CPT data (i.e., without inverse filtering). It can be seen in Fig. 3 that the global function for "any manifestation" has a lesser probability, on average, than the Canterbury function for "minor manifestation." As an example, the *LPI* values corresponding to 50% probability of exceedance, globally and in Canterbury, are respectively 5.5 and 7.5. One possible explanation is that historical criteria for documenting global case histories as "yes" (i.e., some manifestation) tend to fall between the Green et al. (2014) criteria for "minor" and "moderate" manifestations. This is plausible (e.g., given the lack of near-real-time remote sensing for historic case histories) and could result in the noted discrepancy between Canterbury and global results, which exists for most, but not all, of the developed functions. In can also be seen in Fig. 3 that finite-sample uncertainty is relatively larger for the global dataset. This may be attributable to the global data's greater geologic, geomorphic, and seismologic diversity and/or because the global field-data (e.g., CPTs, *PGA*s) were collected over many decades by different investigators. There are also far fewer global case histories; all else being equal, greater finite-

sample uncertainty is thus expected. Given (i) the complications of directly comparing (e.g., via hypothesis testing) the Canterbury and global functions; and (ii) the large finite-sample uncertainty of the latter, we heuristically conclude that the functions resulting from the two datasets are consistent. On this basis, use of the Canterbury functions elsewhere appears reasonable, but more global case-history data is ultimately needed to confirm this, or to draw other conclusions.

Table 4. Summary of fragility-function parameters for "measured" CPT data (i.e., no inverse filtering)

DATASET	Triggering Model	Manifestation Severity	*LPI*						*LPI$_{ISH}$*						*LSN*					
			β_m	β_{84th}	β_{16th}	θ_m	θ_{84th}	θ_{16th}	β_m	β_{84th}	β_{16th}	θ_m	θ_{84th}	θ_{16th}	β_m	β_{84th}	β_{16th}	θ_m	θ_{84th}	θ_{16th}
Canterbury	RW98	*Minor*	2.284	2.252	2.322	3.334	3.215	3.457	3.591	3.553	3.631	1.454	1.390	1.521	1.488	1.524	1.647	15.678	13.991	15.960
		Moderate	2.327	2.281	2.374	21.196	19.889	22.656	2.997	2.905	3.091	18.206	16.511	20.169	2.013	1.992	2.059	69.745	64.109	74.959
		Severe	1.510	1.369	1.660	94.055	74.973	121.900	2.106	1.922	2.296	145.415	107.552	201.984	2.644	2.272	2.861	1499.660	769.413	2302.398
	AIJ01	*Minor*	1.534	1.518	1.550	9.444	9.192	9.707	1.811	1.789	1.834	6.715	6.508	6.932	2.328	2.298	2.362	40.855	39.206	42.652
		Moderate	1.519	1.489	1.551	31.624	30.254	33.132	1.874	1.822	1.924	29.606	27.802	31.552	2.805	2.743	2.871	287.724	262.058	318.069
		Severe	1.155	1.054	1.257	105.814	89.724	126.755	1.630	1.489	1.761	158.853	125.579	201.344	2.569	2.421	2.720	2811.500	2072.920	3881.065
	MEA06	*Minor*	2.009	1.984	2.035	5.335	5.164	5.514	2.311	2.264	2.369	3.995	3.824	4.177	2.112	2.066	2.214	29.359	27.577	30.641
		Moderate	2.158	2.112	2.208	26.434	24.791	28.322	2.448	2.389	2.511	24.840	23.001	26.963	2.815	2.739	2.915	214.996	188.832	245.634
		Severe	1.739	1.607	1.870	139.089	110.419	177.397	2.139	2.005	2.275	201.200	157.247	260.725	3.292	2.953	3.582	6841.716	3541.488	12138.221
	IB08	*Minor*	1.774	1.759	1.791	4.095	3.984	4.209	2.160	2.118	2.217	2.394	2.300	2.485	1.477	1.458	1.499	14.536	14.162	14.911
		Moderate	1.780	1.748	1.812	18.364	17.522	19.273	2.177	2.130	2.224	14.749	13.888	15.691	1.673	1.627	1.719	51.632	48.848	54.633
		Severe	1.443	1.323	1.569	97.345	79.957	121.503	2.040	1.878	2.211	151.680	115.835	204.020	2.234	2.081	2.359	723.631	546.215	927.259
	BI14	*Minor*	1.537	1.526	1.547	5.582	5.446	5.723	1.706	1.692	1.721	3.278	3.185	3.371	1.347	1.325	1.369	16.873	16.441	17.306
		Moderate	1.368	1.344	1.390	19.420	18.716	20.160	1.633	1.602	1.665	14.559	13.936	15.228	1.472	1.424	1.518	50.771	48.079	53.627
		Severe	1.161	1.074	1.249	80.778	70.127	94.439	1.612	1.502	1.725	102.834	85.511	125.581	2.070	1.918	2.183	607.150	459.576	762.474
	GEA19	*Minor*	1.578	1.568	1.590	5.054	4.928	5.185	1.828	1.811	1.848	2.891	2.801	2.980	1.332	1.312	1.355	16.996	16.570	17.428
		Moderate	1.416	1.391	1.440	18.220	17.533	18.951	1.764	1.729	1.798	14.059	13.413	14.752	1.486	1.440	1.532	51.793	49.090	54.691
		Severe	1.188	1.101	1.278	76.484	66.335	89.686	1.709	1.589	1.831	107.736	88.089	133.568	2.087	1.946	2.199	630.637	484.870	790.043
Global	RW98	*Any Manifestation*	1.392	1.328	1.482	6.129	5.198	7.208	1.963	2.006	2.352	2.903	1.618	4.083	1.158	1.091	1.236	13.096	11.414	14.978
	AIJ01		1.131	1.101	1.175	9.330	8.109	10.683	1.320	1.274	1.387	6.243	5.202	7.370	1.239	1.179	1.306	23.958	20.891	27.398
	MEA06		1.109	1.056	1.173	7.311	6.415	8.351	1.437	1.381	1.544	4.378	3.553	5.294	1.163	1.108	1.226	16.859	14.758	19.205
	IB08		1.436	1.377	1.520	6.993	5.875	8.255	2.264	2.161	2.437	3.116	2.150	4.248	1.147	1.079	1.230	13.148	11.507	15.034
	BI14		1.335	1.280	1.414	7.508	6.401	8.807	2.392	2.273	2.423	2.988	2.097	4.342	1.098	1.037	1.173	13.454	11.829	15.314
	GEA19		1.348	1.289	1.430	6.886	5.857	8.039	3.148	2.577	2.790	1.887	1.573	3.626	1.101	1.041	1.182	13.320	11.709	15.159

Table 5. Summary of fragility-function parameters for "true" CPT data (i.e., inverse-filtering applied)

DATASET	Triggering Model	Manifestation Severity	*LPI*						*LPI_{ISH}*						*LSN*					
			β_m	β_{84th}	β_{16th}	θ_m	θ_{84th}	θ_{16th}	β_m	β_{84th}	β_{16th}	θ_m	θ_{84th}	θ_{16th}	β_m	β_{84th}	β_{16th}	θ_m	θ_{84th}	θ_{16th}
Canterbury	RW98	*Minor*	2.634	2.580	2.688	2.652	2.544	2.770	3.835	3.795	3.877	1.079	1.030	1.131	1.520	1.561	1.701	13.922	12.238	14.130
		Moderate	2.619	2.564	2.673	20.975	19.468	22.608	3.859	3.773	3.946	22.971	20.715	25.595	2.184	2.159	2.238	70.048	63.798	75.834
		Severe	1.711	1.539	1.889	106.543	80.699	146.447	2.310	2.091	2.524	173.838	121.238	254.150	2.748	2.431	3.046	1615.648	884.015	2834.070
	AIJ01	*Minor*	1.525	1.509	1.541	7.414	7.220	7.611	1.818	1.793	1.841	5.274	5.112	5.446	2.365	2.331	2.399	33.459	32.097	34.909
		Moderate	1.626	1.595	1.658	26.390	25.193	27.685	2.027	1.982	2.073	25.576	24.070	27.229	2.887	2.823	2.951	249.636	226.816	275.562
		Severe	1.369	1.262	1.475	111.793	93.008	136.043	1.874	1.744	2.009	182.448	144.818	234.428	2.795	2.630	2.950	3345.554	2391.550	4673.310
	MEA06	*Minor*	1.998	1.972	2.025	4.370	4.232	4.518	2.739	2.623	2.862	3.271	3.078	3.496	2.129	2.071	2.235	26.175	24.838	27.279
		Moderate	2.233	2.188	2.279	22.954	21.517	24.574	2.685	2.616	2.762	24.297	22.273	26.704	2.884	2.805	2.986	203.636	179.207	231.624
		Severe	1.778	1.634	1.916	124.116	96.492	160.766	2.317	2.155	2.478	229.648	171.609	312.091	3.424	3.110	3.643	7702.620	4144.428	12137.467
	IB08	*Minor*	1.935	1.914	1.956	3.468	3.368	3.574	3.403	3.363	3.443	1.635	1.569	1.704	1.612	1.596	1.636	13.196	12.740	13.584
		Moderate	2.054	2.019	2.092	18.425	17.483	19.487	2.657	2.587	2.732	16.488	15.202	17.973	1.930	1.879	1.982	55.865	52.253	59.697
		Severe	1.437	1.312	1.576	86.999	70.804	110.977	2.018	1.846	2.194	137.961	103.792	188.058	2.320	2.085	2.481	767.169	509.529	1046.681
	BI14	*Minor*	1.617	1.604	1.630	4.760	4.638	4.885	1.863	1.842	1.889	2.858	2.759	2.955	1.498	1.478	1.522	15.426	14.952	15.885
		Moderate	1.582	1.555	1.607	18.524	17.770	19.324	1.882	1.846	1.919	14.572	13.857	15.355	1.738	1.678	1.793	55.468	51.596	59.330
		Severe	1.151	1.063	1.249	70.628	61.093	83.700	1.602	1.481	1.727	93.419	76.514	116.262	2.155	1.922	2.315	643.064	428.249	874.953
	GEA19	*Minor*	1.670	1.656	1.684	4.315	4.202	4.432	2.168	2.131	2.309	2.432	2.270	2.554	1.480	1.461	1.503	15.521	15.056	15.977
		Moderate	1.652	1.624	1.679	17.665	16.918	18.469	2.074	2.031	2.116	14.577	13.768	15.452	1.749	1.688	1.803	56.573	52.682	60.514
		Severe	1.185	1.092	1.284	67.859	58.323	80.698	1.737	1.600	1.879	103.846	82.701	133.178	2.175	1.947	2.330	670.897	451.459	907.770
Global	RW98	*Any Manifestation*	1.569	1.522	1.631	5.827	4.807	6.990	3.421	2.841	3.012	1.586	1.261	3.058	1.336	1.293	1.389	12.922	11.034	15.132
	AIJ01		1.362	1.330	1.406	8.298	7.058	9.721	1.767	1.807	1.853	4.965	3.657	6.276	1.522	1.474	1.590	20.076	16.748	23.990
	MEA06		1.382	1.345	1.434	6.437	5.397	7.612	1.634	1.606	1.780	4.304	3.270	5.368	1.550	1.529	1.585	15.592	12.803	18.802
	IB08		1.543	1.497	1.606	6.682	5.582	7.941	2.146	2.354	2.537	3.696	1.915	5.175	1.296	1.247	1.363	13.213	11.326	15.384
	BI14		1.442	1.409	1.495	7.561	6.360	8.914	1.930	2.208	2.299	4.344	2.350	5.765	1.257	1.207	1.326	13.633	11.652	15.922
	GEA19		1.425	1.390	1.489	6.898	5.827	8.119	3.808	2.987	3.101	1.765	1.512	4.038	1.270	1.216	1.337	13.464	11.547	15.718

Median fragility functions for *LPI*, computed using each of the six triggering models, are shown for the Canterbury and global datasets in Figs. 4a and 4b, respectively. As see in Fig. 4, the relationship between expected manifestation severity and computed *LPI* is unique to the adopted triggering model. Using *LPI* = 5 as an example, the probabilities of manifestations being at least minor, using measured CPTs from Canterbury, are 57%, 34%, 49%, 54%, 47%, and 50% for the RW98, AIJ01, MEA06, IB08, BI14, and GEA19 models, respectively. Using measured CPTs from the global dataset, the probabilities of any manifestation are 44%, 29%, 37%, 41%, 38%, and 41% using the same respective models. Therefore, *LPI* values tend to be highest when computed using AIJ01 and lowest when computed using RW98, both globally and in Canterbury. Clearly, the fragility coefficients obtained with one model should not be used in conjunction with another.

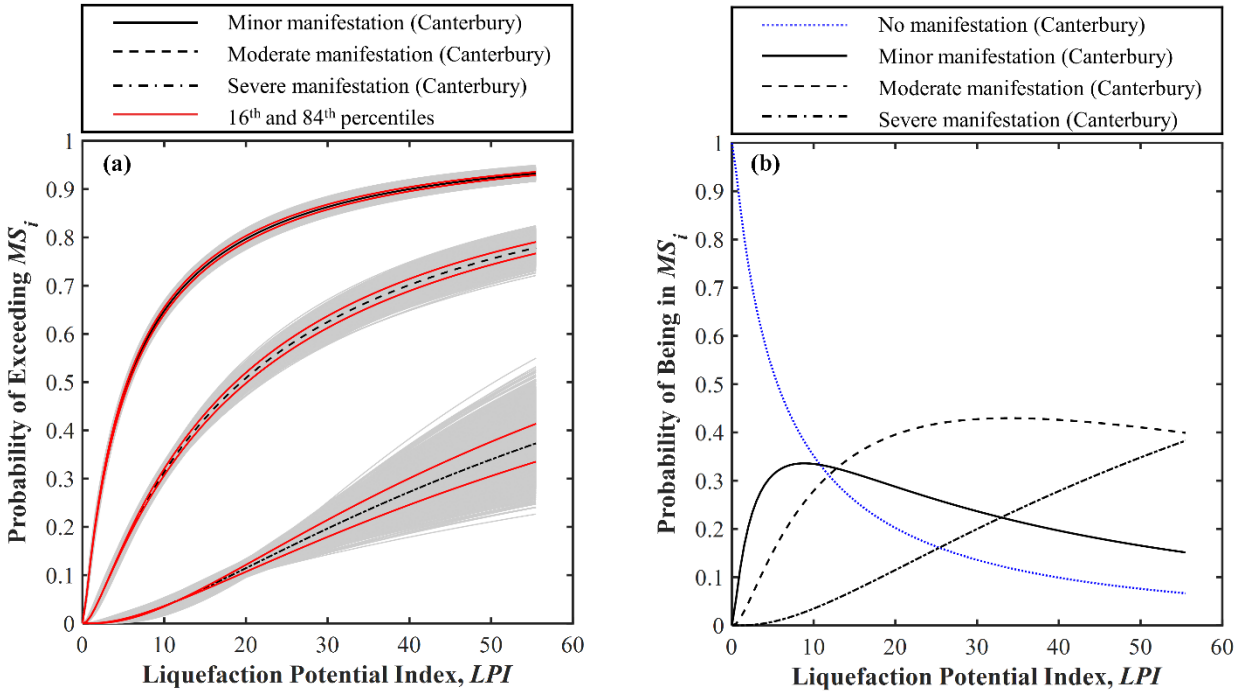


Fig. 2. Probability of manifestations **(a)** exceeding, and **(b)** being in severity class, MS_i, given *LPI* computed using BI14 and Canterbury CPTs (as measured). In **(a)**: grey lines are bootstraps; black lines are median functions.

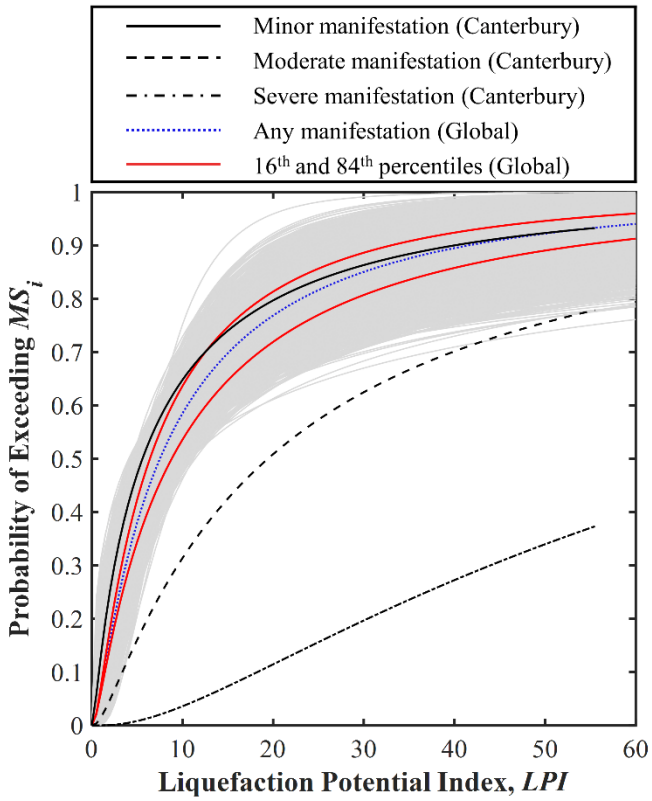


Fig. 3. Probability of manifestations exceeding severity class, MS_i, given *LPI* computed using BI14 and measured CPT data. Global results include bootstrap simulations; Canterbury results are median functions from Fig 1a.

Selecting BI14-*LPI* as a representative example, the effects of CPT inverse-filtering on fragility functions are shown for the Canterbury and global datasets in Figs. 5a and 5b, respectively. In each case, the Boulanger and DeJong (2018) "baseline" parameters were used to compute "true" CPT data. As inferable from Fig. 4a, inverse-filtering has the average tendency to slightly reduce *LPI* values in Canterbury (notably, *LPI* also often increases). Overall, the result is a shift of the formulated functions toward lesser *LPI*. At *LPI* = 5, for example, the probabilities of exceeding minor, moderate, and severe manifestations increase by 4.1%, 4.3%, and 0.25%, respectively, due to inverse filtering. This

shift is often less pronounced in the global functions for predicting any manifestation, ranging from 0.7% for BI14-*LPI* (as shown in Fig. 4b) to 6.43% for AIJ01-*LPI*. While some variation exists, these trends generally also apply to all functions based on LPI_{ISH} and *LSN*. While more rigorous analysis of the inverse-filtering procedure is ongoing, it can be inferred from Fig. 4 (and β values in Table 4 vs. 5) that the procedure does not significantly alter model efficacy (i.e., the ability to segregate sites with and without manifestations), either for better or for worse.

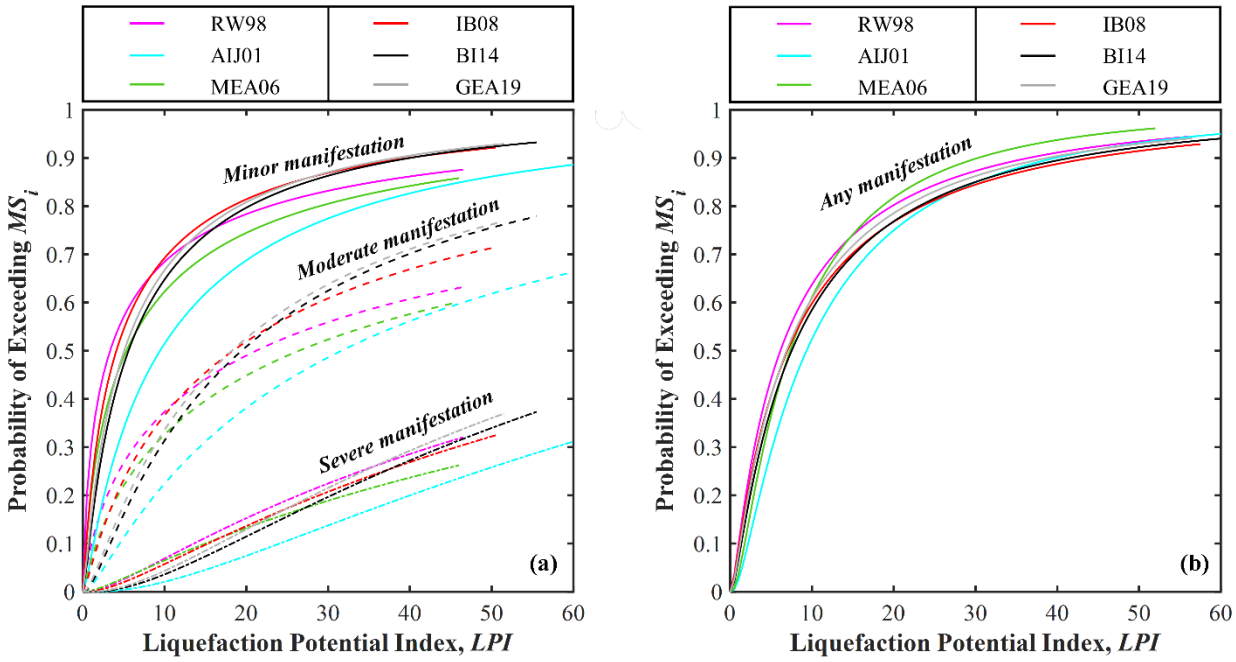


Fig. 4. Probability of manifestations exceeding severity class, MS_i, given *LPI* computed using six different triggering models and CPT data (as measured) from the **(a)** Canterbury and **(b)** global datasets.

A limited parametric study was also performed by varying an influential parameter in the procedure: $z'_{50,ref}$, which controls the "aggression" of the inversion, such that changes to CPT data

increase as $z'_{50,ref}$ increases. Specifically, the "baseline" $z'_{50,ref}$ value of 4.2 was varied from 3.4 to 5.0. While the reader is referred to Boulanger and DeJong (2018) for a full explanation of $z'_{50,ref}$ (and the complete procedure), the results of this analysis are shown in Fig. 5 using BI14-*LPI* and the Canterbury data.

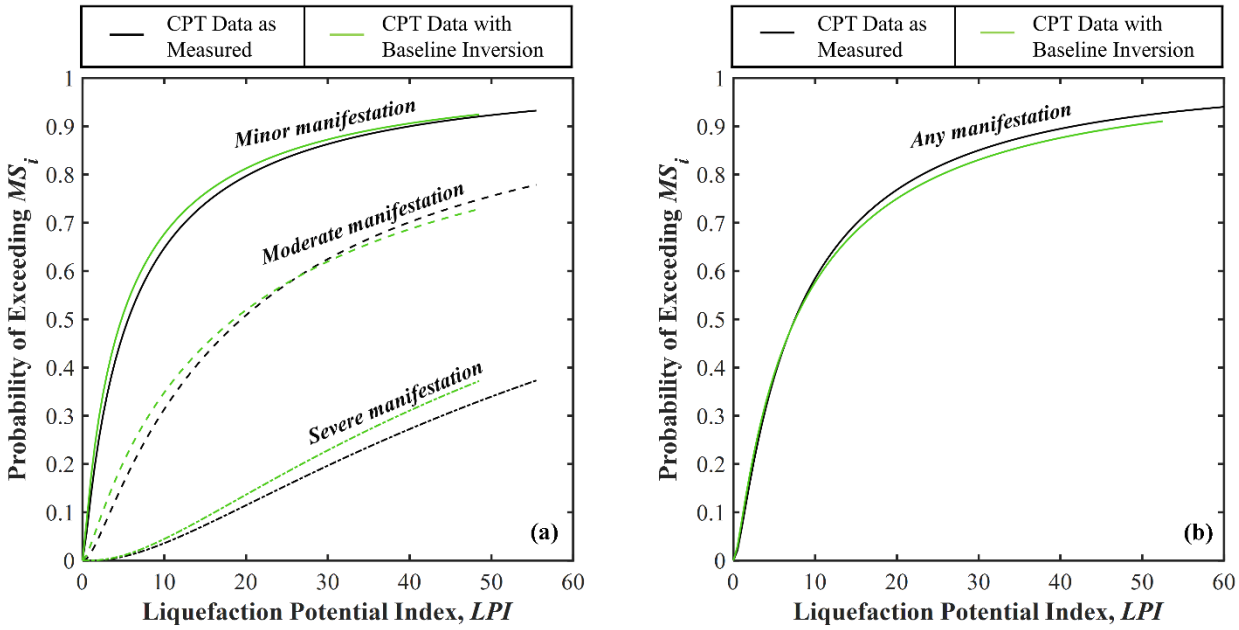


Fig. 5. Probability of manifestations exceeding severity class, MS_i, given *LPI* computed using BI14 with and without inverse filtering of CPT data from the **(a)** Canterbury and **(b)** global datasets.

As expected, larger values of $z'_{50,ref}$ tend to further diminish computed *LPI*, shifting the fragility functions in corresponding fashion. In addition, the functions in Fig. 5 do not suggest marked improvement with respect to predictive performance, regardless of the $z'_{50,ref}$ value. However, the intention of this study is not to recommend use or disuse of the Boulanger and DeJong (2018) procedure, nor is it to intensively analyze the procedure's performance. Rather, the intention is to provide users with fragility functions conditioned on an array of liquefaction analytics. In this regard,

the study's only recommendation is to employ fragility coefficients in a manner perfectly consistent with their development. As shown in Figs. 4-6, failure to do so would invariably introduce some degree of error.

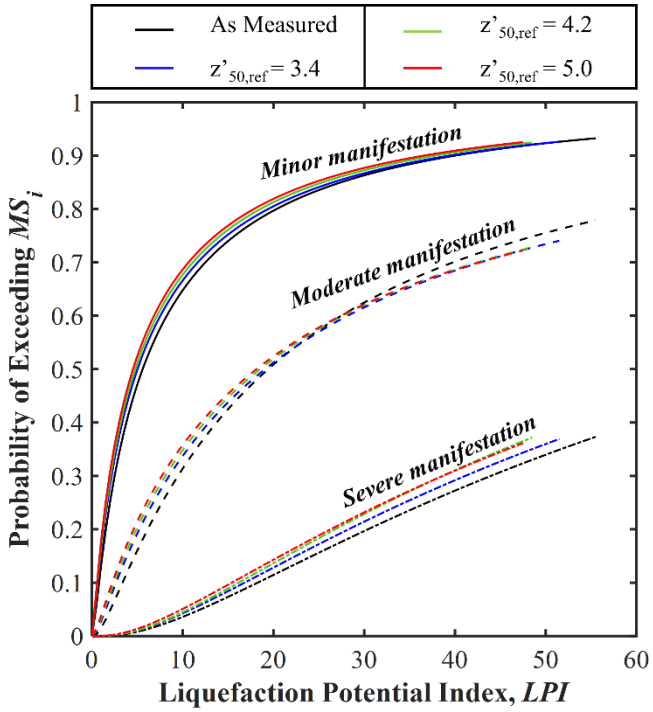


Fig. 6. Probability of manifestations exceeding severity class, MS_i, given *LPI* computed using BI14 with and without inverse filtering of Canterbury CPT data; results for three variants of the inversion procedure are shown.

Demonstration

Fragility-function application will be demonstrated via (i) a scenario earthquake simulation in Christchurch, NZ; and (ii) within a probabilistic seismic-hazard analysis in Seattle, USA.

Scenario Earthquake Simulation: Alpine Fault, New Zealand

New Zealand's 600-km long Alpine Fault (AF) represents a major seismic hazard for the South Island. Believed capable of producing M_w8 earthquakes and to have a 29% probability of rupture in the next 50 years (Cochran et al. 2017), the next AF event will undoubtedly be catastrophic for many. Yet, because there is no historic account of any AF earthquake (the last occurred ca. 1717), the potential extent and severity of ground failure is highly uncertain. Accordingly, to provide an example with utility beyond this paper, the fragility functions are first applied in conjunction with simulated ground-motions (Bradley et al. 2017a) from an $M_w7.9$ AF scenario earthquake. Specifically, measured CPTs from Canterbury will be used to predict liquefaction manifestations in the city of Christchurch and its environs. The physics-based simulation of Bradley et al. (2017a), which explicitly models kinematic fault rupture, wave propagation, and the 3D velocity structure of the subsurface, was obtained from the SeisFinder portal (Bradley et al. 2017b). For this demonstration, predictions will be merged from the 18 fragility functions defined in Table 4 and based on the Canterbury dataset. This ensemble approach, in which models are weighted in proportion to their predictive capabilities, is conventional in seismic hazard analysis and has the advantage of avoiding large "swings" as a result of changing a single adopted model. While the 18 functions could be given equal weighting (i.e., 18^{-1}, or 0.055), a scheme derived from receiver-operating-characteristic (ROC) analysis (Geyin et al. 2020) was adopted. Specifically, weights were assigned in proportion to the area under the ROC curve – a measure of diagnostic efficiency – for each of the 18 models on which fragility functions were conditioned. Taking predictive capabilities into account, the functions received weights ranging from 0.049 to 0.059 (the weighting scheme is detailed in Tables S2 and S3). Presented in Fig. 7 are the resulting probabilities of minor, moderate, and severe manifestations. Notably, manifestations were previously shown to be overpredicted by popular liquefaction analytics in areas of Southwest Christchurch during the 2010-2011 Canterbury sequence (Maurer et al. 2014a).

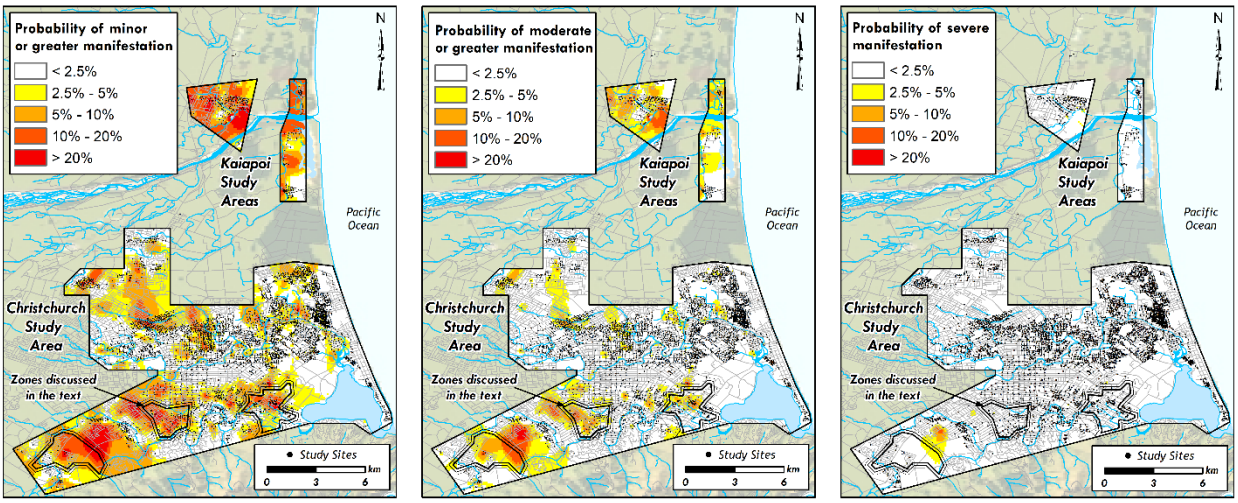


Fig. 7. Probabilities of **(a)** minor, **(b)** moderate, and **(c)** severe liquefaction manifestations in Christchurch, NZ, during an Alpine Fault earthquake. Results in each panel are obtained by weighting predictions from 18 functions. Basemap from Land Information New Zealand (2020).

The non-trivial probabilities of moderate and severe manifestations in these general areas, as loosely delineated on Fig. 7, should thus be viewed skeptically. Nonetheless, the analyses do suggest at least a small probability of *some* manifestation across much of the study area. These manifestations, in general, would likely have minor severity.

Return Period of Liquefaction Manifestations: Seattle, USA

The return period (T_R) of liquefaction manifestations at a site of interest can be computed using: (i) the fragility functions developed herein; (ii) CPT data from the site; and (iii) the *PGA* hazard curve, which describes the mean annual frequency of *PGA* exceeding a given value at the site's location. Using fragility functions conditioned on *LPI* as an example, T_R is computed as:

$$\lambda_{MS} = {}^{1}/_{T_R} = \int_{LPI=0}^{\infty} F_{MS}(LPI) \cdot \left|\frac{d\lambda_{LPI}}{dLPI}\right| \cdot dLPI \quad (6)$$

where F_{MS} (LPI) is the fragility function for a particular *MS* (e.g., minor, moderate, or severe) and defined by Eq. (4); λ_{MS} is the mean annual exceedance frequency of the *MS* (the reciprocal of which is T_R, in years); and $|d\lambda_{LPI}/dLPI|$ is the absolute value of the derivative of the *LPI* hazard curve, which describes the mean annual frequency of *LPI* exceeding a given value (λ_{LPI}), and which is computed from CPT data and the site-specific *PGA* hazard curve. Specifically, to compute a site's λ_{LPI}, the *PGA* hazard curve must first be deaggregated by earthquake magnitude. This requirement results from use of both *PGA* and magnitude as inputs for predicting liquefaction triggering (true of all six triggering models used in this study). λ_{LPI} is then computed as:

$$\lambda_{LPI} = \sum_{j=1}^{N_{Mw}} \sum_{i=1}^{N_{PGA}} P\left(LPI > lpi \middle| pga = pga_i, m = m_j\right) \Delta\lambda_{pga_i, m_j} \quad (7)$$

where N_{Mw} and N_{PGA} are respectively the number of magnitude and *PGA* increments into which the computed seismic hazard is subdivided; $\Delta\lambda_{pgai,mj}$ is the incremental annual-exceedance rate for intensity measure, pga_i, and magnitude, m_j, which follows an established procedure in performance-based liquefaction modeling (Kramer and Mayfield 2007); and $P\left(LPI > lpi \middle| pga = pga_i, m = m_j\right)$ is the binomial probability that *LPI* exceeds some threshold value, *lpi*, conditioned on *PGA* and *M*. This overall approach to computing λ_{LPI} is similar to those presented by Goda et al. (2001) and Green et al. (2020).

To demonstrate this process, a CPT from the South-of-Downtown (SODO) district of Seattle, USA, is analyzed. Deaggregated *PGA* hazard-curve data was obtained for the CPT's location (47.587130, -122.331487), which has D/E seismic site class, from the USGS (2019) Unified Hazard Tool, wherein the 2008 US National Seismic Hazard Model was adopted (Petersen et al., 2008). Using Eqs. 1 and 7, in conjunction with measured CPT data and the Boulanger and Idriss (2014) triggering model, the *LPI* hazard curve was computed and is shown in Fig. 8. This figure, which follows an approach analogous to those demonstrated by Goda et al. (2001) and Green et al. (2020)

shows the expected annual rates at which different *LPI* values will be exceeded. Lastly, using this *LPI* hazard curve within Eq. 6, and computing F_{MS}(*LPI*) with appropriate coefficients from Table 4, the computed return periods of minor, moderate, and severe liquefaction manifestations are 60, 127, and 709 years, respectively. Thus, and assuming a Poisson process, the probabilities of minor, moderate, and severe manifestations occurring at least once in the next 100 years are 81%, 54%, and 13%, respectively.

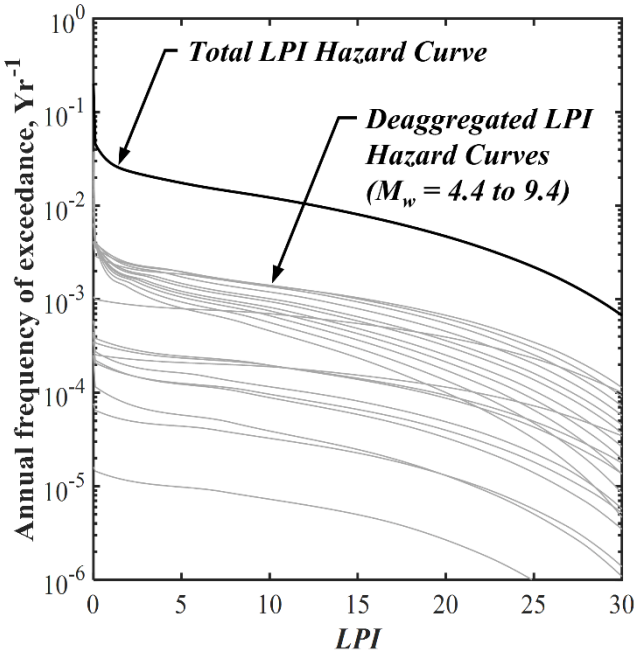


Fig. 8. *LPI* hazard curves deaggregated on earthquake magnitude, such that the total *LPI* hazard curve is equal to the sum of the *LPI* hazard curves for all magnitudes.

Repeating this process for 43 CPTs from the Washington State Department of Natural Resources (2019), the return period of minor manifestations is mapped in Fig. 9 for a 3 km^2 area of SODO. In the context of hazard mapping, planning, and policy, this information is arguably of much greater value than that derived from "probabilistic" analyses or maps focusing on a classification threshold,

the shortcomings of which were previously discussed. Historically, manifestations were observed in the Fig. 9 study area – particularly that with lowest computed return period – following earthquakes in 1949, 1965, and 2001 (Chleborad and Schuster 1990; Bray et al. 2001).

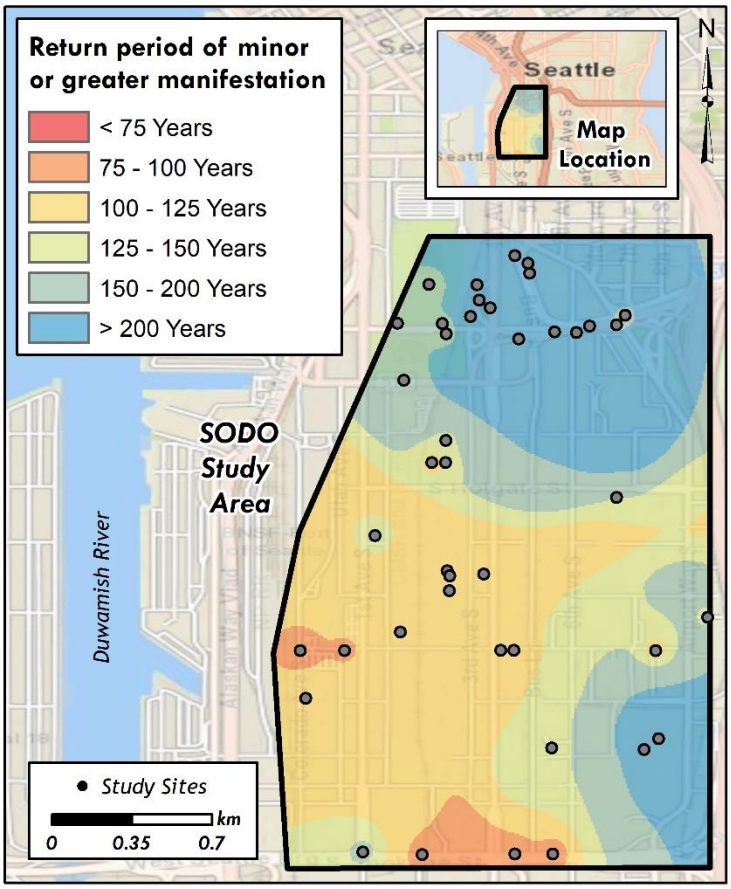


Fig. 9. Computed return period of liquefaction manifestations across SODO study area. Basemap from Esri (2020).

These manifestations were generally "minor" per the Green et al. (2014) criteria. Assuming a Poisson process, and adopting (i) the computed return period of 60 years (see above); and (ii) a 170-year exposure window beginning 1850 (ca. Seattle's founding), the probability of three observations is 22% (the only quantity more likely is two, having 23% probability). With respect to the computed

return period for severe manifestations (i.e., ~700 years), paleoliquefaction from the nearby Duwamish River (Davis et al. 2019) suggests that "severe" manifestations have occurred at least once, and possibly twice, during the last 1200 years. Assuming a Poisson process, it could similarly be shown that these observations are consonant with the computed return period. While discrepancies between observed and predicted return periods would not necessarily discredit the latter, their close agreement nonetheless gives credence to the developed functions and results.

Conclusions

The severity of liquefaction manifested at the ground surface is a pragmatic proxy of damage potential for infrastructure, making it well-suited for hazard mapping, planning, policy, and preliminary site-assessment. Towards this end, empirical fragility functions were formulated to predict the probability of liquefaction manifestations on free-field level ground. These functions are extensions to popular deterministic liquefaction models, such that users need only select fragility coefficients from Table 4 or 5; options were provided for six CPT-based triggering models, one CPT-inversion filter, and three manifestation models. Fragility functions separately developed from case histories globally and in Canterbury were found to be heuristically consistent. This lends permissibility to the use of Canterbury functions elsewhere, which would allow for manifestation severity to be predicted in higher resolution when desirable, yet additional global case-history data is ultimately needed to confirm this, or to draw other conclusions.

While the proposed functions have a variety of uses, they are not intended to predict lateral spreading, which is a distinct phenomenon influenced by factors not considered in this study, nor can the functions explicitly predict damage to specific infrastructure assets. In this regard, liquefaction could trigger at-depth and damage infrastructure without otherwise manifesting at the surface, or could manifest at the surface without causing asset damage. Asset-specific assessments of

liquefaction potential and consequence are thus judicious. Moreover, these functions do not replace the need for improved analytics that more effectively predict the triggering and manifestation of liquefaction. As evident from all fragility functions developed herein, the models on which these functions are conditioned have significant potential for improvement. As one example, interbedded low-permeability soils may complicate prediction of a soil profile's cumulative response by affecting the onset of liquefaction triggering and/or the morphology of manifestation (e.g., Fiegel and Kutter 1994; Brennan and Madabhushi 2005; Özener, et al. 2008). These affects are not considered by the models utilized herein, which may thus perform less efficiently on profiles with complex stratigraphy (e.g., Juang et al.,2005; Maurer et al. 2015c; Yost et al. 2019; Cubrinovski et al. 2019).

Moreover, the results of this study are tied to the data analyzed, which in effect is the present sum of CPT case histories. The applicability of these results to other case-history data – particularly that with different parameter space (e.g., soils with atypical composition, minerology, age, etc.) – or to other models and procedures, is unknown. In addition, the presented findings should be considered in the context of model regionality and possible bias. Ultimately, additional data will confirm or update the fragility functions developed herein.

Data Availability

Some or all data, models, or code generated during the study are available from the corresponding author, including Tables 4 and 5 as well as all data associated with the Canterbury case-history dataset. Tables S1-S3 are available online in the ASCE library (www.ascelibrary.org) and was compiled from data that may be available in full or part from the Next-Generation Liquefaction Project (Brandenberg et al. 2020). In addition, all calculations demonstrated herein, including CPT processing, may be performed using *Horizon* (Geyin and Maurer, 2020), a freely available open-source program developed by the authors.

Acknowledgements

The presented study is based upon work supported by the National Science Foundation (NSF), US Geological Survey (USGS), and Pacific Earthquake Engineering Research Center (PEER) under Grant Nos. CMMI-1751216, G18AP-00006, and 1132-NCTRBM, respectively. 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, USGS, or PEER.

References

Ahmadi, M.M., and Robertson, P.K. (2005). "Thin-layer effects on the CPT q_c measurement." *Canadian Geotechnical Journal*, 42(5), 1302-1317.

Anselin, L. (1995). "Local Indicators of Spatial Association - LISA." *Geograph Analysis*, 27 (2): 93–115.

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

Bradley, B.A. (2010). "Epistemic uncertainties in component fragility functions." *EQS*, 26(1), 41-62.

Baker, J.W. (2015). "Efficient analytical fragility function fitting using dynamic structural analysis." *EQS*, *31*(1), 579-599.

Bradley, B.A. (2013). "Site-specific and spatially-distributed ground motion intensity estimation in the 2010-2011 Christchurch earthquakes." *Soil Dynamics and Earthquake Engineering*, 48, 35-47.

Bradley, B.A., Bae, S.E., Polak, V., Lee, R.L., Thomson, E.M. and Tarbali, K. (2017a). "Ground motion simulations of great earthquakes on the Alpine Fault: effect of hypocenter location and

comparison with empirical modelling." *New Zealand Journal of Geology and Geophysics*, 60(3), 188-198.

Bradley, B. A., Savarimuthu, S., Lagrava, D., Huang, J., Motha, J., Polak, V., & Bae, S. (2017b). "SeisFinder: A web application for extraction of data from computationally-intensive earthquake resilience calculations." *SCEC Annual Meeting*; < https://quakecoresoft.canterbury.ac.nz/seisfinder/>

Brandenberg, S. J., Zimmaro, P., Stewart, J. P., Kwak, D. Y., Franke, K. W., Moss, R. E., … Kramer, S. L. (2020). "Next-generation liquefaction database." *Earthquake Spectra*, Doi: 10.1177/8755293020902477

Bray, J. D., Sancio, R., Kammerer, A. M., Merry, S., Rodriguez-Marek, A., Khazai, B., Chang, S., Bastani, A., Collins, B., Hausler, E., Dreger, D., Perkins, W.J., 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 Center, UCB, University of Arizona, Washington State University, Shannon and Wilson Inc., and Leighton and Associates.*

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

Boulanger, R.W. and DeJong, J.T. (2018). "Inverse filtering procedure to correct cone penetration data for thin-layer and transition effects." *Cone Penetration Testing 2018*, Hicks, Pisano, and Peuchen, eds., Delft University of Technology, The Netherlands, 25-44.

Boulanger, R.W., Khosravi, M., Cox, B.R., DeJong, J.T. (2018). "Liquefaction Evaluation for an Interbedded Soil Deposit: St. Teresa's School, Christchurch, New Zealand." *IACGE 2018*

Geotechnical and Seismic Research and Practices for Sustainability, Chongquing, China, 21- 22 October 2018.

Brennan, A.J., and Madabhushi, S.P. (2005). "Liquefaction and drainage in stratified soil. *Journal of Geotechnical and Geoenvironmental Engineering*, 131(7), 876-885.

Buck, J.R., Daniel, M.M, and Singer, A.C, 2002. *Computer Explorations in Signals and Systems Using MATLAB®*, 2nd Edition. Upper Saddle River, NJ: Prentice Hall.

Chleborad, A.F., and Schuster, R.L. (1990). "Ground failure associated with the Puget Sound region earthquakes of April 13, 1949, and April 29, 1965." *USGS Open-File Report 90-687, 140 p.*

Chen, Q., Wang, C., and Juang, C.H. (2016). "CPT-based evaluation of liquefaction potential accounting for soil spatial variability at multiple levels." *J Geotech and Geoenviron Engineering*, 142(2), 04015077.

Cochran, U.A., Clark, K.J., Howarth, J.D., Biasi, G.P., Langridge, R.M., Villamor, P., Berryman, K.R., and Vandergoes, M.J. (2017). "A plate boundary earthquake record from a wetland adjacent to the Alpine fault in New Zealand refines hazard estimates." *Earth and Planetary Science Letters*, *464*, 175-188.

Cramer, C. H., Rix, G. J., & Tucker, K. (2008). "Probabilistic liquefaction hazard maps for Memphis, Tennessee." *Seismological Research Letters*, *79*(3), 416-423.

Cramer, C.H., Bauer, R.A., Chung, J.W., David Rogers, J., Pierce, L., Voigt, V., Mitchell, B., Gaunt, D., Williams, R.A., Hoffman, D. and Hempen, G.L., (2017). "St. Louis area earthquake hazards mapping project: Seismic and liquefaction hazard maps." *Seismological Research Letters*, *88*(1), 206-223.

Cramer, C. H., Dhar, M. S., & Arellano, D. (2018). “Update of the Urban Seismic and Liquefaction Hazard Maps for Memphis and Shelby County, Tennessee: Liquefaction Probability Curves and 2015 Hazard Maps.” *Seismological Research Letters, 89*(2A), 688-701.

Cubrinovski, M., Rhodes, A., Ntritsos, N., and van Ballegooy, S., (2019). “System response of liquefiable deposits.” Soil Dynamics and Earthquake Engineering, 124, 212-229.

Davis, E., Atwater, B., Crider, J., and Garrison-Laney, C. (2019). “Seattle liquefaction features along the Duwamish waterway, Washington.” Seismological Society of America Annual Meeting, Seattle, WA.

Diaconis, P. and Efron, B. (1983). “Computer intensive methods in statistics.” Sci. Amer., 248(5), 116–130.

Fiegel, G. L., and Kutter, B. L. (1994). “Liquefaction mechanism for layered soils.” Journal of Geotechnical and Geoenvironmental Engineering, 120(4), 737-755.

Esri. (2020). "World Street Map" [basemap]. Scale Not Given. "World Street Map." https://services.arcgisonline.com/ArcGIS/rest/services/World_Street_Map/MapServer (March 25, 2020).

Geyin, M., and Maurer, B.W. (2019). “An analysis of liquefaction-induced free-field ground settlement using 1,000+ case-histories: observations vs. state-of-practice predictions.” GSP 308, 489-498.

Geyin, M. and Maurer, B.W. (2020). “Horizon: CPT-based liquefaction risk assessment and decision software.” DesignSafe-CI, doi: 10.17603/ds2-2fky-tm46.

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, DOI: 10.1177/8755293019899951.

Gathro, J. D., Bwambale, B., Andrus, R. D., & Heidari, T. (2018). “Liquefaction Probability Curves for Three Surficial Sand Deposits near Charleston, South Carolina.” Geotechnical Earthquake Engineering and Soil Dynamics V: Liquefaction Triggering, Consequences, and Mitigation, 374-383. Reston, VA.

Goda, K., Atkinson, G.M., Hunter, J.A., Crow, H. and Motazedian, D. (2011). “Probabilistic liquefaction hazard analysis for four Canadian cities.” Bull. of the Seismological Soc. of America, 101(1), 190-201.

Green, R.A., Cubrinovski, M., Cox, B., Wood, C., Wotherspoon, L., Bradley, B, and Maurer, B. (2014). “Select Liquefaction Case Histories from the 2010-2011 Canterbury Earthquake Sequence.” Earthquake Spectra, 30(1), 131-153.

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.

Green, R.A., Bommer, J., Stafford, P.J., Maurer, B.W., Kruiver, P.P., Edwards, B., Rodriguez-Marek, A., de Lange, G., Oates, S.J., Storck, T., Omidi, P., Bourne, S.J., van Elk, J.F. (2020). “Liquefaction hazard of the Groningen region of the Netherlands due to induced seismicity.” *Journal of Geotechnical and Geoenvironmental Engineering*, Accepted – In Press.

Hayati, H., and Andrus, R. D. (2008). “Liquefaction potential map of Charleston, South Carolina based on the 1886 earthquake.” Journal of Geotechnical and Geoenvironmental Engineering, 134(6), 815–828.

Holzer, T. L. (2008). “Probabilistic liquefaction hazard mapping.” Geotechnical Earthquake Engineering and Soil Dynamics IV, 1-32. Reston, VA: American Society of Civil Engineers.

Holzer, T.L., Noce, T.E., Bennett, M.J., Tinsley, J.C., III, and Rosenburg, L.I. (2005). “Liquefaction at Oceano, California, during the 2003 San Simeon earthquake.” BSSA, 95(6), 2396–2411.

Holzer, T. L., Bennett, M. J., Noce, T. E., Padovani, A., and Tinsley, J. C., III. (2006). “Liquefaction hazard mapping with LPI in the greater Oakland, California, area.” Earthquake Spectra, 22(3), 693–708.

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.” Proc., 11th International Conference on Soil Mechanics and Foundation Engineering, San Francisco, CA, USA, 1, 321-376.

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.

Iwasaki, T., Arakawa, T., Tokida, K. (1984). ‘Simplified procedures for assessing soil liquefaction during earthquakes.” Soil Dynamics and Earthquake Engineering, 3(1), 49-58.

Kramer, S.L., Mayfield, R.T. (2007). “Return Period of Soil Liquefaction.” Journal of Geotechnical and Geoenvironmental Engineering, 133(7): 802-813.

Kwak, D. Y., Stewart, J. P., Brandenberg, S. J., & Mikami, A. (2016). “Characterization of seismic levee fragility using field performance data.” EQS, 32(1), 193-215.

Land Information New Zealand. (2020). "Topographic Data" [basemap]. Scale Not Given. "Topographic Data." https://data.linz.govt.nz/data/category/topographic (March 25, 2020).

Lenz, A.J. and Baise, L.G. (2007). “Spatial variability of liquefaction potential in regional mapping using CPT and SPT data.” Soil Dynamics and Earthquake Engineering, 27(7), 690–702.

Luna, R. and Frost, J.D. (2000). “Treasure Island’s Spatial Liquefaction Evaluation.” GSP, 93, 306-320.

Lunne, T., Robertson, P.K., and Powell, J.M. (1997). Cone Penetration Testing in Geotechnical Practice. Blackie Academic & Professional, London, U.K.

Maurer, B.W., Green, R.A., Cubrinovski, M., and Bradley, B.A., (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, B.W., Green, R.A., Cubrinovski, M., and Bradley, B.A, (2014b). “Assessment of aging correction factors for liquefaction resistance at sites of recurrent liquefaction.” 10th U.S. National Conference on Earthquake Engineering, July 20-26, Anchorage, USA.

Maurer, B.W., Green, R.A., and Taylor, O.D.S. (2015a). “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., (2015b). “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., Cubrinovski, M., and Bradley, B. A. (2015c). “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, 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.

Moss, R.E.S. (2003). “CPT-based probabilistic assessment of seismic soil liquefaction initiation.” Doctor of Philosophy dissertation, Univ. of California, Berkeley, CA.

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.

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. National Research Council, The National Academies Press, Washington, DC.

NZGD (2019). "New Zealand Geotechnical Database." < https://www.nzgd.org.nz/>

Özener, P., Özaydin, K., and Berilgen, M. (2008). "Numerical and physical modeling of liquefaction mechanisms in layered sands." Geotechnical Earthquake Engineering and Soil Dynamics IV, 1-12.

Petersen, M. D., Frankel, A. D., Harmsen, S. C., Mueller, C. S., Haller, K. M., Wheeler, R. L., ... & Luco, N. (2008). "Documentation for the 2008 update of the United States national seismic hazard maps." United States Geological Survey.

Porter, K. (2019). A Beginner's Guide to Fragility, Vulnerability, and Risk. University of Colorado Boulder, 119 pp., http://spot.colorado.edu/~porterka/Porter-beginnersguide.pdf

Rashidian, V., and Gillins, D.T. (2018). "Modification of the liquefaction potential index to consider the topography in Christchurch, New Zealand." Engineering Geology 232, 68-81.

Rix, G.J. and Romero, S. (2001). "Liquefaction Susceptibility Mapping in Memphis/Shelby County, TN" United States Geological Survey Final Technical Report, Award No. 01-HQ-AG-0019.

Robertson, P.K., and Wride, C.E. (1998). "Evaluating cyclic liquefaction potential using cone penetration test." Canadian Geotechnical Journal, 35(3), 442-459.

Robertson, P.K. (2011). "Automated detection of CPT transition zones." Geotechnical News, June 2011, 35-38.

Toprak, S., and Holzer, T. (2003). “Liquefaction potential index: field assessment.” Journal of Geotechnical and Geoenvironmental Engineering, 129(4), 315-322.

Treadwell, D.D. (1976). The influence of gravity, prestress, compressibility, and layering on soil resistance to static penetration. Doctor of Philosophy dissertation, Univ. of California, Berkeley, CA.

USGS (2019). “Unified Hazard Tool.” < https://earthquake.usgs.gov/hazards/interactive/>

van der Linden, T.I., De Lange, D.A. and Korff, M. (2018). “Cone penetration testing in thinly inter-layered soils.” Proceedings of the Institution of Civil Engineers-Geotechnical Engineering, 171(3), 215-231.

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. (2014a). “Assessment of liquefaction-induced land damage for residential Christchurch.” Earthquake Spectra, 30(1), 31-55.

van Ballegooy, S., Cox S.C., Thurlow C., Rutter H.K., Reynolds, T., Harrington, G., Fraser, J., and Smith, T. (2014b). “Median water table elevation in Christchurch and surrounding area after the 4 September 2010 Darfield earthquake: Version 2.” GNS Science Report 2014/18, 2014b.

van Ballegooy, S., Green, R.A., Lees, J., Wentz, F., and Maurer, B.W. (2015). “Assessment of various CPT based liquefaction severity index frameworks relative to the Ishihara (1985) H1-H2 boundary curves.” Soil Dynamics and Earthquake Engineering, 79, 347-364.

Washington State Department of Natural Resources (2019). “Washington Geologic Information Portal.” https://geologyportal.dnr.wa.gov/

Yost, K. M., Cox, B. R., Wotherspoon, L., Boulanger, R. W., van Ballegooy, S., and Cubrinovski, M. (2019). “In situ investigation of false-positive liquefaction sites in Christchurch, New Zealand:

Palinurus Road Case History." Geotechnical Special Publication 308, Meehan et al. (eds), ASCE, 436-451.

Zhang, G., Robertson, P.K., Brachman, R. (2002). "Estimating Liquefaction Induced Ground Settlements from CPT." Canadian Geotechnical Journal, 39, 1168-1180.

APPENDIX E

Publication 5 (Conference Paper 1), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology*, In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering*.	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

AN ANALYSIS OF LIQUEFACTION-INDUCED FREE-FIELD GROUND SETTLEMENT USING 1,000+ CASE-HISTORIES: OBSERVATIONS VS. STATE-OF-PRACTICE PREDICTIONS

Mertcan Geyin, S.M.ASCE[1] and Brett W. Maurer, A.M.ASCE[2]

[1] Graduate Research Assistant, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195; e-mail: mgeyin@uw.edu

[2]Assistant Professor, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195; e-mail: bwmaurer@uw.edu

Abstract

Liquefaction-induced free-field settlements can be attributed to several mechanisms, including volumetric reconsolidation strain, ground loss due to ejecta, and vertical components of lateral spreads and flow slides. In practice, the volumetric-strain settlement is typically predicted using one of several semi-empirical methods, but there is no consensus of which, if any, is most accurate. Moreover, no method has been proposed to predict the magnitude of settlement owed to ejecta, even though this component could be larger than any other. The lack of a complete and trusted model for predicting free-field settlement is due, in part, to the scarcity of case-history data needed to apportion and model its individual components. However, the 2010-2011 Canterbury, New Zealand, earthquake sequence (CES) resulted in a liquefaction dataset of unprecedented size and quality, presenting a unique opportunity to assess and advance liquefaction analytics. Accordingly, this study compiles and analyzes 1,013 high-quality settlement case-histories from the CES, such that the contributing components of settlement can be reasonably accounted for and quantified. Presented herein are preliminary findings of this effort, summarized as follows: (1) popular CPT-based models for

predicting volumetric-strain settlement exhibit heteroskedastic bias, such that smaller total settlements (< ~6 cm) are underpredicted while larger total settlements (> ~6 cm) are overpredicted; (2) these poor predictions cannot be explained by the uncertainty of CPT proxies of relative density, which are central to predictions of volumetric strain; (3) introduction of a depth-weighting function lessens the heteroskedastic prediction bias, showing promise for future study; and (4) prediction-performance is related to the observed volume of liquefaction ejecta, illustrating the need to account for and model this mechanism of settlement. Collectively, the findings suggest popular methods for predicting free-field settlement have significant room for improvement.

INTRODUCTION

Liquefaction-induced free-field settlements are attributable to several sources, including volumetric strain (ε_v) associated with reconsolidation of liquefied soil; ground loss due to liquefied and fluidized soil venting onto the surface (i.e., liquefaction ejecta); and vertical components of deformation resulting from lateral spreads and flow slides. Such settlements were vividly displayed during the 2010-2011 Canterbury, New Zealand, earthquake sequence (CES), which resulted in pervasive and often-severe liquefaction in the city of Christchurch. As a result, more than 3.5 km^2 of land experienced settlements exceeding one meter in a single earthquake. Beyond the immediate, widespread damage to infrastructure, this subsidence also greatly increased the risk of flooding on Christchurch's coastal-alluvial plains, a consequence since realized repeatedly.

In practice, the settlement owed to ε_v is often predicted using one of several semi-empirical methods, but there is no consensus of which is most accurate. Moreover, no method - simple or otherwise - has been proposed to predict the magnitude of settlement owed to ejecta, even though this

component could be larger than any other. As a result, the predicted ε_v settlement is often implicitly treated as a "catch-all" prediction, despite the knowledge that, if accurate, it could significantly underpredict total settlement. The lack of a complete and trusted model for predicting free-field settlement is due, in part, to the scarcity of case-history data needed to apportion and model its individual components. However, the CES resulted in a liquefaction dataset of unprecedented size and quality, presenting a unique opportunity to assess and advance liquefaction analytics. In this regard, one prior study of CES ground settlements (Tonkin and Taylor, 2013) found no apparent relationship between the observed total settlement and that expected from volumetric strain. However, that study accounted neither for horizontal ground movement nor for liquefaction ejecta, both of which were pervasive in Christchurch, and each of which could lead to large discrepancies between the observed settlement and that suggested by volumetric-strain alone. Also, complicating factors result in some CES events having relatively larger uncertainties that affect both the measured and predicted settlements, clouding subsequent analyses.

Accordingly, this study compiles and analyzes 1,013 settlement case-histories from the M_w6.2, 22 Feb 2011 Christchurch earthquake, such that the contributing components of settlement can be reasonably accounted for and quantified. This is accomplished by analyzing measurements of both horizontal and vertical ground displacements, and by classifying the volume of ejecta at each case-history location. In addition, the cases studied herein represent only the highest quality data available from the CES. Presented herein are preliminary results of this ongoing effort. Specifically, we: (1) compare observed settlements to those predicted by a popular cone-penetration-test (CPT) method (which considers only ε_v); (2) determine whether predictions could be improved by regional calibration of CPT proxies for relative density (D_r), which are central to predictions of ε_v; (3) assess the use of a depth-weighting function for predicting settlement, which in current practice is typically

not used; and (4) evaluate the need to account for and model the settlement due to liquefaction ejecta. In the following, the methods that will be used to predict ε_v settlement are first discussed. The case-studies to be analyzed are then summarized, followed by analyses of observed vs. predicted liquefaction-induced free-field settlements.

Settlement Due to Post-Liquefaction Volumetric Strain (ε_v)

Current CPT-based methods for estimating ε_v are generally rooted in the work of Nagase and Ishihara (1988), who conducted cyclic simple shear tests on saturated sands of varying D_r. Based largely on these test results, Ishihara and Yoshimine (1992) proposed a series of curves, shown in Figure 1a, for estimating ε_v as a function of D_r and the factor of safety against liquefaction triggering (FS_{liq}). Using a correlation proposed by Tatsuoka et al. (1990) to convert from D_r to normalized CPT tip resistance (q_{c1N}), Zhang et al. (2002) converted the Ishihara and Yoshimine (1992) curves into a form suitable for use with CPT-based procedures for computing FS_{liq}, as shown in Figure 1b. However, to better capture trends identified by Ishihara and Yoshimine (1992), additional curves were developed by Maurer et al. (2015a) for FS_{liq} = 0.95, FS_{liq} = 0.96 … FS_{liq} = 0.99 to supplement those of Zhang et al. (2002). As evident in Figure 1b, these curves are needed to accurately represent the behavior identified by Ishihara and Yoshimine (1992), whereas linear interpolations between the FS_{liq} = 0.9 and FS_{liq} = 1 curves often fail to do so. It is also noteworthy that the Zhang et al. (2002) curves were developed using a correlation to q_{c1N} (i.e. normalized tip resistance, *uncorrected* for fines), but are proposed for use with values of q_{c1Ncs}. This simplification, justifications for which are given by Zhang et al. (2002) and Idriss and Boulanger (2008), is used throughout this study. Once a profile of depth

vs. expected ε_v is developed, ground-surface settlement is predicted by integrating over depth (Zhang et al., 2002):

$$S = \sum_{i=1}^{j} \epsilon_{vi} \ \Delta z_i \tag{1}$$

Where S is free-field ground settlement; ε_{vi} is post-liquefaction volumetric strain for sublayer i, Δz_i is the thickness of the sublayer i; and j is the number of soil sublayers. Since its inception, the Zhang et al. (2002) method has been widely used and is arguably the most popular CPT-based approach in engineering practice. It will be evaluated as part of the present study.

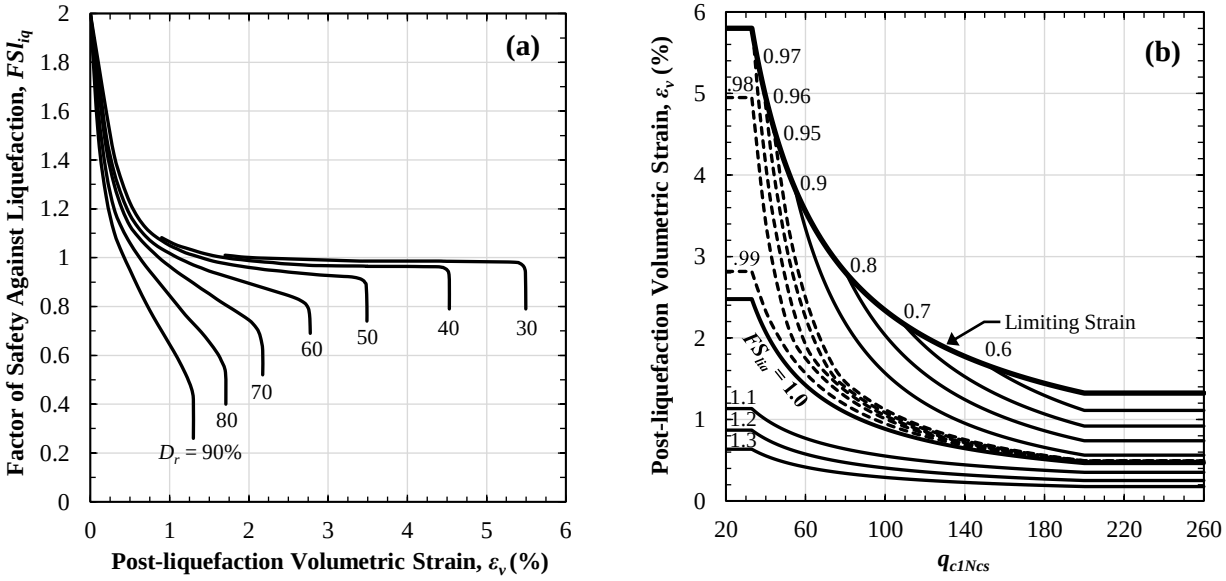


Figure 1. (a) Curves for estimating post-liquefaction volumetric-strain (ε_v) as a function of the factor of safety against liquefaction (FS_{liq}) and initial relative density (D_r) (after Ishihara and Yoshimine 1992); **(b)** curves derived from (a) for estimating ε_v as a function of FS_{liq} and equivalent clean sand normalized CPT tip resistance (q_{c1Ncs}), as proposed by Zhang et al. (2002), with additional curves developed by Maurer et al. (2015a).

Notably, while the Zhang et al. (2002) method is widely used in practice, it is rigidly tied to the Tatsuoka et al. (1990) D_r - q_{c1n} correlation for Toyoura sand:

$$D_r\,(\%) = -85 + 76\log(q_{c1N}) \tag{2}$$

Thus, if Equation (2) is inaccurate for soils other than Toyoura (Japan) sand, the ε_v estimates from Zhang et al. (2002) could be erroneous. Alternative D_r - q_{c1n} correlations include that of Idriss and Boulanger (2003), adapted from Salgado et al. (1997):

$$D_r = 0.465\left(\frac{q_{c1N}}{C}\right)^{0.264} - 1.063 \tag{3}$$

where C is a soil-specific constant. Idriss and Boulanger (2003) used C = 0.9 in analyzing global liquefaction case studies, whereas C ranged from 0.64 to 1.55 for soils studied by Salgado et al. (1997). Plotted in Figure 2 are Equations (2) and (3), wherein 3 values of C are adopted for the latter (C = 0.64, 0.9, 1.55). It can be seen that large differences exist among the selected correlations, which could lead to different estimates of ε_v, and thus, different predictions of settlement. Accordingly, to account for the uncertain applicability of the Tatsuoka et al. (1990) correlation to soils affected by the CES, three variants of Zhang et al. (2002) will be developed using: (1) the Salgado et al. (1997) correlation with C values of 0.64, 0.9, and 1.55; and (2) the Yoshimine et al. (2006) solution, which approximates the Ishihara and Yoshimine (1992) curves in Figure 1a. The Yoshimine et al. (2006) solution has the advantage of being a continuous function, whereas Zhang et al. (2002) uses discrete functions requiring interpolation. This study will evaluate both the original Zhang et al. (2002) method and the three aforementioned variants.

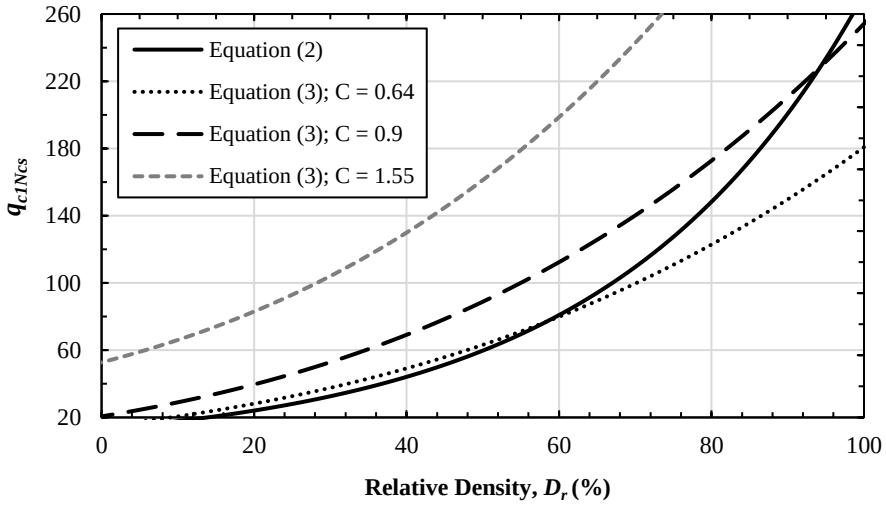


Figure 2. Correlations relating equivalent clean sand normalized CPT tip resistance (q_{c1Ncs}) to relative density (D_r).

Data and Methodology

Ground Deformation Data

Vertical and horizontal ground deformation data was obtained from the New Zealand Geotechnical Database (NZGD, 2012) for the M_w6.2, 22 Feb 2011 Christchurch earthquake. These deformations were measured via airborne LIDAR following the earthquake and include local liquefaction-induced movements as well as regional tectonic movements associated with the event. To isolate liquefaction-induced deformations, tectonic movements (NZGD, 2012) were first subtracted from LIDAR measurements. While airborne LIDAR enabled deformations to be efficiently measured across a very large area, it comes with the trade-off of a +/- 7 cm measurement accuracy (NZGD, 2015). Due to this uncertainty, many negative settlements (i.e., inferred ground heave) are observed across the study area, which may not reflect actual conditions. Despite this relatively large measurement uncertainty,

the LIDAR data allows for significantly more settlement case-histories to be compiled than in all previous earthquakes combined.

While settlement data is available from four major CES events – the 4 Sept 2010 M_w7.1, 22 Feb 2011 M_w6.2, 13 June 2011 M_w5.9, and 23 Dec 2011, M_w5.9 earthquakes - we choose to focus only on the February 2011 earthquake for the following reasons. *First*, the 13 June and 23 December data are complicated by the fact that multiple, similar-magnitude events occurred only minutes-to-hours apart. As a result, pore pressures were elevated at the time of latter events (complicating predictions) and LIDAR data captures the effects of multiple earthquakes (complicating measurements). *Second*, because pre-CES reference LIDAR is of lower quality, the measurement uncertainty of settlements induced by the 4 September event (15 cm) is twice that of subsequent events (7 cm). Therefore, the settlement case-histories compiled from the February 2011 event represent the highest quality data available from the CES.

CPT Soundings

This study utilizes CPTs performed at sites where (1) the liquefaction-induced ground deformation was measured, as described above; and (2) the severity of liquefaction ejecta was reliably documented, as described subsequently. In compiling these CPT case-histories, soundings were first rejected: (1) if the depth of "pre-drill" exceeded the depth to ground water; (2) if believed to have prematurely terminated on shallow gravels (see Maurer et al., 2014); and (3) if more than 20 cm of horizontal ground movement was measured at the CPT location. This last criterion was applied to remove artificially large settlements caused or augmented by lateral spreading. In other words, we aim to study liquefaction-induced deformations that are purely vertical in nature. At the end of this process, a total of 1,013 settlement case-histories were compiled.

Liquefaction Severity

Observations of liquefaction and the severity of ejecta were made by the authors for each CPT location following the Christchurch earthquake. CPT sites were assigned one of four classifications (Table 1), as described in Green et al. (2014). Of the 1,013 cases compiled, 19.6% are cases of "no ejecta" and 80.4% are cases where ejecta was observed and classified.

Table 4. Criteria used to classify the severity of liquefaction ejecta in this study.

Classification	Severity ID	Criteria
No Ejecta	0	No observed liquefaction ejecta or lateral spreading
Minor Ejecta	1	Small, isolated liquefaction features less than a vehicle width; <5% of ground surface is covered by ejecta; no lateral spreading.
Moderate Ejecta	2	Groups of liquefaction features greater than a vehicle width; 5-40% of ground surface is covered by ejecta; streets are generally passable; no lateral spreading.
Severe Ejecta	3	Adjoining large liquefaction features that are greater than a vehicle width; >40% of ground surface is covered by ejecta; streets are generally impassable; no lateral spreading.

Estimation of Ground Water Depth (GWT) and Peak Ground Acceleration (PGA)

GWT depths were sourced from the event-specific regional ground water models of van Ballegooy et al. (2014). These models, which reflect seasonal and local fluctuations across the region, were derived in part using monitoring data from ~1000 piezometers and provide a best-estimate of GWT depths immediately prior to the Christchurch earthquake.

To evaluate FS_{liq} for use in computing $\varepsilon_{v,}$ peak ground accelerations (*PGAs*) were computed using the Bradley (2013) model, which has been used by many studies pertaining to the CES (e.g.,

Maurer et al., 2015b, 2015c, 2017; van Ballegooy et al., 2015). This procedure combines unconditional *PGA* distributions estimated by a ground motion prediction equation, recorded *PGAs* from strong motion stations, and the spatial correlation of intra-event residuals to compute *PGAs* at sites of interest.

Liquefaction Triggering Model and Volumetric Strain (εv) Settlement

FS_{liq} values were computed using the deterministic procedure proposed by Boulanger and Idriss (2014) wherein the soil behavior index (I_c) proposed by Robertson and Wride (1998) was used to identify soils not susceptible to liquefaction; soils with $I_c > 2.5$ were assumed non-liquefiable, per Maurer et al. (2018). Within the Boulanger and Idriss (2014) procedure, fines content (*FC*) is required to compute normalized tip resistances. Accordingly, *FC* values were estimated using an I_c-*FC* correlation developed specifically for soils in Christchurch (Maurer et al., 2018).

As previously discussed, ε_v settlement was separately predicted using the Zhang et al. (2002) method along with three variants of Zhang et al. (2002) that utilize the Yoshimine et al. (2006) approximate solution in conjunction with different D_r - q_{c1Ncs} correlations (Salgado et al., 1997) wherein soil-specific constants of C = 0.64, 0.9, and 1.55 were used.

Results and Discussion

Performance of Existing Settlement Prediction Methods

Plotted in Figure 3 are prediction-residuals vs. predicted settlement for each of the 1,013 case histories. The prediction residual is defined by:

$$Residual = Observed\ Settlement - Predicted\ Settlement \quad (4)$$

As such, positive residuals indicate an underprediction of observed settlement whereas negative residuals indicate an overprediction of observed settlement. As seen in Figure 3a, the Zhang et al. (2002) model for predicting ε_v settlement exhibits heteroskedastic bias, such that smaller total settlements (< ~6 cm) are slightly underpredicted while larger total settlements (> ~6 cm) are significantly overpredicted. Consequently, while the mean residual is near-zero (-5.13 cm), 33.7% of residuals exceed ±10 cm, and 9.3% exceed ±20 cm. One possible cause of these poor predictions is the Tatsuoka et al. (1990) D_r - q_{c1ncs} correlation inherent to Zhang et al. (2002). This correlation could be inaccurate for soils affected by the CES, resulting in erroneous estimates of D_r, and by corollary, erroneous estimates of ε_v. However, as seen in Figures 3b-3d, the use of alternative D_r - q_{c1ncs} correlations provides little-to-no improvement – each variant results in a similar heteroskedastic prediction bias. Thus, while the possibility cannot be eliminated, it is unlikely the poor performance of Zhang et al. (2002) is a result of inaccurate CPT-derived estimates of in-situ D_r. In Figures 3a-3d, the degree of scatter is quantified using the normalized root-mean square error (NRMSE), defined as:

$$NRMSE = \frac{\sqrt{\Sigma e_i^2 / N}}{p_{i-max} - p_{i-min}} \quad (5)$$

Where e is the residual error, as defined in Equation 4; p is the predicted value; the subscript i is the i^{th} case history; and N is the total number of cases.

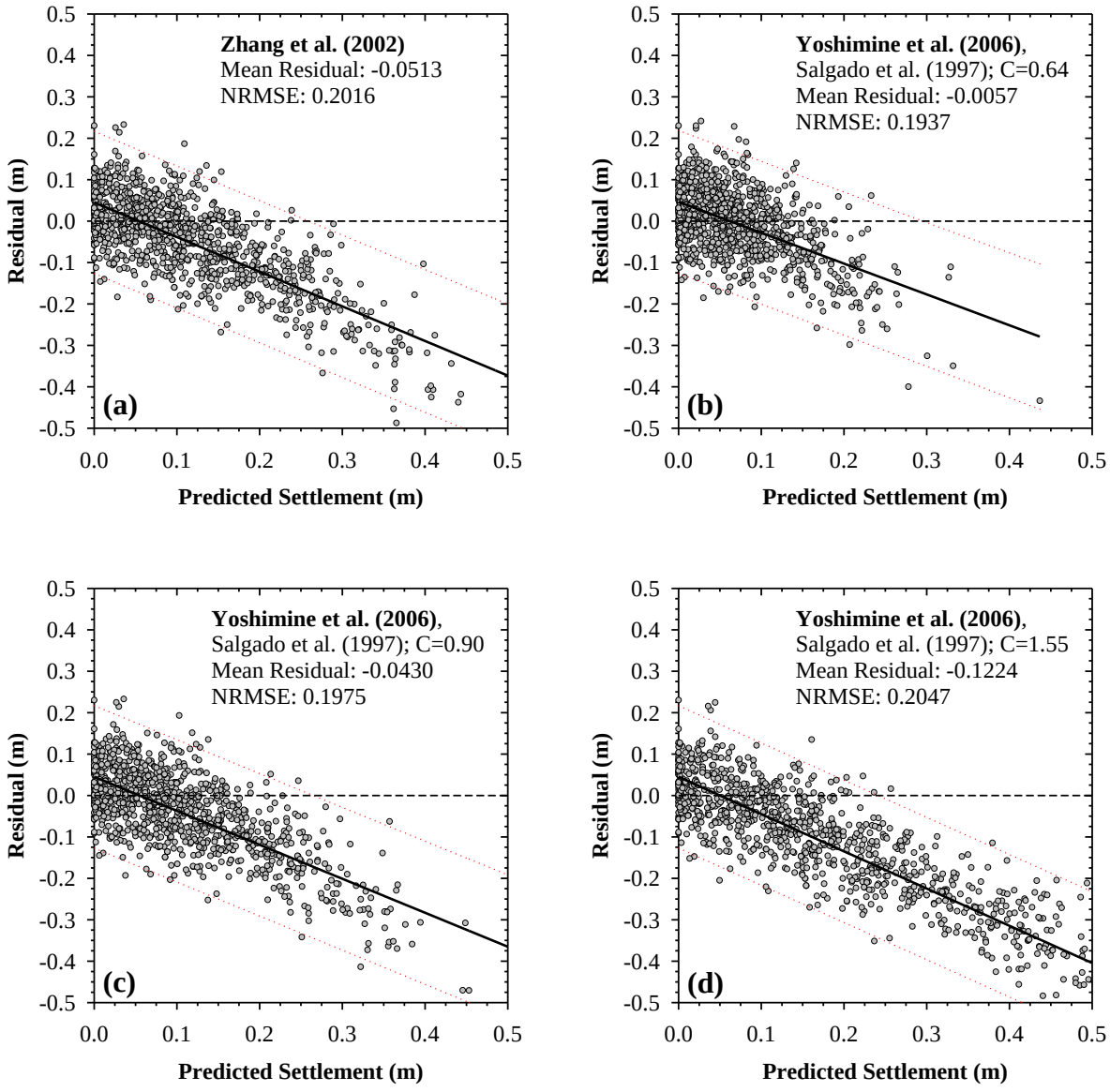


Figure 3. Predicted settlements vs. residual errors with linear-robust regression line and 95% prediction bands: **(a)** original Zhang et al. (2002) method and **(b, c, d)** Yoshimine et al. (2006) with different *C* coefficients using correlation provided in Equation 3.

Introduction of a Depth Weighting Function

In contrast to the liquefaction potential index (LPI) and other similar parameters, depth-weighting functions have traditionally not been used to predict free-field settlement. In other words, settlement is assumed independent of the depth(s) at which liquefaction occurs. However, there are several reasons why this assumption could be unreasonable (Cetin et al., 2009), including lateral heterogeneity of subsurface stratigraphy and soil-arching effects. Accordingly, with the goal of improving the performance shown in Figure 3, a linear depth-weighting scheme was introduced:

$$S = \sum_{i=1}^{j} \epsilon_{vi} \ (1 - 0.1 z_i) \ \Delta z_i \qquad (6)$$

Where S is free-field ground settlement; subscript i is the sublayer index; ε_{vi} is post-liquefaction volumetric strain; z_i is the sublayer depth; Δz_i is the sublayer thickness; and j is the total number of soil sublayers. Using Eq. (6), predicted volumetric-strains at depths less than 10 m were down-scaled with increasing depth, and those at depths exceeding 10 m were assumed not to contribute to surface settlement. Representative results of this application are shown in Figure 4a. While the degree of scatter is increased (compare to Figure 3c), the heteroskedastic bias is significantly diminished, showing promise for more rigorous investigation and model development.

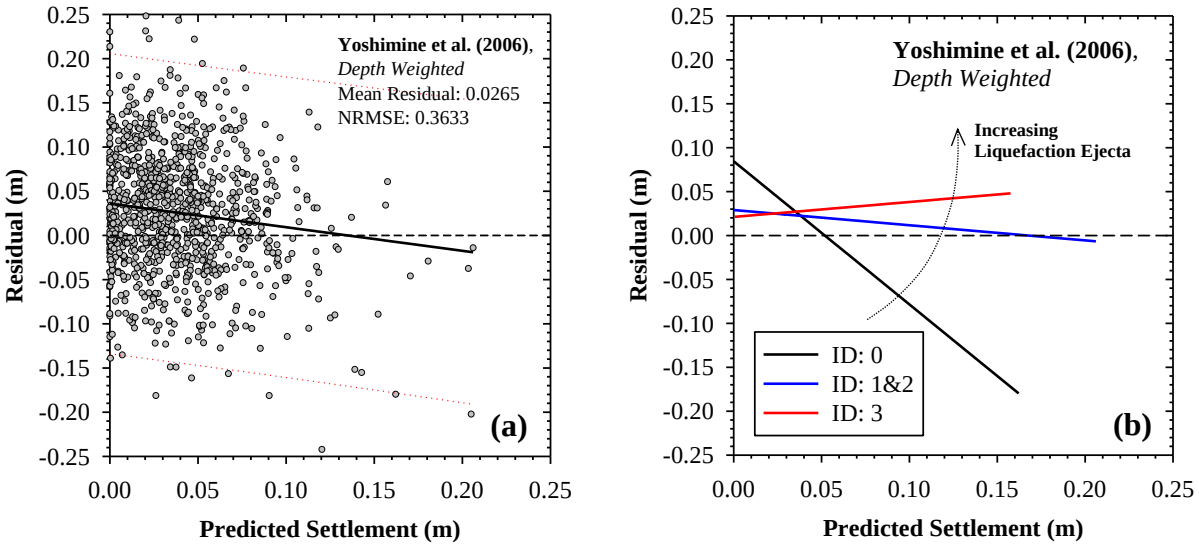


Figure 4. Settlement predictions with depth-weighting (Yoshimine et al. (2006) method with Salgado et al. (1997) *Dr* - *qc1Ncs* correlation; C = 0.90): **(a)** predicted settlement vs. residual errors with linear-robust regression line and 95% prediction bands; **(b)** residual trendlines grouped by observed severity of liquefaction ejecta, as classified using Table 1.

Observed Severity of Liquefaction Ejecta and Predicted Settlements

None of the prediction models evaluated thus far directly account for the contribution of liquefaction ejecta to ground settlement. To evaluate the need to account for and model this component of settlement, predictions were grouped by the observed severity of liquefaction ejecta, as classified using the criteria in Table 1. Representative results of this analysis are shown in Figure 4b. It can be seen that as the observed volume of ejecta increases, residuals tend to decrease, especially for larger predicted settlements. In other words, overpredictions of observed settlement diminish as the volume of ejecta increases. Thus, the general tendency for existing ε_v prediction methods to overpredict total settlement (e.g., due to the lack of depth-weighting schemes) is to some degree counteracted by the fact these methods do not account for ejecta, the omission of which should result in underprediction of total settlement. As shown in Figure 4b, prediction-performance is clearly related to the observed volume of liquefaction ejecta, illustrating the need to account for and model this mechanism of settlement in addition to that resulting from ε_v.

Conclusion

This study compiled and analyzed 1,013 settlement case-histories from the M_w6.2, 22 Feb 2011 Christchurch earthquake, such that the contributing components of settlement could be reasonably accounted for and quantified. Preliminary findings of this effort were presented and are summarized

as follows: (1) popular CPT-based models for predicting volumetric-strain settlement exhibit heteroskedastic bias, such that smaller total settlements (< ~5 cm) are underpredicted while larger total settlements (> ~5 cm) are overpredicted; (2) these poor predictions cannot be explained by the uncertainty of CPT proxies of relative density, which are central to predictions of volumetric strain; (3) introduction of a depth-weighting function lessens the heteroskedastic prediction bias, showing promise for future study; and (4) prediction-performance is related to the observed volume of liquefaction ejecta, illustrating the need to account for and model this mechanism of settlement. Collectively, the findings suggest that popular methods for predicting free-field settlement have significant room for improvement.

References

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, Civil & Environmental Eng., UC Davis, CA.

Bradley, B.A. (2013). "Site-specific and spatially-distributed ground motion intensity estimation in the 2010-2011 Christchurch earthquakes." *Soil Dynamics and Earthquake Engineering*, 48, 35-47.

Cetin, K.O., Bilge, H.T., Wu, J, Kammerer, A.M., and Seed, R.B. (2009). "Probabilistic model for the assessment of cyclically induced reconsolidation (Volumetric) Settlements." *Journal of Geotechnical and Geoenvironmental Engineering* 135(3): 387-398.

Green, R.A., Cubrinovski, M., Cox, B., Wood, C., Wotherspoon, L., Bradley, B, Maurer, B. (2014). "Select Liquefaction Case Histories from the 2010-2011 Canterbury Earthquake Sequence." *Earthquake Spectra*, 30(1): 131-153.

Idriss IM and Boulanger RW. (2003). "Relating K_α and K_σ to SPT blow count and to CPT tip

resistance for use in evaluating liquefaction potential." *Dam Safety Conference*, Minneapolis, MN, USA.

Idriss, I.M. and Boulanger, R.W. (2008). *Soil liquefaction during earthquakes*. Monograph MNO-12, EERI, Oakland, CA, 261 pp.

Maurer, B.W., Green, R.A., Cubrinovski, M., and Bradley, B.A. (2014). "Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand." *Journal of Geotechnical and Geoenvironmental Engineering* 140(7), 04014032.

Maurer, B.W., Green, R.A., Cubrinovski, M., and Bradley, B.A. (2015a). "Investigating the influence of post-liquefaction strain potential on the accuracy of Liquefaction Severity Number (LSN) hazard assessments." *10th Pacific Conference on Earthquake Engineering*, Nov 6-8; Sydney, Australia; Paper No. 796.

Maurer, B.W., Green, R.A., Cubrinovski, M., Bradley, B.A. (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, B.W., Green, R.A., Cubrinovski, M., and Bradley, B.A. (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., van Ballegooy, S., Bradley, B.A., and Upadhyaya, S. (2017). "Performance comparison of probabilistic and deterministic liquefaction triggering models for hazard assessment in 23 global earthquakes." *Geo-Risk 2017: Reliability-based design and code developments* (J. Huang, G.A. Fenton, L. Zhang, and D.V. Griffiths, eds.), Geotechnical Special Publication 283: 31-42.

Maurer, B.W., Green, R.A., van Ballegooy, S., and Wotherspoon, L. (2018). "Development of region-specific soil behavior type index correlations for evaluating liquefaction hazard in Christchurch, New Zealand." *Soil Dynamics and Earthquake Engineering, In Press.*

Nagase, H. and Ishihara, K. (1988). Liquefaction-induced compaction and settlement of sand during earthquakes, *Soils and Foundations*, 28(1):66–76.

NZGD (New Zealand Geotechnical Database). (2012). "Vertical Ground Surface Movements." Map Layer CGD0600 - 23 July 2012, <https://www.nzgd.org.nz/>, (Jan 05, 2018).

NZGD (New Zealand Geotechnical Database). (2015). "LiDAR and Digital Elevation Models." Map Layer CGD0500 - 30 June 2015, <https://www.nzgd.org.nz/>, (Jan 05, 2018).

Robertson, P.K. & Wride, C.E. (1998). "Evaluating cyclic liquefaction potential using the cone penetration test." *Canadian Geotechnical Journal*, 35(3), 442-459.

Salgado R, Mitchell JK, Jamiolkowski, M. (1997). "Cavity expansion and penetration resistance in sands." *J Geotech Geoenviron Eng*, 123(4): 344-354.

Tatsuoka, F., Zhou, S., Sato, T., and Shibuya, S. (1990). "Evaluation method of liquefaction potential and its application." *Report on seismic hazards on the ground in urban areas*. Ministry of Education of Japan, Tokyo, Japan.

Tonkin & Taylor (2013). "Liquefaction vulnerability study." *T&T Ref: 52020.0200/v.1.0*, Christchurch, NZ.

van Ballegooy S, Cox SC, Thurlow C, Rutter HK, Reynolds T, Harrington G, Fraser J, and Smith T. (2014). "Median water table elevation in Christchurch and surrounding area after the 4 September 2010 Darfield earthquake: Version 2." *GNS Science Report 2014/18*, 2014b.

Van Ballegooy, S., Green, R.A., Lees, J., Wentz, F., and Maurer, B.W. (2015). “Assessment of various CPT based liquefaction severity index frameworks relative to the Ishihara (1985) H_1-H_2 boundary curves.” *Soil Dynamics and Earthquake Engineering,* 79, 347-364.

Yoshimine M, Nishizaki H, Amano K, Hosono Y. (2006). “Flow deformation of liquefied sands under constant shear load and its application to analysis of flow slide in infinite slope.” *Soil Dynamics and Earthquake Engineering*, 26: 253-264.

Zhang, G., Robertson, P.K., and Brachman, R. (2002). “Estimating Liquefaction Induced Ground Settlements from CPT.” *Canadian Geotechnical Journal,* 39: 1168-1180.

APPENDIX F

Publication 6 (Journal Paper 4), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology,* In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

EVALUATION OF A CONE PENETRATION TEST THIN-LAYER CORRECTION PROCEDURE IN THE CONTEXT OF GLOBAL LIQUEFACTION MODEL PERFORMANCE

Mertcan Geyin[3] and Brett W. Maurer[1]

Abstract

Engineering geologists routinely perform liquefaction hazard assessments using data from the cone penetration test (CPT). However, the volume of soil mobilized by the CPT acts as a low-pass filter on the true stratigraphy, potentially removing information such as the data defining a thin layer of soil or the interface between two dissimilar soils. The Boulanger and DeJong (2018) CPT inversion procedure, which aims to correct these effects, is herein evaluated in the context of CPT-based liquefaction model performance. Using over 15,000 case-histories from 24 earthquakes parsed into 2 datasets, 18 different liquefaction models are studied, resulting in 36 performance trials. In 1 of these trials, the CPT inversion procedure increases model efficiency to a statistically significant degree, but in 23 others it significantly decreases efficiency. This decline in performance tends to grow as profiles become more stratified. To explore remedies, a liquefaction triggering curve is rederived from inverted CPT data, such that its training and forward implementation are made consistent. Nonetheless, this exacerbates the decline in prediction efficiency. Ultimately, the results of this study are not a direct assessment of the pioneering Boulanger and DeJong (2018) procedure. However, the results do provide evidence that this procedure – when applied to existing CPT-based liquefaction models – may provide no demonstrable performance benefit.

[3] Department of Civil and Environmental Engineering, University of Washington, Seattle USA.

Keywords: cone penetration test; thin layer corrections; liquefaction model

Introduction

Cone Penetration Test (CPT) measurements, which principally include cone-tip resistance (q_c) and sleeve friction (f_s), may be used to infer various soil properties and behaviors. The CPT, for example, has significant benefits over other tests on which liquefaction models have been based (NRC 2016). In this regard, engineering geologists routinely carry out liquefaction hazard studies, both locally and regionally, using CPT data (e.g., among many, Juang et al. 2009; Heidari and Andrus, 2010; Khoshnevisan et al., 2015; Zhang et al. 2016, 2018; Chen et al., 2016; Gheibi et al., 2016; Hasek and Gassman, 2019; Bastin et al., 2020; Norini et al. 2021). However, because the CPT is an intermediate-to-large-strain penetration test, its measurements are still potentially obfuscated by the volume of mobilized soil. That is, q_c and f_s are influenced by soil conditions both relatively proximal to, and distal from, the CPT sensors. While CPT measurements are recorded at discrete depths – typically 1-2 cm intervals – they sample a mobilized zone of influence that may extend ~10-30 cone diameters from the q_c sensor (Boulanger and DeJong 2018), with the zone growing and shrinking as a function of soil properties. For the industry standard 10 cm^2 cone, this equates to a zone 0.35-1.05 m thick. This mobilized zone effectively acts as a "low-pass filter" on the true soil stratigraphy, filtering information from the low spatial wavelengths, such as the data defining a thin layer of soil or the interface between two dissimilar soils. These spatial smoothing effects, which are often called "thin layer" and "transition" effects, have been investigated by many authors (e.g., van der Linden 2018; Ching et al. 2015; Robertson 2011; Ahmadi and Robertson 2005; Lunne et al. 1997; Treadwell 1976). Although chart-based methods have been proposed to manually correct CPT data for these effects, Boulanger and DeJong (2018) proposed what may be the first fully automated approach. This

"inverse filtering and interface detection" procedure attempts to correct CPT measurements to their "true" values. Since the direct measurements reflect conditions averaged over a volume, rather than at discrete points, their correction would invariably improve CPT site characterization. As an example, CPT data is shown in Fig. 1, with and without correction by the Boulanger and DeJong (2018) procedure.

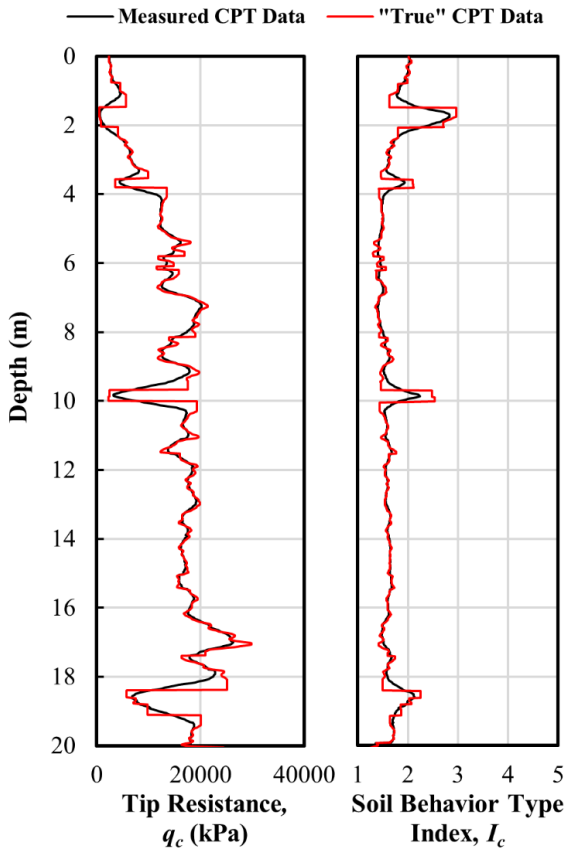


Fig. 1. CPT data with (i.e. "true") and without (i.e. measured) inverse-filtering and interface correction per Boulanger and DeJong (2018), as implemented in the software *Horizon* (Geyin and Maurer 2020a).

The efficacy of the Boulanger and DeJong (2018) procedure has yet to be rigorously studied in the literature, but its adopted can change the computed liquefaction hazard at a site, with the magnitude

and direction of change dependent on many factors. Moreover, this procedure was recently incorporated into software programs that predict liquefaction and its consequences (e.g., *CLiq* by Geologismiki 2020; *Horizon* by Geyin and Maurer 2020a). Accordingly, the objective of this paper is to assess the Boulanger and DeJong (2018) procedure in the context of CPT-based liquefaction model performance. That is, to determine whether use of the procedure improves prediction-model efficiency. Using field case histories from 24 earthquakes parsed into 2 datasets, 18 different liquefaction models will be studied, resulting in 36 performance trials. Uncertainty due to finite sampling will be accounted for and used to establish statistical significance. Ultimately, the results of these trials spur further inquiries that expand on the initial objective, and which are introduced later in the paper.

In the following, the Boulanger and DeJong (2018) procedure is first succinctly summarized. The 18 liquefaction models in which this procedure will be evaluated, and the field case-history data to which it will be applied, are then identified. Lastly, the methodologies for evaluating predictive performance are described and trial results are presented and discussed.

Summary of the Boulanger and DeJong (2018) Procedure

The correction of CPT data can be idealized as the conversion of measured q_c and f_s values, or q^m and f^m, to the "true" values q^t and f^t, which would be obtained if measurements reflected conditions at discrete points. This can be viewed as an inverse problem, with the goal of determining the "true" values via inversion of what was measured. As illustrated in Fig. 2 (in this case for tip resistance), q^m may deviate from q^t near layer interfaces because the measurements are influenced by materials both above and below the interface, even though the measurement is recorded at a discrete depth. In the case shown in Fig. 2, q^m is artificially increased in the upper soft clay as the cone approaches, and is influenced by, the underlying dense sand. The opposite occurs in the dense sand as the cone

approaches the underlying soft clay. That is, q^m is artificially decreased when the cone begins to sense the softer material. These errors are commonly referred to as "transistion" or "interface" effects. It can also be seen in Figure 2 that should the dense sand be insufficiently thick for the cone to receive no influence, at least momentarily, from the soft clay either above or below it, then q^m in the dense sand will never reach its true value. Of course, the opposite would occur in a soft, loose stratum sandwhiched by denser strata. While these latter errors have the same root cause as transition effects (i.e., measurement over a volume) they are often separetely referred to as "thin layer" effects.

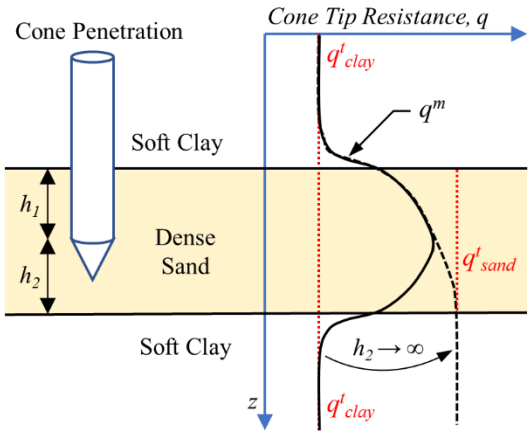


Fig. 2. Conceptual schematic showing cone penetration in layered soil. The measured cone-tip resistance (q^m) deviates from the "true" resistance (q^t) that would be measured in each material if measurements reflected conditions at discrete points (modified from Mo et al. 2017).

Nomenclature aside, the Boulanger and DeJong (2018) procedure aims to correct these collective effects by viewing q^m as the convolution of q^t with a physical low-pass filter (w_c) over a zone of influence (a depth window 60 times that of the cone diameter (d_c), centered at the cone tip). With this approach, q^m is computed as:

$$q^m(z) = q^t(z) * w_c(z) \tag{1}$$

where q^m, q^t, and w_c are a function of depth, z, and the asterisk indicates convolution, which is the integral of the point-wise multiplication of $q^t(z)$ and $w_c(z)$, as a function of the amount that one of the functions is shifted relative to the other. While complete details will not be given here, $w_c(z)$ is a function of two other terms: w_1, which decreases the relative influence of any soil as its distance from the cone tip increases; and w_2, which weights the influence of soil above or below the cone tip based on whether those soils are stronger or weaker than that immediately at the tip. Specifically, it is assumed that q^m receives more influence from the soil immediately near the tip when that soil is relatively weaker than the surrounding soil. Conversely, it is assumed that q^m receives less influence from the soil immediately near the tip when that soil is relatively stronger than the surrounding soil. In essence, the Boulanger and DeJong (2018) technique identifies the q^t that, when convolved with w_c, best predicts q^m. As part of this procedure, additional filters are used to smooth high frequency noise considering the CPT sampling interval, thereby increasing the speed and likelihood of convergence to an optimal solution.

Following estimation of q^t via inversion of q^m, Boulanger and DeJong (2018) propose, at least for the time being, that f^t be estimated from q^t. Specifically, the proposed "inversion" of f^m follows the assumption that both inverted and measured pairs of normalized tip resistance (Q) and normalized sleeve friction ratio (F) lie along the same radial line originating from the origin of the Soil Behavior Type Index, I_c, proposed by Robertson and Wride (1998), which maps in Q-F space. In effect, this approach changes I_c near interfaces, and in-turn the inferred soil type (e.g., the inferred susceptibility to liquefaction) but results in relatively minimal changes to I_c otherwise (i.e., away from interfaces). Following the initial development of q^t and f^t profiles, a separate procedure detects and corrects interfaces based on the rate of change of q^t with respect to depth. In total, the Boulanger and DeJong (2018) procedure has five parameters, for which recommended "baseline" values were provided:

$z'_{50,ref}$ = 4.2; m_{z50} = 0.5; m_z = 3; m_q = 2; and m_t = 0.1. While complete details are provided in Boulanger and DeJong (2018), these baseline parameters were herein adopted to compute "true" CPT data. It is plausible that these values could be calibrated at the site-specific level (e.g., using high resolution borings adjacent to a CPT). This would change the magnitude of correction and the sensing and development distances, thereby potentially improve identification of thin layers. However, the information compiled for this study was generally insufficient to attempt calibration and, when available, provided insufficient statistical support to justify it. As part of the CPT processing methodology, statistical cross-correlation (Buck et al. 2002) was used to align tip- and sleeve-measurements, both for the measured and "true" CPT data. All CPT processing, including implementation of the Boulanger and DeJong (2018) procedure, was completed using the open-source software *Horizon* (Geyin and Maurer 2020a).

Liquefaction Models

The Boulanger and DeJong (2018) procedure will be evaluated in the context of CPT-based liquefaction model performance. That is, whether use of the procedure improves or worsens their prediction efficiency. Towards that end, six triggering models based on the so-called "simplified stress-based" framework, first envisioned by Whitman (1971) and Seed and Idriss (1971), are herein adopted: Green et al. (2019), Boulanger and Idriss (2014), Idriss and Boulanger (2008), Moss et al. (2006), Architectural Institute of Japan (2001), and Robertson and Wride (1998). However, because these models predict triggering at specific depths below ground, a true evaluation of their performance requires subsurface exploration or instrumentation (i.e., to assess whether predicted and actual responses agree). This could potentially be achieved in advance of an earthquake using buried sensors (e.g., Holzer et al. 2007) or after an earthquake using vision penetrometers (Raschke and Hryciw 1997) or geoslicers (Nakata and Shimazaki 1997). Yet, case histories with such data are

exceedingly rare and still may not result in definitive interpretations of what did, and did not, liquefy (e.g., Takada and Atwater 2004). As a result, nearly all existing case-histories document only whether liquefaction manifestations were observed at the ground surface. Accordingly, to compare subsurface predictions of liquefaction against surface observations, the predictions from each triggering model will be input to three different models that predict surficial manifestations of liquefaction: Maurer et al. (2015a), van Ballegooy et al. (2014), and Iwasaki et al. (1978), who developed models termed *LPI*, *LSN*, and LPI_{ISH}, respectively. In this study, "liquefaction model" therefore refers to the use of two models in series: one triggering model and one manifestation model. The independent performance of these models can simply not be assessed via any practical, objective means. A summary of the 18 models to be used (6 triggering models x 3 manifestation models), and the symbols that will be used to identify them, is provided in Table 1.

Table 1. Summary of Liquefaction Triggering and Manifestation Models used in this Study.

Model	**Type**	**Symbol**
Robertson & Wride (1998)	Triggering	RW98
Arch. Institute Japan (2001)	Triggering	AIJ01
Moss et al. (2006)	Triggering	Mea06
Idriss & Boulanger (2008)	Triggering	IB08
Boulanger & Idriss (2014)	Triggering	BI14
Green et al. (2019)	Triggering	Gea19
Iwasaki et al. (1978)	Manifestation	*LPI*
van Ballegooy et al. (2014)	Manifestation	*LSN*
Maurer et al. (2015)	Manifestation	LPI_{ISH}

Liquefaction Case-History Data

A total of 15,223 liquefaction case histories from 24 earthquakes will be analyzed, as listed in Table 2. However, because most of these cases are from three earthquakes in New Zealand's Canterbury province, analyses will be carried out separately for these and the remaining 21 events, henceforth referred to as the "Canterbury dataset" and "global dataset," respectively. Each case history includes estimates of groundwater depth and peak ground acceleration (*PGA*), CPT data, and observations of the presence or absence of surficial liquefaction manifestations.

Table 2. Summary of Liquefaction Case-Histories Analyzed (after Geyin et al. 2020b)

Date	Event	Country	Magnitude (M_w)	Number of Case Histories
16/6/1964	Niigata	Japan	7.60	3
9/2/1971	San Fernando	USA	6.60	2
4/2/1975	Haicheng	China	7.00	2
27/7/1976	Tangshan	China	7.60	10
15/10/1979	Imperial Valley	USA	6.53	7
9/6/1980	Victora (Mexicali)	Mexico	6.33	5
26/4/1981	Westmorland	USA	5.90	9
26/5/1983	Nihonkai-Chubu	Japan	7.70	2
28/10/1983	Borah Peak	USA	6.88	3
2/3/1987	Edgecumbe	New Zealand	6.60	23
24/11/1987	Elmore Ranch	USA	6.22	2
24/11/1987	Superstition Hills	USA	6.54	8
18/10/1989	Loma Prieta	USA	6.93	67
17/1/1994	Northridge	USA	6.69	3
16/1/1995	Hyogoken-Nambu	Japan	6.90	21
17/8/1999	Kocaeli	Turkey	7.51	16
20/9/1999	Chi-Chi	Taiwan	7.62	34
8/6/2008	Achaia-Ilia	Greece	6.40	2
4/4/2010	Baja	Mexico	7.20	3
11/3/2011	Tohoku	Japan	9.00	7
20/5/2012	Emilia	Italy	6.10	46
4/10/2010	Darfield	New Zealand	7.10	5371
22/2/2011	Christchurch	New Zealand	6.20	4806
14/2/2016	Christchurch	New Zealand	5.70	4771

The Canterbury data was sourced from Geyin et al. (2020a, 2021), who compiled liquefaction case-histories from the: (i) M_w7.1, 4 Sept. 2010 Darfield earthquake; (ii) M_w6.2, 22 Feb. 2011

Christchurch earthquake; and (iii) M_w5.7, 14 Feb. 2016 Christchurch earthquake. The Geyin et al. (2020a, 2021) database built upon earlier compilations from Canterbury (Maurer et al. 2014, 2015b), resulting in 15,890 case histories. Of those, 14,948 were ultimately selected for analysis in the present study. Cases were excluded when: (i) the predominant manifestation of liquefaction was lateral spreading, since the manifestation models adopted herein are not intended to predict it; (ii) the depth of CPT pre-drill significantly exceeded that of the groundwater, since CPT data in the pre-drill zone must be estimated rather than measured (e.g., by extrapolating upward from just below the pre-drill); or (iii) the estimated median *PGA* was less than 0.075 g, since such cases may not provide meaningful tests of prediction efficiency, given that the absence of liquefaction is easily predicted by judgement. With respect to manifestations, Geyin et al. (2020a, 2021) classified each case history as "none," "marginal," "moderate," "severe," "lateral spreading," or "severe lateral spreading" using criteria from Green et al. (2014), wherein classifications were based on a circular sample centered on each CPT, with approximate radius of 10 m. In this study, liquefaction-model performance will be judged on the ability to predict surficial manifestations of liquefaction on free-field level ground. Cases with other expressions of liquefaction, such as lateral spreading, foundation movements, or evidence from ground-motions are removed because the adopted liquefaction models are not designed to predict these expressions. To facilitate model assessment, the Geyin et al. (2020a, 2021) case histories are binomially classified as "No Manifestation" and "Manifestation," where the latter are cases with either "minor," "moderate," or "severe" manifestations. Of the resulting cases assembled from Canterbury, 65% are "No Manifestation" and 35% are "Manifestation." The reader is referred to Geyin et al. (2020a, 2021) for further details relevant to data collection and processing, and where the complete Canterbury dataset may be obtained.

To compare results from Canterbury with regions elsewhere, 275 case histories from 21 global earthquakes will be analyzed in parallel. These cases were sourced from Geyin and Maurer (2021), who compiled case-history data from the literature. Whereas documentation of liquefaction in Canterbury was aided by remote sensing, case histories elsewhere are often preserved in less detail, occasionally with few details about the nature of manifestation. Thus, while manifestations were again classified binomially using the Green et al. (2014) criteria, uncertainty is unavoidably present. Among these 275 case histories, 42% are "No Manifestation and 58% are "Manifestation." The reader is referred to Geyin et al. (2021) for further details, and where the complete global dataset may be obtained. The Canterbury and Global datasets were previously studied by Geyin et al. (2020b), who compared the efficacies of various geospatial and geotechnical liquefaction models, and by Geyin and Maurer (2020b), who developed fragility functions for predicting the probability of liquefaction-induced ground failure. In the current study, these datasets are used to rigorously evaluate whether the Boulanger and DeJong (2018) procedure improves CPT-based liquefaction model performance.

Liquefaction Model Methodology

All calculations described in this section were performed with the software *Horizon* (Geyin and Maurer 2020a), which has been used in previous research (e.g., Geyin et al. 2020b). The six triggering models listed in Table 1 were each used to compute the factor-of-safety against liquefaction (FS_{liq}) vs. depth. As part of this process, liquefaction susceptibility was first inferred using the CPT soil-behavior-type index (I_c) (Robertson and Wride 1998). Using a criterion developed from lab and field data in Canterbury (Maurer et al. 2019), soils with $I_c > 2.5$ were assumed not susceptible. However, because this value is within the range of common, generic thresholds (e.g., 2.4-2.6) (Youd et al. 2001), it was also applied to the global dataset. Ultimately, the most salient findings of this study are found to be independent of this criterion, as will be further discussed. For soils deemed susceptible, the

IB08, BI14, and Gea19 models (see Table 1) consider the influence of fines-content (*FC*) on liquefaction resistance. Accordingly, an $I_c - FC$ model specific to Canterbury (Maurer et al. 2019) was used for the Canterbury dataset while a generic $I_c - FC$ correlation (Boulanger and Idriss 2014) was used for the global dataset. The six triggering models were otherwise implemented as proposed by the respective publications in Table 1.

The predictions of FS_{liq} vs. depth made by each triggering model were then input to *LPI*, LPI_{ISH}, and *LSN*, which all have the same objective - to characterize the cumulative response of a soil profile in the free field, thereby predicting manifestations of liquefaction at the ground surface. For brevity, the formulae describing *LPI*, *LSN*, and LPI_{ISH} are not given here but are provided in Geyin et al. (2020c). Their implementation is exactly as described in Geyin et al. (2020c) and mirrors that in popular practice.

Performance Evaluation

Receiver-operating-characteristic (ROC) analyses are a popular diagnostic tool to evaluate models (e.g., Fawcett 2006; Zou 2007) and are widely used in geoscience and geoengineering (e.g., among many, Lin et al. 2021; Upadhyaya et al. 2021; Ju et al. 2020; Sarma et al. 2020). In this study, ROC analyses will be used to: (i) quantify the efficiency of liquefaction models; and (ii) assess whether "true" CPT data improves that efficiency to a statistically significant degree. In all classification problems (e.g., predicting whether sites have observations of liquefaction), "positive" and "negative" observations overlap when plotted as a function of a diagnostic model index (e.g., *LPI*, *LSN*, etc.). As an example, two distributions are plotted in Figure 3a. ROC curves plot the true-positive prediction rate (R_{TP}) (i.e., the rate at which positives, or liquefaction manifestations, are correctly predicted) versus the false-positive prediction rate (R_{FP}) (i.e., the rate at which negatives, or a lack of liquefaction manifestations, are correctly predicted) for a range of classification "thresholds," which are used to

predict outcomes. Values above a threshold predict positives and those below a threshold predict negatives. Figure 3b depicts the relationship between the observations, thresholds, and ROC curve.

As a model segregates the distributions of positive and negative outcomes more efficiently (i.e., the distributions have less overlap), a corresponding ROC curve trends toward a point at the coordinates (0,1) in ROC space, indicating the existence of a threshold value that perfectly separates the distributions (i.e., the model is perfectly efficient). Conversely, random guessing appears as a 1:1 line in ROC space, in which case the model has no utility and the positive and negative distributions perfectly overlap. For this reason, the area under a ROC curve (*AUC*) is widely adopted to characterize model efficiency (e.g., Fawcett 2006). *AUC* also has statistical significance. In this case, it is the likelihood that sites with manifestations have larger model values than sites without manifestations. In this regard, it is equivalent to the nonparametric Wilcoxon statistic (Hanley and Mc Neil 1982). In the case of a perfectly efficient model, *AUC* = 1.0 (or 100%), whereas with random guessing *AUC* = 0.5 (or 50%). Better prediction models thus have higher *AUC*. With this approach, false-positive predictions (liquefaction is predicted but is not observed) and false-negative predictions (liquefaction is not predicted but is observed) are given equal importance. In other words, *AUC* reflects the overall misprediction rate, rather than treating either false positives or false negatives as being more important. Another desirable feature of *AUC* is its relative insensitivity to class imbalance. Suppose a hypothetical dataset includes 1 positive case and 99 negative cases. For this dataset, a hypothetical model that predicts negative outcomes 100% of the time would be 99% accurate, even though the model is objectively useless. For this reason, model accuracy – while commonly reported – is a poor metric unless the positive and negative classes are equal in size. Other oft-reported performance metrics that similarly focus on only positive or negative predictions, such as sensitivity and specificity (e.g., Powers 2011), would be similarly inappropriate for the aims of this study. *AUC*,

however, would appropriately characterize this model's lack of utility by finding a value near 0.5. Accordingly, *AUC* will be used to quantify model performance in each of the 36 performance trials.

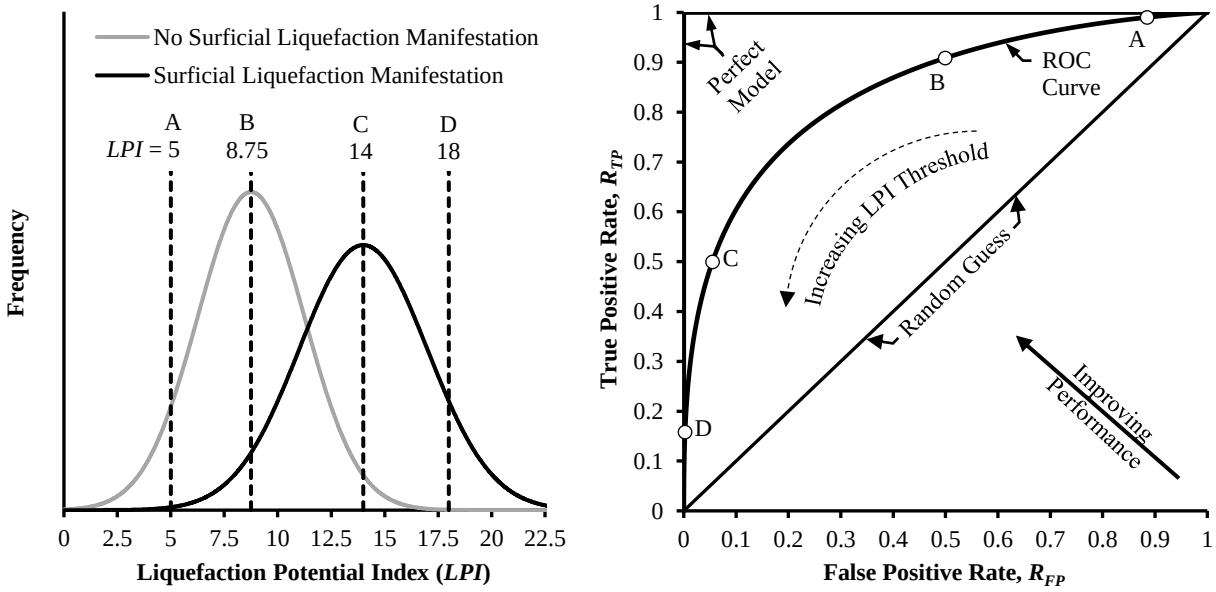


Fig. 3. ROC analyses: (a) positive and negative observations vs. computed *LPI*; (b) derivative ROC curve, and depiction of how the efficiency of a diagnostic test is assessed via *AUC* (after Geyin et al. 2020b).

To account for the finite availability of case histories within each dataset (i.e., the Canterbury and global datasets), bootstrap simulations (e.g., Diaconis and Efron 1983) will be used to quantify the finite-sample uncertainty of *AUC* for each model. This will characterize the sensitivity of results to the data chosen for study and be used to assess whether differences in *AUC* could arise from chance (i.e., due to finite samples) and not because one model is truly better. Tests of significance will be carried out via the ROC-specific DeLong et al. (1988) methodology. This approach computes the P-values, or probabilities, that two *AUC* samples could have come from the same distribution. In each of the 36 trials, model performance using measured CPT data will be compared to that using "true"

CPT data to determine which, if any, is statistically better. Since this method is predicated on *AUC* normality, both Anderson-Darling and Lilliefors tests (Anderson and Darling 1952; Lilliefors 1967) were used to confirm, with 95% confidence, that all *AUC* samples came from a normally distributed family.

Results and Discussion

Using the data and methodology above, liquefaction manifestations were predicted for CPT-based case histories by 18 models, both with and without correction by the Boulanger and DeJong (2018) procedure. To illustrate how model performance will be evaluated using ROC analysis, results will first be shown for one model, following which summary statistics from all models are presented. In Figures 4a and 4b, results using the Gea19 – *LPI* model (i.e., the Gea19 triggering model and *LPI* manifestation model) are shown for the global dataset using measured and "true" CPT data, respectively. In each figure, 95%-confidence intervals (CIs) were computed from a total of 10,000 bootstrapped samples. The median ROC curve is the same as that resulting from an analysis of all data without resampling. It can be seen, for this model and dataset, that use of the Boulanger and DeJong (2018) correction procedure decreases the median *AUC* by 1.62% (i.e., predictions are less efficient), counter to what ideally should occur.

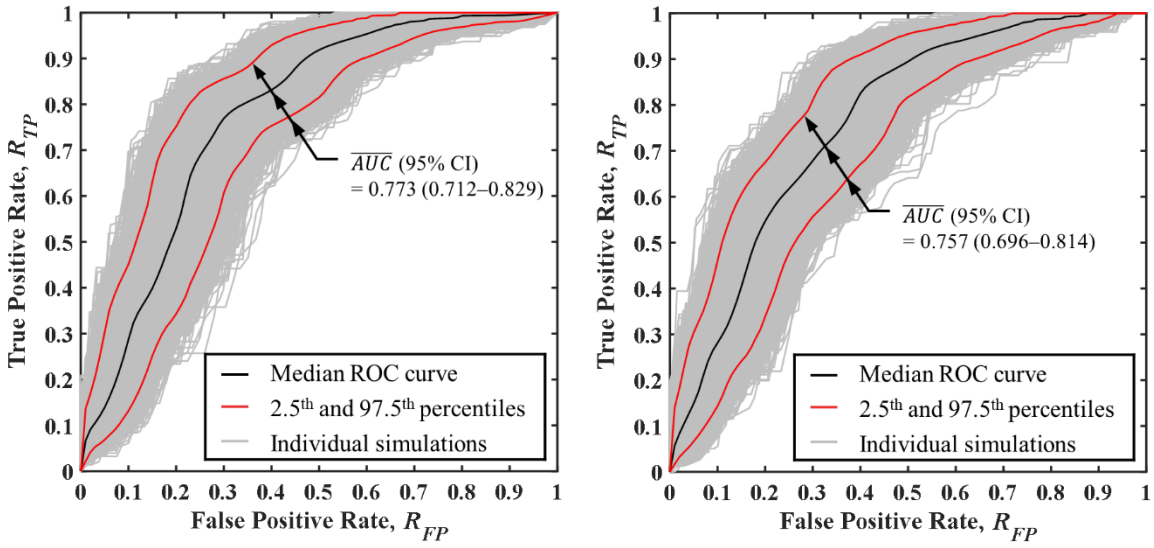


Fig. 4. ROC analysis of Gea19-*LPI* model performance in predicting surficial manifestations of liquefaction for the global dataset: (a) measured CPT data; and (b) "true" CPT data.

From identical ROC analyses of all 18 models, summary statistics are compiled in Figures 5a and 5b for the global and Canterbury datasets, respectively, wherein the resulting shift in *AUC* from using the Boulanger and DeJong (2018) procedure is plotted. Shown for each of 36 trials is the median shift in *AUC* and the 95% CI on that shift, sorted by the year each model was proposed.

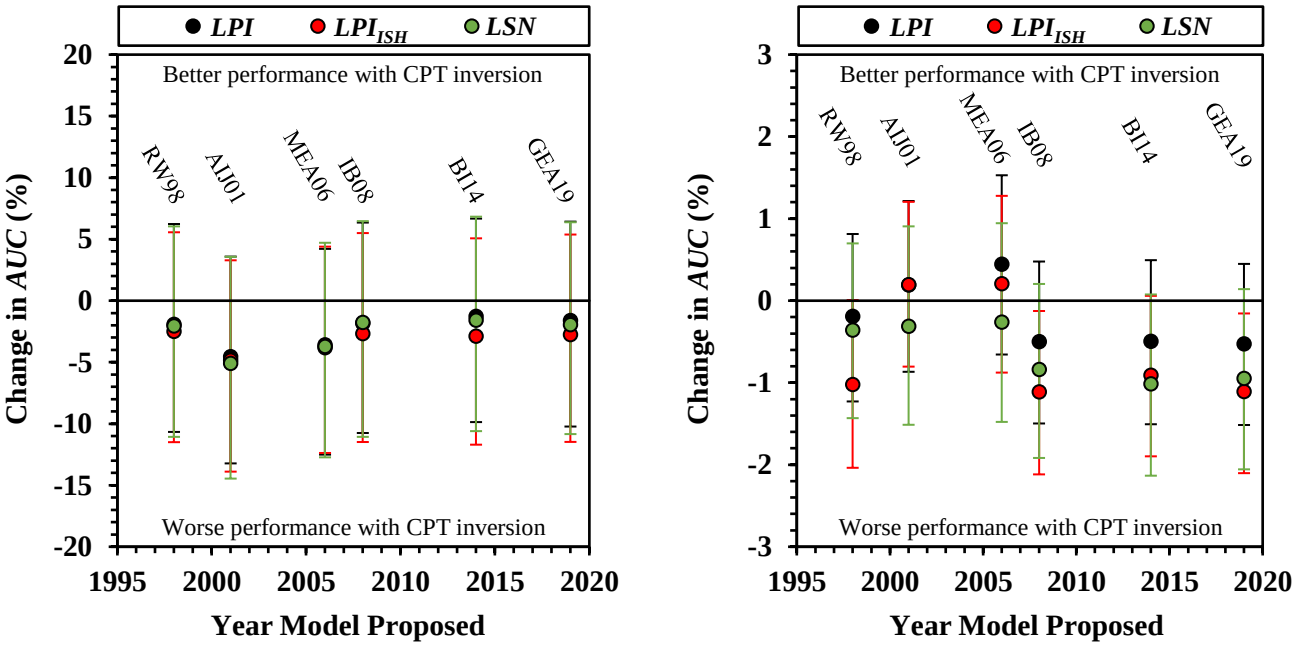


Fig. 5. Summary of change in liquefaction-model performance – quantified by *AUC* – for 18 models to which the Boulanger and DeJong (2018) procedure was applied, ordered by year proposed: (a) global dataset; and (b) Canterbury dataset. Markers denote median shift in *AUC*; bars are 95% confidence intervals on that shift; all model acronyms are identified in Table 1.

As shown in Figure 5, the behavior exhibited by Gea19-*LPI* in Figure 4 (i.e., the decline in median *AUC* using "true" CPT data) is also true of all 18 models when tested on the global dataset, and true of 14 models when tested on the Canterbury dataset. Considering all models, the average changes in *AUC* owed to using "true" CPT data are respectively -2.8% and -0.47% for the global and Canterbury datasets; the average 95% CIs on these changes are respectively -11.50% to 5.29% and -1.51% to 0.53%. The larger finite-sample uncertainty of the global dataset is expected given its much smaller size, all else being equal. Its larger uncertainty may also be further augmented by the relative geomorphic and seismologic diversity of the global dataset and/or because the data collection methods (e.g., ground-motion estimates, CPTs) varied somewhat with time and place.

When assessing the apparent, general decline in performance using "true" CPT data, it should be noted that some of the global cases used to test performance in Figure 5a were also previously used to train triggering models when they were originally developed. That is, the models were previously trained using the same measured CPT data that is now being using to test them, which might reasonably result in bias that favors the measured data and disfavors the "true" data obtained from Boulanger and DeJong (2018). In this regard, the percentage of test data previously used in training varies from 0% (e.g., RW98) to ~75% (e.g., Gea19). However, the matter of bias is more complicated. *First*, the developers of triggering models, to the degree then possible, manually corrected for interface and thin-layer effects using chart-based solutions or judgement, mitigating the possibility of the aforementioned bias. *Second*, while a case history may have been used to train a triggering model, this training was performed using a judgement-based manifestation model (i.e., an analyst selected the so-called "critical layer" using their judgement to analyze an observation at the ground surface; they did not use *LPI*, *LSN*, or LPI_{ISH} in-reverse to select it). Similarly, *LPI*, *LSN*, and LPI_{ISH} were not formulated or optimized using case-history data, but rather, were developed heuristically and then retrospectively shown to provide useful predictions on field data. In this respect, the training and test datasets used herein might be considered wholly different. Complications aside, it can be seen in Figure 5a that the 18 models perform relatively similarly on the global dataset, despite the large variability in possible bias. Regarding the Canterbury data, the BI14 and Gea19 triggering models, when originally proposed, were trained on a dataset of which 20% was from Canterbury, whereas all other models were trained independent of Canterbury data. It can be seen in Figure 5b that while the BI14 and Gea19 models do exhibit worse performance with CPT inversion, relative to others, these models perform very similar to the unbiased IB08 model, to which they owe their analytical provenance. Thus, there is no readily apparent difference between models with and without

conceivable bias. Nonetheless, the possibility that models may perform better if retrained on "true" CPT data will be explored later.

It was found that in 32 of the 36 trials performed, *AUC* decreased because of CPT inversion. To determine whether these changes in *AUC* are statistically significant, P-values were computed per DeLong et al. (1988) to compare the performance of each model with and without CPT inversion. These values are given in Table 3 and are the probabilities that two *AUC* samples (i.e., one model with and without CPT inversion) come from the same parent distribution. Thus, when P-values are small, an observed difference in *AUC* is more likely the result of CPT inversion and less likely a consequence of finite-sample uncertainty. Small P-values occur when (i) two *AUC* values are dissimilar; and/or (ii) the uncertainties of those *AUC* values are small. The common significance level of 0.05 is adopted for these analyses, meaning that two models are classified as "significantly" different if the difference is at least 95% probable. Using this approach, Table 3 summarizes, for each of 36 trials, whether CPT inversion made the efficiency of a liquefaction model better or worse and whether that change was significant.

Notable observations from Table 3 are: (i) of the 32 out of 36 trials in which performance decreased using "true" data, those decreases were significant in 23 trials and insignificant in the remaining 9; (ii) of the 4 out of 36 trials in which performance increased, those increases were significant in 1 trial and insignificant in the remaining 3. While these findings would invariably change if the adopted significance threshold were changed (say, from 0.05 to 0.10), the findings would nonetheless suggest a lack of demonstrable improvement using the Boulanger and DeJong (2018) procedure, as implemented herein. Additional inquiries will be investigated in the following.

Table 3. P-Value Matrix to Compare Model Performance With and Without CPT Inversion.

Triggering Model	Manifestation Model	*AUC* Better or Worse with CPT Inversion? (p-value)[1]	
		Global Dataset	**Canterbury Dataset**
RW98	*LPI*	Worse (0.0765)	Worse (0.0424)
	LPI_{ISH}	Worse (0.0333)	Worse (0.0001)
	LSN	Worse (0.0804)	Worse (0.0003)
AIJ01	*LPI*	Worse (0.0004)	Better (0.0511)
	LPI_{ISH}	Worse (0.0005)	Better (0.0972)
	LSN	Worse (0.0002)	Worse 0.0121)
MEA06	*LPI*	Worse (0.0044)	Better (0.0007)
	LPI_{ISH}	Worse (0.0087)	Better (0.1823)
	LSN	Worse (0.0077)	Worse (0.0772)
IB08	*LPI*	Worse (0.077)	Worse (0.0001)
	LPI_{ISH}	Worse (0.025)	Worse (0.0001)
	LSN	Worse (0.1534)	Worse (0.0001)
BI14	*LPI*	Worse (0.2565)	Worse (0.0001)
	LPI_{ISH}	Worse (0.0123)	Worse (0.0001)
	LSN	Worse (0.2111)	Worse (0.0001)
GEA19	*LPI*	Worse (0.1386)	Worse (0.0001)
	LPI_{ISH}	Worse (0.0165)	Worse (0.0001)
	LSN	Worse (0.109)	Worse (0.0001)

[1]Probability that *AUC* using measured CPT data and *AUC* using "true" CPT data could be from the same distribution. Values less than 0.05 are classified as statistically "significant", in which case significance is indicated with shading.

General Shifts in Perceived Hazard

Towards identifying conditions under which "true" CPT data changes the efficiency of liquefaction models, the general causes of shifting predictions are next investigated. Plotted in Figure 6 are Gea19-*LPI* values computed using measured CPT data versus "true" data, for both the global and Canterbury datasets. The Boulanger and DeJong (2018) procedure, on average, minimally alters the computed Gea19-*LPI* value, with average changes of +0.04 and -0.76 for the global and

Canterbury datasets, respectively. For a minority of cases, however, the computed value changes by as much as ±5 or ±10. Of course, profiles inferred to be relatively homogenous from CPT data plot near the 1:1 line in Figure 6, since the Boulanger and DeJong (2018) procedure will minimally alter such data. By corollary, profiles plotting well above or below the 1:1 line tend to be interbedded. From an investigation of all such outliers, two representative cases are next highlighted.

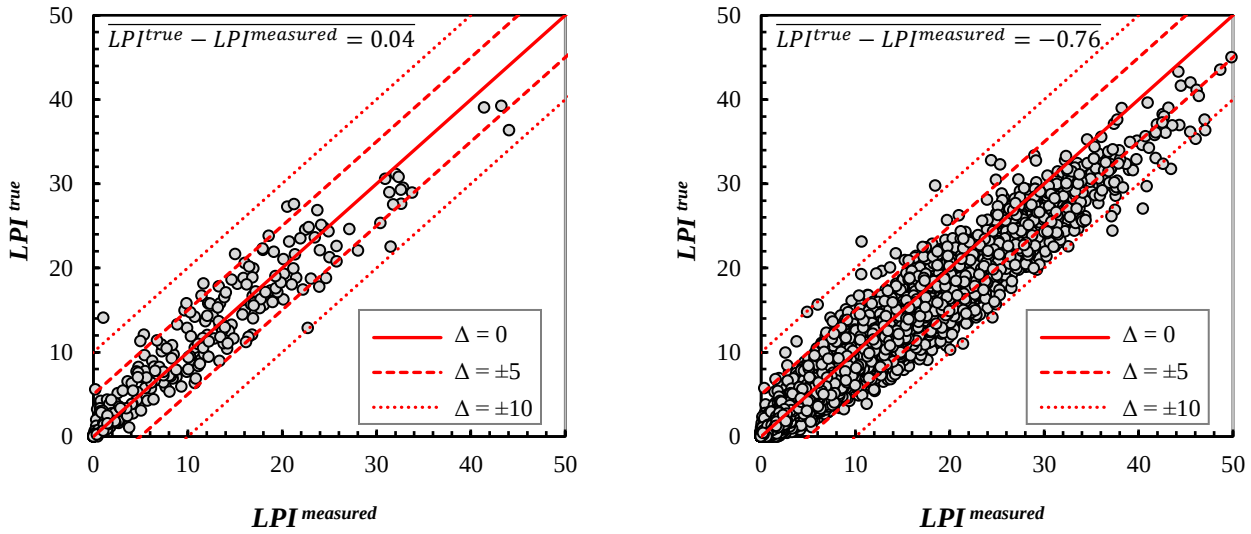


Fig. 6. Computed Gea19-*LPI* values using measured CPT data versus "true" CPT data corrected by the Boulanger and DeJong (2018) procedure: (a) global dataset; and (b) Canterbury dataset.

Plotted in Figure 7 is a case history from Canterbury in which the computed Gea19-*LPI* value increased from 5.7 to 16.8 following CPT inversion. The Boulanger and DeJong (2018) procedure both decreased and increased the measured q_c in thin layers, albeit by relatively small margins. However, the perceived hazard increased substantially because of a downward I_c shift in some layers. Specifically, where the procedure perceives that a thin layer of susceptible sand is sandwiched between softer unsusceptible soils, which may artificially decrease measured q_c and increase measured I_c, the procedure attempts to correct these effects. Thus, while the corrective increase in q_c

increases liquefaction resistance, the decrease in I_c changes the inferred soil type to one that is susceptible to liquefaction. It stands to reason, then, that CPT inversion may increase the perceived hazard for one level of seismic loading (e.g., that in Figure 7) but decrease the perceived hazard for a different, lesser level of seismic loading.

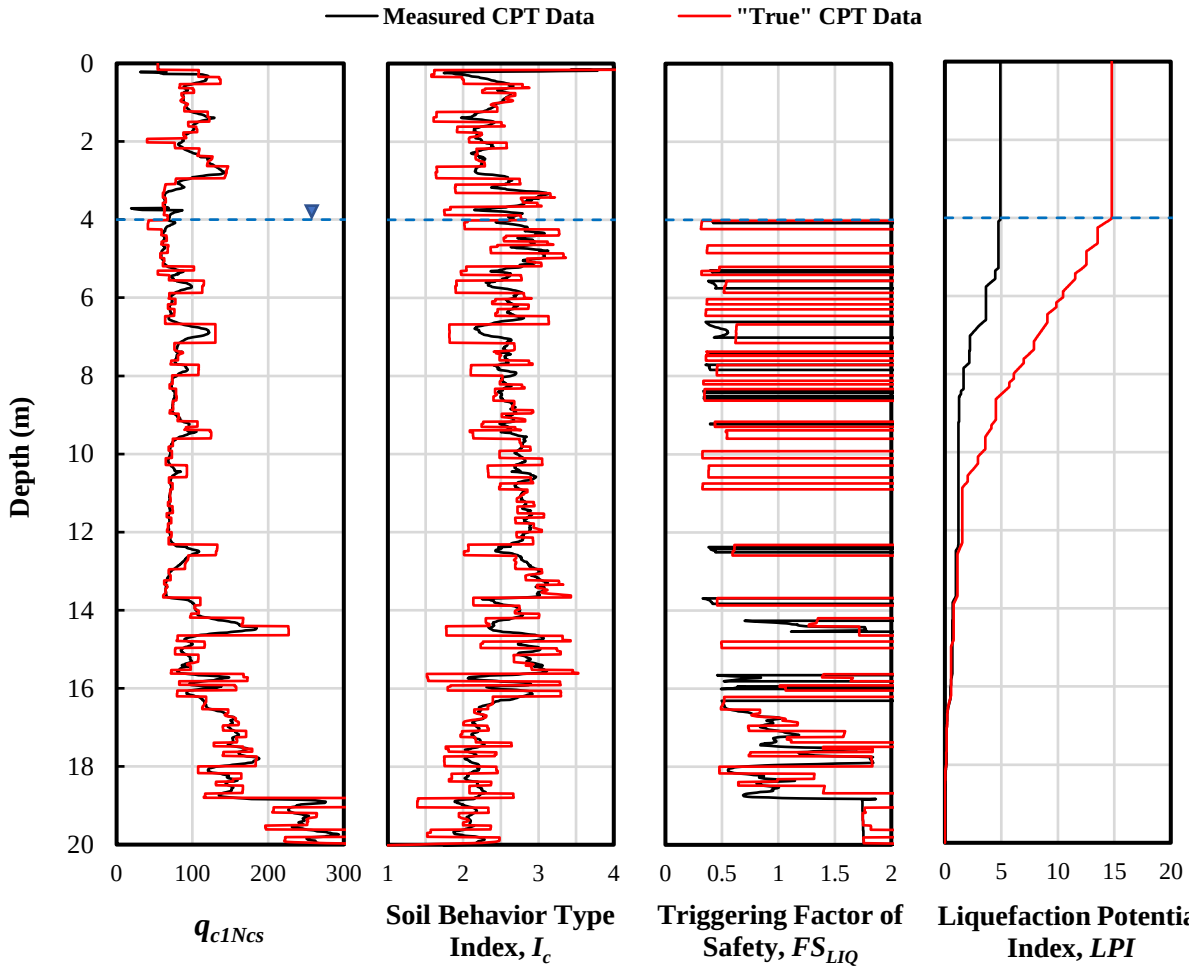


Fig. 7. Computed Gea19-*LPI* values using measured CPT data versus "true" CPT data corrected by the Boulanger and DeJong (2018) procedure [Canterbury case history #5510 from Geyin et al. (2020a, 2021)].

Plotted in Figure 8 is a case history from the global dataset in which the computed Gea19-*LPI* value decreased from 34.5 to 22.51 following CPT inversion. In this case the predominant effect of

the Boulanger and DeJong (2018) procedure was to increase measured q_c in some thin layers by a significant margin (much more than in Figure 7). This may be due to the inferred interbedded layers in Figure 8 being relatively thinner and surrounded by relatively softer material. While these corrections are conceptually reasonable, their accuracy cannot be directly assessed without "true" CPT data from calibration chamber tests or numerical simulations. Nonetheless, the preceding examples demonstrate that large upward or downward shifts in the perceived hazard can occur in highly interbedded profiles, particularly when the involved soils are transitional in nature with I_c values near the threshold for discriminating susceptibility.

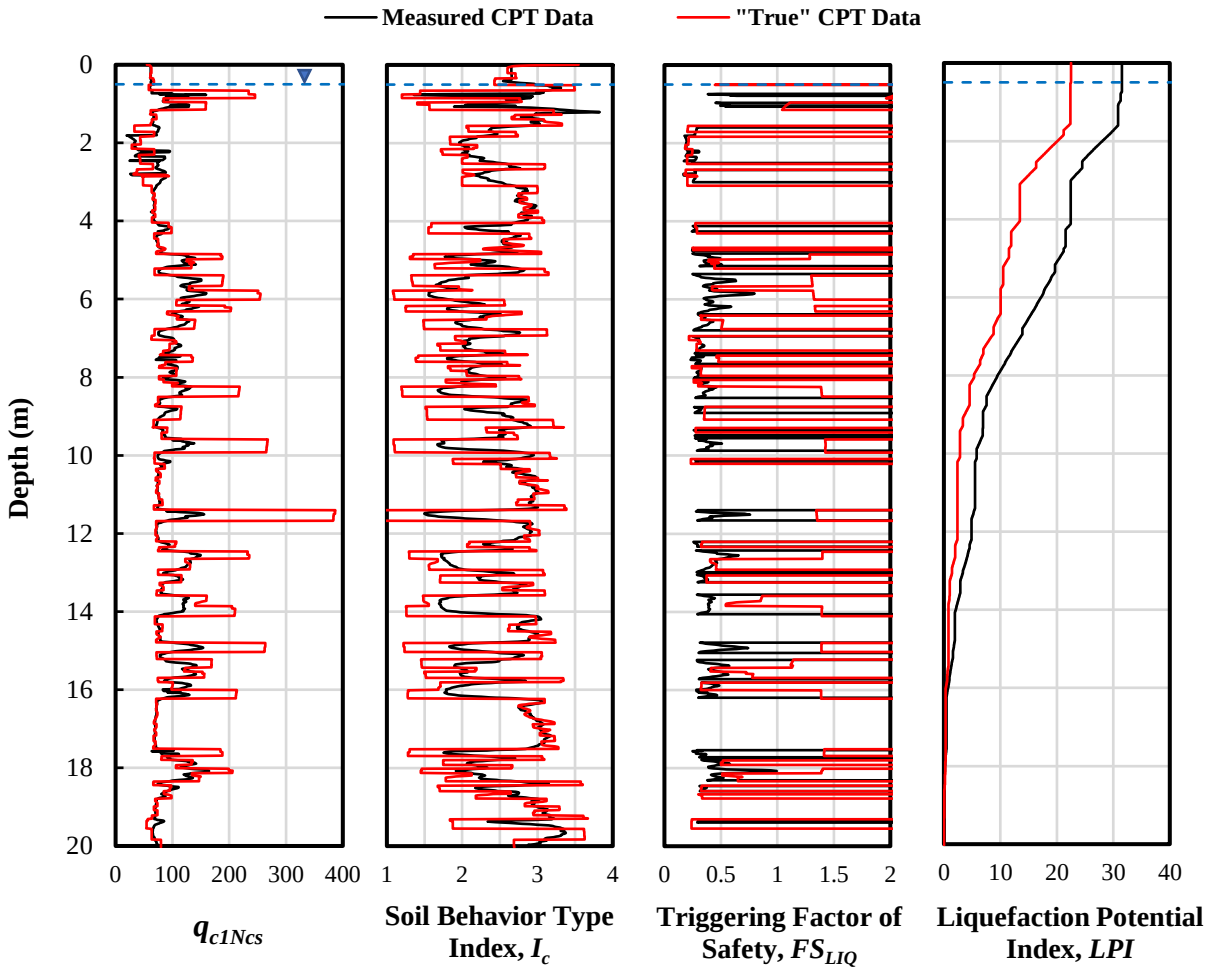


Fig. 8. Computed Gea19-*LPI* values using measured CPT data versus "true" CPT data corrected by the Boulanger and DeJong (2018) procedure [global case history #95 from Geyin and Maurer 2021)].

Dependence of Results on the I_c Cutoff Used to Infer Liquefaction Susceptibility

Given that large shifts in the perceived hazard are often associated with I_c values moving above or below the $I_c = 2.5$ threshold for discriminating susceptibility, the sensitivity of previous results to this threshold should be assessed. In Figure 7, for example, Gea19-*LPI* increased significantly because I_c values initially just above 2.5 (and thus inferred to be unsusceptible) were adjusted by the Boulanger and DeJong (2018) procedure to just below 2.5. Had the $I_c = 2.5$ threshold been increased in this case and others (e.g., to account for uncertainty in whether sleeve friction is inverted correctly), then CPT inversion could result in less drastic changes to the perceived hazard. To investigate whether the results summarized in Figure 5 and Table 3 would change if a different I_c cutoff were used, all previous analyses were repeated using cutoffs ranging from 2.5 to 3.5 in increments of 0.1. The results of these analyses, which are summarized in Figure 9, indicate that prior findings were insensitive to the I_c cutoff. That is, CPT inversion tends to slightly decrease model efficiency independent of the cutoff. While CPT inversion *does*, for some of the 18 models, slightly improve efficiency when very high cutoffs (e.g., $I_c > 3.2$) are used, those efficiencies are initially much less than optimal. Thus, while CPT inversion may technically increase model performance in this range of the I_c-cutoff domain, there is no compelling reason to employ these high cutoffs, or to use CPT inversion with the liquefaction models assessed herein.

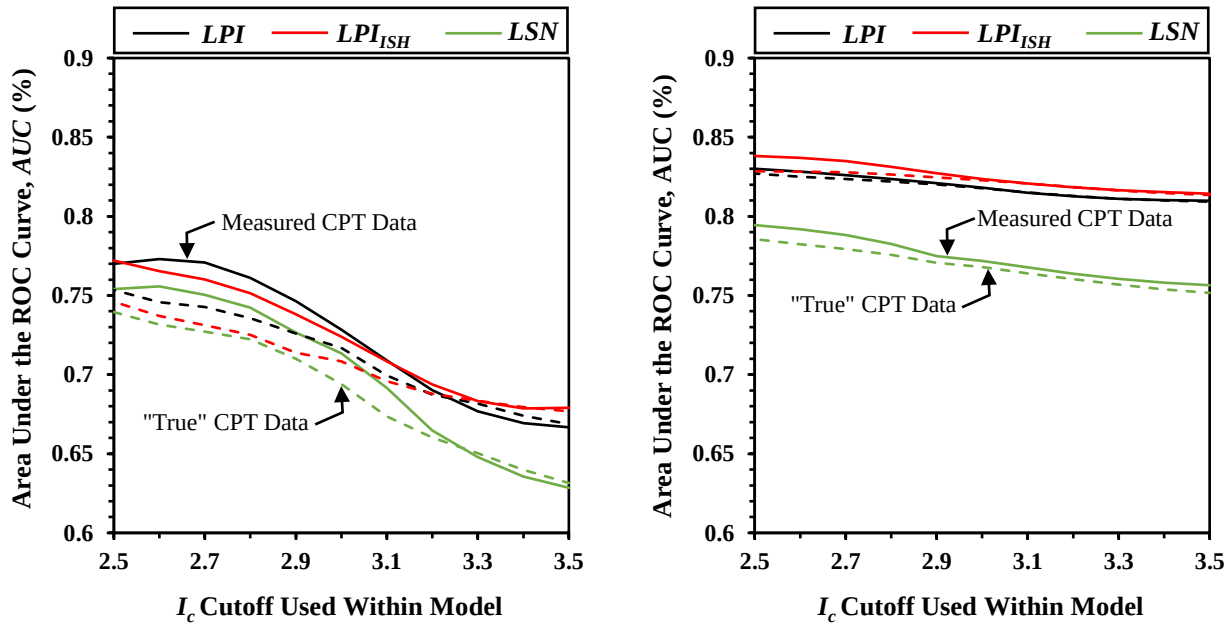


Fig. 9. Liquefaction-model performance – quantified by *AUC* – with and without CPT inversion, plotted as a function of the I_c cutoff used to infer liquefaction susceptibility: (a) global dataset; and (b) Canterbury dataset. Solid lines denote the average *AUC* across all six triggering models using measured CPT data; dotted lines denote the same average using "true" CPT data.

Correlations Between Geomorphology and Shifts in Model Efficiency

While large shifts in the perceived hazard may occur in profiles inferred to be highly interbedded, it is not yet clear whether the observed, general decline in liquefaction-model efficiency is directly connected to such profiles. For this investigation we focus on the Canterbury dataset, which provides case histories from different geologic units, with differing degrees of "interbeddedness," having experienced similar ground motions. Figure 10a shows the expected, surficial geologic units in the vicinity of Christhurch as mapped by Brown (1975) and Brown and Weeber (1992), and the locations of CPT soundings in the Canterbury dataset. The units in Figure 10a are: (A) Alluvial sand and silt

of overbank deposits; (B) Peat swamps now drained; (C) Fixed dune sand and beach sand deposits; (D) Saline sand, silt and peat of drained lagoons and estuaries; and (E) Fluviatile gravel, sand, and silt of historic flood channels.

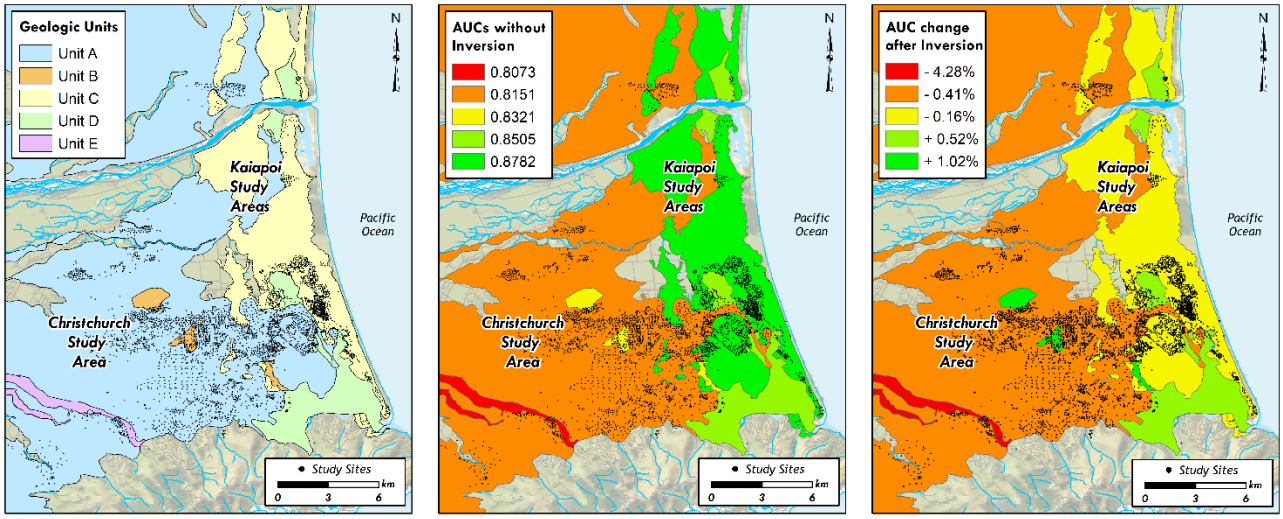


Fig. 10. (a) Mapped geologic units of the Canterbury database, as described in Table 4, and CPT locations; (b) Intra-unit *AUC*s computed from Gea19-*LPI* predictions using measured CPT data; (c) the change to intra-unit *AUC* values due to CPT inversion by the Boulanger and DeJong (2018) procedure.

Intra-unit ROC analyses were performed on the five geologic units in Table 4, resulting in *AUC* values for each model within each unit. These values convey the efficiency with which a model separates cases with and without manifestations of liquefaction, independent of cases in all other units. Results for the Gea19-*LPI* model using measured CPT data are mapped in Figure 10b and are representative of all other models. A spatial dependence may be observed, such that *AUC* values are highest in the east (e.g., dune and beach deposits) and 6-7% lower in the west (e.g., alluvial sand and silt overbank deposits). The less efficient performance of liquefaction models in western Christchurch was previously noted by other investigators (e.g., Beyzaei et al. 2017; McLaughlin 2017; Boulanger

et al. 2019). Mapped in Figure 10c are the shifts in *AUC*, within each unit, resulting from CPT inversion. While a statistically significant overall decrease in *AUC* was previously computed for the study area, performance actually locally increased in some units (where performance initially tended to be better) and locally decreased in others (where performance initially tended to be worse).

In the search for traits by which study sites in Canterbury may be further segregated, a quantitative "interbeddedness index," or the number of inferred soil type changes per meter, is proposed. Here, we adopt the following I_c boundaries proposed by Robertson and Wride (1998), which segregate soils into different behavior types: I_c = 1.31, 2.05, 2.6, 2.95, and 3.6. As an example, Robertson and Wride (1998) proposed that I_c = 2.05 separates silty sand from sandy silt. We limit this index to the upper 10 m of each CPT, given that *LPI*, *LSN*, and LPI_{ISH} assume that surface expression is largely a result of response in the upper 10 m of a profile. Adopting this index, the Canterbury dataset was binned based on inferred soil type changes per meter, wherein a moving bin width of 0.5 changes per meter, and a moving bin increment of 0.25 changes per meter, were adopted. As was done previously, 10,000 bootstrap simulations were performed in each bin to quantify finite-sample uncertainty.

Using the Gea19-*LPI* model as a representative example, *AUC* values are shown in Figure 11a as a function of interbeddedness. It can be seen that as the number of inferred soil type changes per meter increases from zero to seven, the median *AUC* decreases by ~10%. Thus, prior to CPT inversion, liquefaction models tend to perform worse as profiles become more interbedded. In this respect, it is well documented in lab, numerical, and field research that interbedded low-permeability soils can influence the triggering and surface expression of liquefaction (e.g., Fiegel and Kutter 1994; Ozutsumi et al. 2002; Brennan and Madabhushi 2005; Juang et al. 2005; Özener, et al. 2008; Maurer et al. 2015c; Cubrinovski et al. 2019). Despite this, none of the triggering or manifestation models evaluated herein explicitly consider this influence. As such, the performance exhibited in Figure 11a

could be considered unsurprising but could also be partly ameliorated via correction of thin layer effects. With respect to this possibility, Figure 11b shows the resulting change in *AUC* due to CPT inversion. As could be expected, there is little to no change in model efficiency for profiles without inferred soil type changes. However, in highly interbedded profiles, inversion tends to exacerbate the discrepancy seen in Figure 11a. That is, when many soil type changes are inferred, CPT inversion decreases liquefaction model efficiency, counter to what ideally would occur. As could be expected, and as shown in Figure 12, there exists in the Canterbury dataset a strong correlation between mapped surficial geology and the inferred interbeddedness of soil profiles. Where interbeddedness is greatest, liquefaction models tend to perform worse initially and tend to be made worse by CPT inversion.

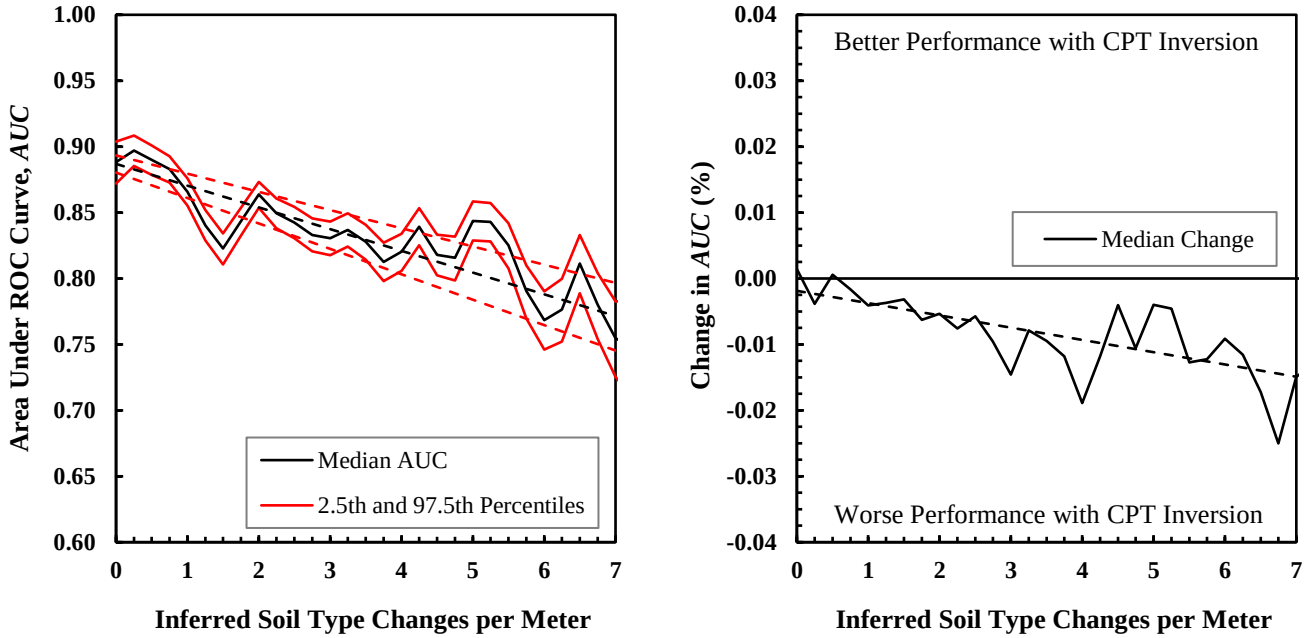


Fig. 11. Gea19-*LPI* model performance as a function of profile interbeddedness (i.e., inferred soil type changes per meter): **(a)** *AUC* prior to CPT inversion; **(b)** Change in *AUC* due to CPT inversion

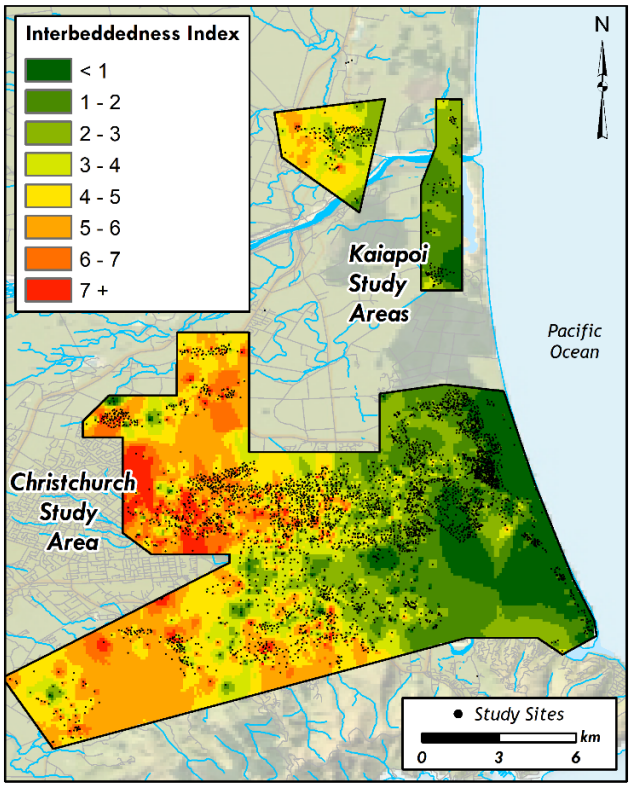


Fig. 12. Profile interbeddedness (i.e., inferred soil type changes per meter, as described in the text) across the Canterbury database study area.

Development and Assessment of a Liquefaction Model Trained on "True" CPT Data

As discussed, the results in Figure 5 (i.e., changes in *AUC* due to CPT inversion, considering all 18 models) do not suggest testing bias against "true" CPT data. Nonetheless the possibility persists that liquefaction models could perform better if they were both trained and tested on such data. To explore this possibility, the BI14 triggering model is herein retrained on case history data to which CPT inversion has been applied. Boulanger and Idriss (2014) define their limit-state triggering curve as:

$$CRR_{M=7.5,\sigma'_v=1atm} = \exp\left(\frac{q_{c1Ncs}}{113} + \left(\frac{q_{c1Ncs}}{1000}\right)^2 - \left(\frac{q_{c1Ncs}}{140}\right)^3 + \left(\frac{q_{c1Ncs}}{137}\right)^4 - C_o + \varepsilon_{\ln(R)}\right) \quad (5)$$

Where C_o is a fitting parameter that serves to scale the relationship and which has a recommended value of 2.60, and where $e_{ln(R)}$ is normally distributed with a mean of 0.0 and recommended standard deviation of $s_{ln(R)} = 0.20$. Due to historical precedent, Boulanger and Idriss (2014) proposed that their deterministic triggering model (which was herein evaluated as BI14) correspond to a triggering probability (P_L) of ~16%. As such, their deterministic model is defined by the equation above when $s_{ln(R)} = -0.20$ and is plotted in Figure 13 as a black line (note that by removing the $\varepsilon_{\ln(R)}$ term in Eq. (5), the deterministic curve is defined using $C_o = 2.8$). Adopting the optimization/training routine of BI14 exactly as prescribed therein, the BI14 triggering model was first retrained with measured CPT data using 81% of the BI14 dataset (the authors were unable to obtain the raw CPT data for the remaining 19%). Owing to this difference, an optimal C_o of 2.87 was found (as compared to the value of 2.80 proposed by BI14), creating a new baseline for comparison. This triggering curve in shown in Figure 13 as a dashed orange line.

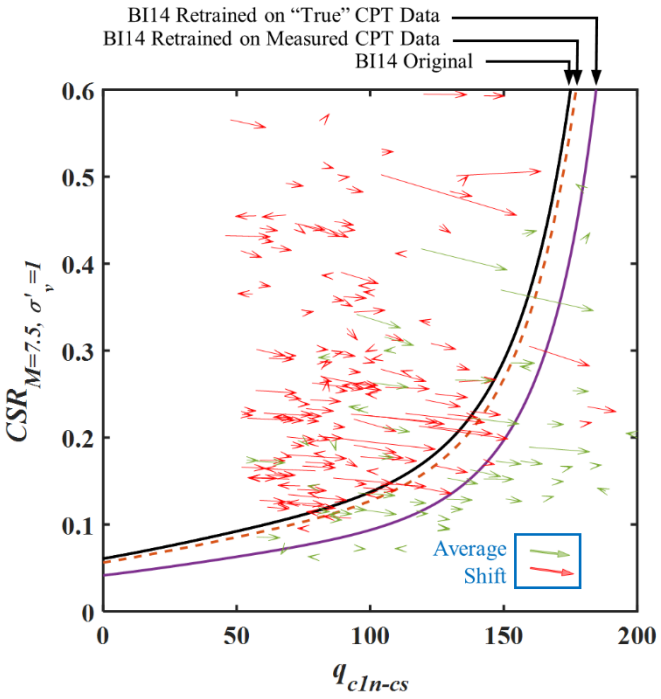


Fig. 13. Retraining of the Boulanger and Idriss (2014) triggering curve using "true" CPT data from, as described in the text. Vectors indicate the change to case-history datapoints resulting from CPT inversion.

Next, the Boulanger and DeJong (2018) inversion procedure was applied to the partial BI14 database, and "critical layers" proposed by BI14 for those cases were resampled. The resulting shift of each case-history point in triggering space is shown in Figure 13, wherein red vectors indicate shifts for cases with manifestations of liquefaction and green vectors indicate shifts for cases without manifestations of liquefaction. While CPT inversion both decreased and increased q_{c1Ncs} (cone tip resistance, normalized for overburden pressure and adjusted for fines content), the average changes were +16.24 and +14.67 for cases with and without manifestations, respectively. The corresponding changes in $CSR_{M=7.5,\ \sigma'_V=1}$ (cyclic stress ratio, normalized for overburden pressure and adjusted for earthquake magnitude) were respectively -0.007 and -0.005. This lesser change in *CSR* is

unsurprising, given the lesser dependence of *CSR* on soil properties, relative to q_{c1Ncs}. Repeating the BI14 optimization procedure, but training on "true" CPT data, the optimal C_o was 3.18 (compared to 2.87 using measured data). Given the shift of this curve, which is shown in Figure 13 as a solid purple line, it appears that "true" data is not optimally compatible with a model trained on measured data. Accordingly, using the version of BI14 newly trained on "true" data, the preceding ROC analyses of case-histories were repeated on the global and Canterbury "true" datasets. However, relative to the *AUC*s obtained by applying the original BI14 model to measured data, those obtained by applying the new BI14 model to "true" data were less. Using BI14-*LPI* as an example, *AUC* was 1.2% less in both the global and Canterbury datasets.

The final question, then, is whether any triggering curve can be found to improve predictions of liquefaction manifestations when inverted CPT data is used? To investigate, the C_0 parameter in Eq. (5) was varied from 1.5 to 5.0 in increments of 0.1, which is analogous to changing the probability associated with liquefaction triggering from 16% to various other values. For each C_0, BI14-*LPI* predictions were recomputed for each case history using "true" data, after which *AUC* values were computed for the global and Canterbury datasets. The results of these analyses, which are summarized in Figure 14, show that while *AUC*s computed from "true" data may be slightly improved using C_0 values different from the BI14 default of 2.8, these *AUC*s are still less than those computed from measured data. Thus, there exists no triggering curve obtainable via C_0 recalibration for which CPT inversion demonstrably improves the efficiency of liquefaction predictions.

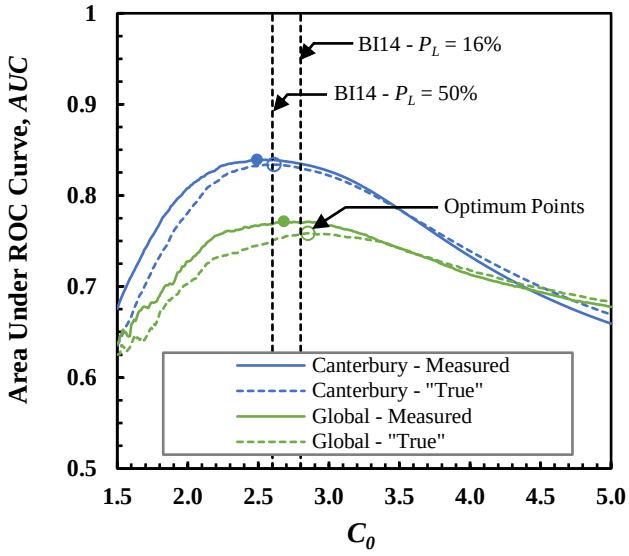


Fig. 14. BI14-*LPI AUC* values with and without inversion, as a function of the BI14 parameter C_0.

Caveats and Limitations

The presented findings are inherently tied to the datasets studied herein. The applicability of these findings to other datasets (e.g., soils of unusual age, minerology, composition, etc.) – or to other methodologies, is unknown. The phrases “true CPT data” and “CPT inversion” refer specifically to the Boulanger and DeJong (2018) inversion procedure using “baseline” parameters, as implemented in the software *Horizon*. As previously discussed, it is plausible these parameters could be calibrated at the site-specific level to potentially improve identification and correction of thin layers. As an example, it is inherently a challenge to distinguish graded strata (i.e., fining upwards or downwards) from relatively distinct material interfaces, since both appear as a rate of change in CPT data, and since both naturally occur. While the procedure includes a baseline parameter for flagging material interfaces, its selection invariably includes subjectivity that ideally could be confirmed by borehole sampling or knowledge of local geology. In the present study, all globally available liquefaction case histories are studied. As a result, site-specific calibration was not undertaken, but could conceivably

improve performance. In addition, it should be emphasized that the Boulanger and DeJong (2018) procedure was not directly evaluated, and as such, nothing can be directly concluded about its efficacy. That is, its ability to accurately correct CPT data for multiple thin-layer effects was not assessed. This would require CPT calibration chamber data or numerical simulations of such data, both in uniform and layered deposits. When analyzing field case histories, as done in this study, only the combined performance of a CPT inversion procedure *and* a liquefaction model can be quantified. In this regard, the Boulanger and DeJong (2018) procedure might improve liquefaction predictions using some triggering and/or manifestation models other than those utilized in this study. Similarly, the procedure might provide utility in other geotechnical applications. That is, the lack of utility observed herein could be due to fundamental limitations in the liquefaction models, rather than to limitations of the procedure itself. However, given that model performance at-best increased insignificantly (but most often, decreased significantly), it is unlikely that minor adjustments to liquefaction models would alter this outcome. Ultimately, additional data will confirm or update the findings presented herein and summarized below.

Conclusions

The Boulanger and DeJong (2018) CPT inversion procedure was evaluated in the context of CPT-based liquefaction model performance. Using field case-histories parsed into 2 datasets, 18 different liquefaction models were studied, resulting in 36 performance trials. In only 1 trial did the CPT inversion procedure increase model efficiency to a statistically significant degree, while in 23 others it significantly decreased efficiency. This decline in performance, which was independent of the I_c cutoff used to infer liquefaction susceptibility, tended to grow as profiles became more stratified, opposite of what ideally should occur. To explore possible remedies, a liquefaction triggering curve was rederived from inverted CPT data, such that its training and forward implementation were made

consistent. Nonetheless, this exacerbated the decline in prediction efficiency when applied to field case histories. Moreover, no readily conceivable triggering curve could be found to improve predictions of liquefaction when inverted CPT data was used. Ultimately, the results of this study are not a direct assessment of the pioneering Boulanger and DeJong (2018) procedure. However, the results do provide strong evidence that this procedure may provide no demonstrable performance benefit when applied to existing CPT-based liquefaction models. This conclusion should be weighed against caveats and limitations discussed in the preceding section.

Acknowledgements

This study is based on work supported by the National Science Foundation (NSF), US Geological Survey (USGS), and Pacific Earthquake Engineering Research Center (PEER) under Grant Nos. CMMI-1751216, G18AP-00006, and 1132-NCTRBM, respectively. The authors also wish to acknowledge the numerous researchers who contributed to the data studied herein. Much of this data was collected under the sponsorship of either the New Zealand Earthquake Commission (EQC) or one of the three organizations above. 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, USGS, PEER, or EQC.

Data Availability

All data analyzed in this study is publicly available. In addition, many calculations performed as part of the study, including CPT processing and liquefaction modelling, were carried out using *Horizon*, a freely available program developed by the authors.

References

Ahmadi MM and Robertson PK (2005) Thin-layer effects on the CPT q_c measurement. *Canadian Geotechnical Journal* 42(5): 1302-1317.

Anderson TW and Darling DA (1952) Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics* 23(2): 193-212.

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

Bastin, S., van Ballegooy, S., Mellsop, N., & Wotherspoon, L. (2020). Liquefaction case histories from the 1987 Edgecumbe earthquake, New Zealand–Insights from an extensive CPT dataset and paleo-liquefaction trenching. *Engineering Geology*, 271: 105404.

Beyzaei CZ, Bray JD, van Ballegooy S, Cubrinovski M, & Bastin S (2018) Depositional environment effects on observed liquefaction performance in silt swamps during the Canterbury earthquake sequence. *Soil Dynamics and Earthquake Engineering* 107: 303-321.

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 DeJong JT (2018) Inverse filtering procedure to correct cone penetration data for thin-layer and transition effects. *Cone Penetration Testing 2018*, Hicks, Pisano, and Peuchen, eds., Delft University of Technology, The Netherlands: 25-44.

Boulanger RW, Khosravi M, Cox BR and DeJong JT (2019) Liquefaction Evaluation for an Interbedded Soil Deposit: St. Teresa's School, Christchurch, New Zealand. IACGE 2018: Geotechnical and Seismic Research and Practices for Sustainability: 686-704.

Brennan AJ and Madabhushi, SP (2005) Liquefaction and drainage in stratified soil. *JGGE* 131(7): 876-885.

Brown LJ (1975) Water well data. Sheet 576/7-8, Belfast-Styx. New Zealand Geological Survey, Report 72.

Brown LJ and Weeber JH (1992) Geology of the Christchurch urban area: Institute of Geological and Nuclear Sciences Geological Map 1. Institute of Geological and Nuclear Sciences Limited, Lower Hutt, New Zealand, scale, 1(25,000), 1.

Buck JR, Daniel MM, and Singer AC (2002) *Computer Explorations in Signals and Systems Using MATLAB®*, 2nd Edition. Upper Saddle River, NJ: Prentice Hall.

Chen, Q., Wang, C., & Juang, C. H. (2016). Probabilistic and spatial assessment of liquefaction-induced settlements through multiscale random field models. *Engineering Geology, 211*, 135-149.

Ching J, Wang JS, Juang CH, & Ku CS (2015) Cone penetration test (CPT)-based stratigraphic profiling using the wavelet transform modulus maxima method. *Canadian Geotechnical Journal 52*(12): 1993-2007.

Cubrinovski M, Rhodes A, Ntritsos N, and van Ballegooy S, (2019). System response of liquefiable deposits. SDEE 124: 212-229.

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.

Diaconis, P and Efron B (1983) Computer intensive methods in statistics. *Sci. Amer*, 248(5): 116–130.

Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874.

Fiegel GL and Kutter BL, (1994). Liquefaction mechanism for layered soils. *JGGE* 120(4): 737-755.

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.

Geologismiki (2020). CLiq v3, CPT soil liquefaction software, https://geologismiki.gr/products/cliq/. Geologismiki Geotechnical Software, Serres, Greece.

Geyin M, and Maurer BW (2019). An analysis of liquefaction-induced free-field ground settlement using 1,000+ case-histories: observations vs. state-of-practice predictions. Geotechnical Special Publication 308: 489-498.

Geyin M and Maurer BW (2020a) Horizon: CPT-based liquefaction risk assessment and decision software. DesignSafe-CI, doi: 10.17603/ds2-2fky-tm46.

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

Geyin M and Maurer BW (2021) CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1). DesignSafe-CI. https://doi.org/10.17603/

Geyin M, Maurer BW, Bradley BA, Green RA, and van Ballegooy S. (2020a) CPT-Based Liquefaction Case Histories Resulting from the 2010-2016 Canterbury, New Zealand, Earthquakes: A Curated Digital Dataset (Version 2). DesignSafe-CI. https://doi.org/10.17603/ds2-tygh-ht91.

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. *Earthquake Spectra*, In Press.

Green RA, Cubrinovski M, Cox B, Wood C, Wotherspoon L, Bradley B and Maurer B (2014) Select Liquefaction Case Histories from the 2010-2011 Canterbury Earthquake Sequence. *Earthquake Spectra* 30(1): 131-153.

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 Eng* 17(8): 4539-4557.

Hanley JA & McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology 143*(1): 29-36.

Hasek, M. J., & Gassman, S. L. (2019). Cyclic resistance ratio of pleistocene-age sands from the South Carolina coastal plain. *Engineering Geology*, *251*, 158-171.

Heidari, T., & Andrus, R. D. (2010). Mapping liquefaction potential of aged soil deposits in Mount Pleasant, South Carolina. *Engineering Geology*, *112*(1-4), 1-12.

Holzer TL and Youd TL (2007) Liquefaction, ground oscillation, and soil deformation at the wildlife array, California. *Bulletin of the Seismological Society of America* 97(3): 961-976.

Idriss IM and Boulanger RW (2008). Soil liquefaction during earthquakes. *Monograph MNO-12* 2008; Earthquake Engineering Research Institute, Oakland, CA, 261 pp.

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

Ju, N., Huang, J., He, C., Van Asch, T. W. J., Huang, R., Fan, X., ... & Wang, J. (2020). Landslide early warning, case studies from Southwest China. *Engineering Geology*, *279*: 105917.

Juang, C. H., Lu, C. C., & Hwang, J. H. (2009). Assessing probability of surface manifestation of liquefaction at a given site in a given exposure time using CPTU. *Engineering Geology*, 104(3-4): 223-231.

Khoshnevisan, S., Juang, H., Zhou, Y. G., & Gong, W. (2015). Probabilistic assessment of liquefaction-induced lateral spreads using CPT—focusing on the 2010–2011 Canterbury earthquake sequence. *Engineering Geology, 192*, 113-128.

Lilliefors HW (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association 62*(318): 399-402.

Lin, A., Wotherspoon, L., Bradley, B., & Motha, J. (2021). Evaluation and modification of geospatial liquefaction models using land damage observational data from the 2010–2011 Canterbury Earthquake Sequence. *Engineering Geology*, *287*: 106099.

Lunne T, Robertson PK and Powell JM (1997) *Cone Penetration Testing in Geotechnical Practice*. Blackie Academic & Professional, London, U.K.

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

Maurer BW, Green RA and Taylor ODS (2015a) 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 and Bradley B (2015b) Assessment of CPT-based methods for liquefaction evaluation in a liquefaction potential index framework. Géotechnique 65(5): 328-336.

Maurer BW, Green RA, Cubrinovski M and Bradley BA (2015c) 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, van Ballegooy S, and Wotherspoon L (2019). Development of region-specific soil behavior type index correlations for evaluating liquefaction hazard in Christchurch, New Zealand. *SDEE* 117: 96-105.

McLaughlin K (2017) Investigation of false-positive liquefaction case history sites in Christchurch, New Zealand. M.S. Thesis. The University of Texas at Austin, Austin, TX.

Mo PQ, Marshall AM and Yu HS (2017). Interpretation of cone penetration test data in layered soils using cavity expansion analysis. *JGGE 143*(1): 04016084.

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. *JGGE* 132(8): 1032-1051.

Nakata T and Shimazaki K (1997) Geo-slicer, a newly invented soil sampler, for high-resolution active fault studies. *Journal of Geography* 106(1): 59-69 (in Japanese).

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.

Norini, G., Aghib, F. S., Di Capua, A., Facciorusso, J., Castaldini, D., Marchetti, M., ... & Piccin, A. (2021). Assessment of liquefaction potential in the central Po plain from integrated geomorphological, stratigraphic and geotechnical analysis. *Engineering Geology*, 282: 105997.

Özener P, Özaydin K, and Berilgen M (2008). Numerical and Physical Modeling of Liquefaction Mechanisms in Layered Sands. *Geotechnical Earthquake Engineering and Soil Dynamics* IV, 1-12.

Ozutsumi O, Sawada S, Iai S, Takeshima Y, Sugiyama W, & Shimazu T (2002). Effective stress analyses of liquefaction-induced deformation in river dikes. *SDEE* **22**(9): 1075-1082.

Powers, D.M.W. (2011). Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies* (1): 37–63.

Raschke SA, and Hryciw RD (1997). Vision cone penetrometer for direct subsurface soil observation. *JGGE* **123**(11): 1074-1076.

Robertson PK (2011) Automated detection of CPT transition zones. *Geotechnical News*, June: 35-38.

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

Sarma, C. P., Dey, A., & Krishna, A. M. (2020). Influence of digital elevation models on the simulation of rainfall-induced landslides in the hillslopes of Guwahati, India. *Engineering Geology*, *268*: 105523.

Seed HB & Idriss IM (1971). Simplified procedure for evaluating soil liquefaction potential. *ASCE Journal of Soil Mechanics and Foundations Division* 97(9): 1249–1273.

Takada K and Atwater BF (2004). Evidence for liquefaction identified in peeled slices of Holocene deposits along the lower Columbia River, Washington. *BSSA* 94(2): 550-575.

Treadwell DD (1976) The influence of gravity, prestress, compressibility, and layering on soil resistance to static penetration. PhD Dissertation, Univ. of California, Berkeley, CA.

Upadhyaya, S., Maurer, B. W., Green, R. A., & Rodriguez-Marek, A. (2021). Selecting the optimal factor of safety or probability of liquefaction triggering for engineering projects based on misprediction costs. *Journal of Geotechnical and Geoenvironmental Engineering*, *147*(6): 04021026.

van Ballegooy S, Malan P, Lacrosse V, Jacka ME, Cubrinovski M, Bray JD, O'Rourke TD, Crawford SA, and Cowan H, (2014). Assessment of liquefaction-induced land damage for residential Christchurch. *Earthquake Spectra* 30(1): 31-55.

van der Linden TI, De Lange DA and Korff M (2018) Cone penetration testing in thinly inter-layered soils. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering* 171(3): 215-231.

Whitman RV, (1971) Resistance of soil to liquefaction and settlement, *Soils & Foundations* 11(4): 59–68.

Youd TL, Idriss IM, Andrus RD, Arango I, Castro G, Christian JT, Dobry R, Finn WDL, Harder LF, Hynes ME et al. (2001) Liquefaction resistance of soils: summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils. *JGGE* 127: 817–833.

Zhang, J., Chen, F.Y., Juang, C.H., & Chen, Q. (2018). Developing joint distribution of amax and Mw of seismic loading for performance-based assessment of liquefaction induced structural damage. *Engineering Geology*, 232: 1-11.

Zhang, J., Juang, C.H., Martin, J.R., & Huang, H.W. (2016). Inter-region variability of Robertson and Wride method for liquefaction hazard analysis. *Engineering Geology*, 203: 191-203.

Zou KH (2007) Receiver operating characteristic (ROC) literature research. *On-line bibliography*, http://www.spl.harvard.edu/archive/spl-pre2007/pages/ppl/zou/roc.html

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	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology,* In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

AN AI DRIVEN, MECHANISTICALLY GROUNDED GEOSPATIAL LIQUEFACTION MODEL FOR RAPID RESPONSE AND SCENARIO PLANNING

Geyin, M.[4], Maurer, B.W., and Christofferson, K.[2]

Abstract

Geospatial models for predicting soil liquefaction infer subsurface traits via satellite remote sensing and mapped information, rather than directly measure them with subsurface tests. Field tests of such models have demonstrated both provocative potential and significant shortcomings. Informed by these tests, this paper develops geospatial models driven by algorithmic learning (benefiting from machine- and deep-learning insights, or ML/DL) but pinned to a physical framework (benefiting from mechanics and the knowledge of regression modelers). With this approach, subsurface test data are predicted remotely within the framing of a popular mechanistic model for probabilistically predicting ground failure. This has three advantages: (i) a mechanistic underpinning; (ii) a significantly larger training set, with the model principally trained on in-situ test data, rather than on ground failures; and (iii) insights from ML/DL, with greater potential for geospatial data to be exploited. While liquefaction is a physical phenomenon best predicted by mechanics, subsurface traits lack theoretical links to above-ground parameters, but correlate in complex, interconnected ways - a prime problem for ML/DL. Preliminary models are developed from a modest U.S. dataset, are tested on recent earthquakes, and are shown to be efficient. The models are coded in free, simple-to-use Windows software. The only input is a ground-motion raster, downloadable minutes after an earthquake and available for countless future earthquake scenarios. Ultimately, the proposed approach could be

[1] Department of Civil and Environmental Engineering, University of Washington, Seattle USA.
[2] Global Innovation Exchange, University of Washington, Seattle USA

expanded to further improve geospatial liquefaction models using additional training data and new predictor variables.

Introduction

As evidenced by earthquakes occurring globally each year, reliable predictions of soil liquefaction are needed both prior to an earthquake for efficient planning and mitigation, and immediately after an earthquake for informing response, reconnaissance, and recovery. Such predictions would thus ideally have the capability of being made: (i) rapidly (e.g., in near-real-time after an event); (ii) at high resolution (e.g., consistent with the scale of individual assets); and (iii) over the regional extents impacted by large earthquakes (e.g., that of a metropolis or transportation system). Problematically, state-of-practice liquefaction models require relatively costly geotechnical data, such as that from the cone-penetration-test (CPT). Given the infeasibility of in-situ testing across vast areas, regional-scale predictions of liquefaction have traditionally relied on geologic maps, from which generic areal classifications of liquefaction susceptibility may be assumed (e.g., Youd and Hoose, 1977). Such an approach is simple, but also unacceptably uncertain for most intents. Alternatively, statistical distributions of geotechnical data within geologic units may be developed and used to predict liquefaction (e.g., Holzer et al., 2011). This is more data-driven and likely to be more accurate, but many in-situ tests are still required within each mapped unit of interest. Moreover, predictions using this approach have the coarse resolution of geologic maps and assume intra-unit uniformity (i.e., local conditions are not considered), meaning an entire city could potentially receive the same prediction. It thus remains a persistent challenge to model liquefaction (or any geohazard) in a manner that is regional and rapid, yet high resolution and accurate. The existence of a model having these traits is conceivable, however, given the growth of community geotechnical datasets, remote sensing, and algorithmic learning (i.e., machine- and deep-learning, or ML/ML).

Towards that end, interest has grown in prediction models that use inputs readily available from satellite remote-sensing and existing, mapped information. In contrast to geotechnical methods, "geospatial" models can predict liquefaction rapidly, at infinitely many locations. This is made possible using geospatial proxies of soil properties relevant to liquefaction (i.e., above-ground inferences of below-ground conditions). While the concept of such a model is not new (e.g., Kramer, 2008; FEMA, 2013), the model of Zhu et al. (2017), lightly modified by Rashidian and Baise (2020), is arguably the most rigorously formulated and widely accepted. It is also implemented in the United States Geological Survey (USGS) "PAGER" system, which provides content on possible earthquake impacts (Wald et al., 2008). In a recent study, Geyin et al. (2020) tested the Zhu et al. (2017) model against 18 CPT-based models using ~15,000 liquefaction case histories (essentially all CPT-based case histories globally available to-date). These analyses elucidated both the provocative potential of geospatial data, as well as significant room, and potential means, for improving existing geospatial models.

Informed by these analyses, this paper aims to develop an improved geospatial model driven by algorithmic learning (benefiting from ML/DL insights) but pinned to a physical framework (benefiting from mechanics and the knowledge of regression modelers). In the following, the typology of liquefaction models is succinctly summarized (to place this paper, and the methods it will utilize, in context). Next, tests of the Zhu et al. (2017) model are summarized and lessons for improvement are discussed, as are the advantages of the proposed approach. This approach is then used to develop three ML/DL models that predict the probability of liquefaction-induced ground failure. Finally, these models are tested using unbiased data and implemented in *RapidLiq*, a new Windows software program.

A Succinct Overview of Geotechnical and Geospatial Liquefaction Models

The typology of models for predicting liquefaction roughly consists of three tiers: (Tier 1) fully-empirical models that require only geospatial or geologic information (e.g., Rashidian and Baise, 2020); (Tier 2) semi-mechanistic "stress-based" models that require in-situ test data and are widely used in engineering practice (e.g. Boulanger and Idriss, 2014; Green et al., 2019); and (Tier 3) numerical constitutive models, which require many material and model parameters (e.g., Cubrinovski and Ishihara, 1998; Ziotopoulou and Boulanger, 2016). While improvements to computational throughput have grown the use of "Tier 3" models, their application is still limited to specific sites and special projects, given the required inputs and operator skill. Given the rapid and regional scale aims of the proposed work, "Tier 3" models will not be used herein, which is not to say that such models could not conceivably be implemented at regional scale.

Many "Tier 2" models are popular in engineering practice. These include, among others, Robertson and Wride (1998), Moss et al. (2006), Idriss and Boulanger (2008), Boulanger and Idriss (2014), and Green et al. (2019), which all use data from the CPT to predict liquefaction as a function of earthquake magnitude (M_w) and peak ground acceleration (PGA). However, because these models predict the factor of safety against liquefaction "triggering" (FS_{liq}) at-depth within a profile, the outputs are often used cooperatively with other models that predict manifestations of liquefaction at the surface (i.e., "ground failure"). One popular manifestation model is the liquefaction potential index (LPI), proposed by Iwasaki et al. (1978):

$$LPI = \int_0^{20\,m} F(FS_{liq}) \cdot w(z)\,\mathrm{d}z \tag{1}$$

where $F(FS_{liq})$ and $w(z)$ weight the respective influences of FS_{liq} and depth, z, on surface manifestation. Specifically, $F(FS_{liq}) = 1 - FS_{liq}$ for $FS_{liq} \leq 1$ and $F(FS_{liq}) = 0$ otherwise; $w(z) = 10 -$

$0.5z$. *LPI* thus assumes that surface manifestation depends on the thickness of all liquefied strata in the upper 20 m, the degree to which FS_{liq} in each stratum is less than 1.0, and how near those strata are to the surface. *LPI* can range from zero to 100, with surface manifestations becoming more likely as *LPI* increases (e.g., Maurer et al., 2014; Geyin and Maurer, 2020a). Other similar manifestation models include those of van Ballegooy et al. (2014) and Maurer et al. (2015a).

"Tier 1" geospatial models, which aim to predict liquefaction via readily available predictor variables, have recently received renewed attention. Like "Tier 2" models, geospatial models characterize liquefaction demand via ground-motion intensity measures (IMs) (e.g., *PGA*). But, instead of quantifying liquefaction resistance with in-situ measurements, geospatial models predict below-ground conditions using above-ground information. Examples of such predictors include (among many) the slope and roughness of the surface; the distance to rivers and coasts; and compound-topographic-index, which can be derived from satellite data or existing prediction maps. Geospatial models are well suited for regional scale applications such as: (i) loss estimation and disaster simulation; (ii) city planning and policy development; (iii) emergency response; and (d) post-event reconnaissance (e.g., to remotely identify sites of interest).

The geospatial model originally proposed by Zhu et al. (2017) is a logistic regression model of the form $P(X) = (1 + e^{-X})^{-1}$ where X is a sequence of predictor variables and coefficients, and $P(X)$ is the likelihood of ground failure (i.e., surface manifestation). The model, which was trained on observations of ground-failure, takes on two forms depending on a site's vicinity to a coastline. The equations for model parameter X are in Table 1. The variables are: *PGV* = peak ground velocity (cm/s); V_{S30} = shear-wave velocity of the upper 30-m (m/s) predicted from topography (Wald and Allen, 2007); *dr* = closest distance to a river (km) in the Lehner et al. (2006) dataset; *dc* = distance to coast (km); *dw* = the lesser of *dr* and *dc* (km); *precip* = mean annual precipitation (mm) (Fick and

Hijmans, 2017); and *wtd* = predicted water table depth (m) (Fan and Miguez-Macho, 2013). Following additional testing, Rashidian and Baise (2020) proposed two minor modifications to mitigate false positive predictions: (i) the model's output should be reassigned as zero below a *PGA* of 0.1 g; and (ii) the *precip* input should be capped at 1700 mm/yr.

Table 1. Geospatial Liquefaction Model Equations (Zhu et al., 2017; Rashidian and Baise, 2020).

Model	Model Parameter X
(Coastal)	$12.435 + 0.301 \cdot \ln(PGV) - 2.615 \cdot \ln(V_{s30}) + 5.556 \times 10^{-4} \cdot precip - 0.0287 \cdot (d_c)^{0.5} + 0.0666 \cdot d_r - 0.0369 \cdot d_r \cdot (d_c)^{0.5}$
(Inland)	$8.801 + 0.334 \cdot \ln(PGV) - 1.918 \cdot \ln(V_{s30}) + 5.408 \times 10^{-4} \cdot precip - 0.2054 \cdot d_w - 0.0333 \cdot wtd$

A Test of Geospatial Liquefaction-Model Performance

Using approximately 15,000 liquefaction case histories compiled from 23 earthquakes by Geyin et al. (2021) and Geyin and Maurer (2021a), Geyin et al., 2020 tested the Zhu et al. (2017) geospatial model against 18 different CPT methods for predicting liquefaction surface manifestation. These were comprised of six different triggering models used in series with three different manifestation models. Because most of the case histories were sourced from three events in Canterbury, New Zealand, test cases were parsed into the "Canterbury" and "Global" datasets. Performance was quantified via receiver-operating-characteristic (ROC) analyses – specifically the area under the ROC curve, or *AUC* – which is a popular metric of prediction efficiency (e.g., Fawcett, 2005). Using this metric, a perfectly efficient model achieves an *AUC* of 1.0 whereas a model on par with random guessing achieves an *AUC* of 0.5. P-values were computed per the method of DeLong et al. (1988) to determine whether measured differences in *AUC* could have arisen by chance (i.e., due to finite-sample uncertainty) and not because one model is more efficient than another. While the reader is referred to Geyin et al. (2020) for complete details, the most salient results are summarized as follows.

First, on the “Canterbury” dataset, the geospatial model performed significantly better than 16 out of 18 CPT models, with a measured *AUC* of 0.84. Against the top two CPT models, it was statistically indifferent, and thus either outperformed or matched all 18 CPT models. This is a surprising and provocative result, given the relative costs of the required model inputs. Second, on the “Global” dataset, all geotechnical models performed significantly better than the Zhu et al. (2017) model, with the latter performing only somewhat better than random guessing with an *AUC* of 0.55. This might be expected, given: (i) the variation of geomorphic, topographic, and climatic environs in a global dataset; and (ii) the challenge, given this variation, of accurately predicting below-ground conditions from above-ground parameters. Inherently, the CPT models - being directly based on subsurface tests - *should* be more portable across environs. Nonetheless, the performance of the seminal Zhu et al. (2017) model in Canterbury demonstrates the provocative potential of geospatial data.

Limitations of Existing Geospatial Modelling Approaches

By way of the study above, possible shortcomings of the Zhu et al. (2017) and Rashidian and Baise (2020) geospatial model, henceforth referred to as RB20, were identified. It should be emphasized that this model has transformed the perception of geospatial modelling for geohazards. Nonetheless, like all models it has shortcomings that could be improved upon. *First*, RB20 was trained directly on outcomes (i.e., observations of ground failure) rather than on the mechanistic causes of those outcomes (i.e., subsurface engineering properties). While this lack of a mechanistic underpinning can be overcome with vast training data (e.g., how voice transcription apps predict words without understanding language), current ground failure inventories are arguably too sparse. Specifically, both “positive” and “negative” cases (i.e., sites with and without observed liquefaction) are needed in which predictor variables span the range of possible values. That is, the parameter space of all predictor variables should be fully populated. Yet, while liquefaction is common in earthquakes,

ground failure inventories are slow to grow (relative, for example, to that of in-situ test data). Given the adopted approach, inadequate training data can result in a divergence from mechanistic principles (e.g., prediction of liquefaction given shaking too weak, from a mechanistic perspective, to induce liquefaction).

Second, RB20 uses just 5-variables. Four represent capacity (distance to surface water; precipitation; and mapped V_{S30} and groundwater depth) and one represents demand (*PGV*). Notably, none of these variables are likely to correlate to the type of soil, or by corollary, to the susceptibility of the soil to liquefaction. This was a common cause of mispredictions identified in the Geyin et al. (2020) study, with RB20 expecting susceptibility if the ground is flat, saturated, and near water. However, such profiles can consist mostly of soils less- or un-susceptible to liquefaction (e.g., clays, peats, or gravels). Moreover, we find in many such cases that geologic maps accurately predict the presence of such soils. An improved model might thus use mapped geologic data, when available, and/or other yet unidentified proxies of soil type. *Third*, and following from the above, RB20 is a traditional regression equation. This method of modelling inherently requires hypotheses of what is believed to matter and how (beliefs that are unnecessary with ML/DL). The efficient prediction of subsurface traits likely requires more than four geospatial variables, yet regression limits the number easily modeled. Algorithmic learning would allow more geospatial predictor variables to be used, with greater potential for those variables/data to be exploited fully.

The Proposed Modelling Approach and its Potential Benefits

This paper proposes a next-generation geospatial model that is driven by algorithmic learning but pinned to an established mechanistic framework. Specifically, ML/DL models will be used to predict *LPI* values computed from in-situ geotechnical tests subjected to a range of hypothetical ground motions. In effect, the models will predict below-ground geotechnical data using above-ground

geospatial information. Multiple models will be developed and ensembled, thereby avoiding large "swings" on account of which model is chosen (as is common in prediction of ground motions, hurricane tracks, etc.). To use these models in the forward direction, the predicted *LPI* values will be input to existing fragility functions (Geyin and Maurer, 2020a) that predict the probability of liquefaction manifestation (i.e., "ground failure") as a function of *LPI*. These functions were trained on a large database of well-documented liquefaction case histories compiled from 24 global earthquakes. Thus, the ultimate output is a predicted probability of ground failure. An illustrative synopsis of the proposed approach is shown in Fig. 1.

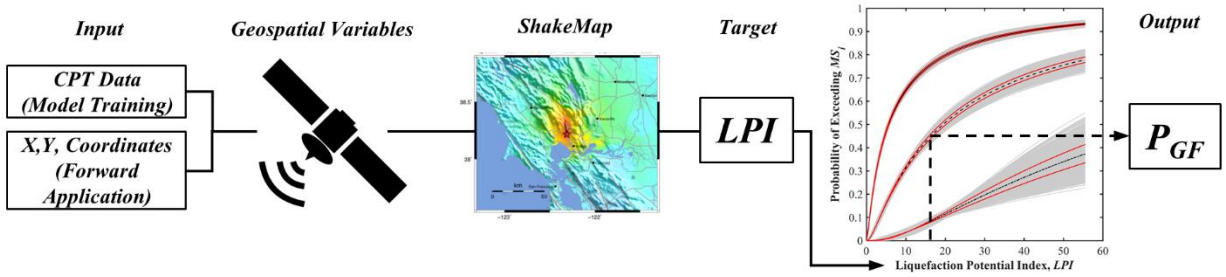


Fig 1. Synopsis of the proposed modelling approach to predicting probability of ground failure (P_{GF}).

This approach has several potential advantages:

- The principal prediction target is transferred from ground failure (with relatively sparse training data) to subsurface measurements (for which the potential training set is vast). Because the location of in-situ tests need not have experienced an earthquake (i.e., be a liquefaction case history), significantly more training data is available. Given the rise of community geotechnical datasets – both internationally and in the U.S. – the gap between the number of subsurface tests

and the number of liquefaction case histories will likely grow. We hypothesize this larger training set will be advantageous, both now and in the future.

- Liquefaction is a physical phenomenon best predicted by mechanics. Much has been learned about liquefaction over the last 50 years. This knowledge is embedded in current state-of-practice liquefaction triggering and manifestation models. We hypothesize that anchoring to these models, which provide a mechanistic foundation, will be advantageous, given their validated ability to model liquefaction response as a function of soil and profile traits (e.g., subsurface stratigraphy, soil density, fines-content, plasticity, saturation, ground motion duration and intensity).
- Whereas liquefaction is best predicted by mechanics, subsurface traits lack theoretical links to above-ground parameters (i.e., geospatial data), but surely correlate to them in complex, interconnected ways. This is a prime problem for ML/DL, which can provide learning insights that are simply unlikely, if not altogether infeasible, with regression approaches. We hypothesize that ML/DL provides the potential for geospatial data to be exploited more fully.
- The models are updated relatively easily as additional training data (in-situ tests) become available. In the short term, some geospatial variables could be viewed by the learning algorithms as relatively unimportant, either because they truly are unimportant, or because there is insufficient training data to elucidate their predictive value. The RB20 model for predicting ground-failure is also retrainable, but we hypothesize that meaningful growth in the ground-failure dataset will take place at a slower pace (e.g., with data from a few events annually that impact a small fraction of earth).

Data and Methodology

In the current effort, two models will be developed using relatively modest sets of training data and predictor variables compiled in the United States (US). As will be discussed, the proposed approach could be extended using additional training data and new predictor variables at regional, national, or global scales. Nonetheless, the models developed herein, which are applicable to the US, will be shown to be as efficient as other models and thus warrant application and further evaluation, even if preliminary in nature.

While several in-situ geotechnical tests could be used within the proposed approach, we choose CPT data given that: (i) it has inherent advantages over other tests upon which liquefaction models have been based (NRC, 2016); (ii) the Geyin and Maurer (2020a) fragility functions were trained on CPT-based case histories; and (iii) a US national CPT database is readily available in native digital format. Specifically, the USGS national database of 1,712 CPTs (USGS, 2021) was adopted for analysis. This dataset provides somewhat well-distributed measurements, as mapped in Fig. 2, in a range of environments, generally in high-seismicity regions. Approximately 5% are from sites where liquefaction case histories were compiled following modern earthquakes. Given the limited dataset, some regions of the US are unrepresented in model development, as shown in Fig. 2. Ultimately, however, tests of the derivative geospatial models are not clearly suggestive of regional bias (i.e., the models perform well in regions with no training data). Nevertheless, it stands to reason that an expanded dataset would result in better models. Of the 1,712 CPTs, 20% were randomly selected and reserved for unbiased model testing, while the remaining 80% were used for model training.

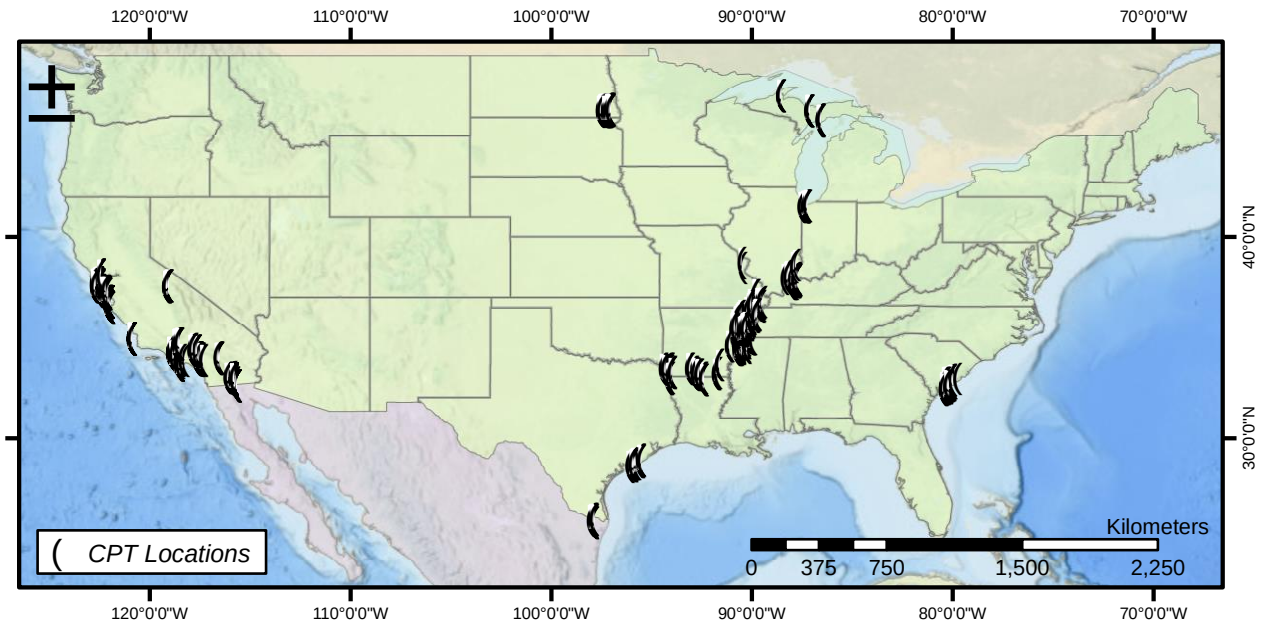


Fig. 2. Spatial distribution of CPT training and test data.

Next, each CPT was subjected to 152 combinations of *PGA* and M_w, with *PGA* ranging from 0.0 g to 1.0 g and M_w ranging from 4.5 to 9.0. These represent loadings of interest that could hypothetically impact a site. With the proposed approach, however, it is irrelevant whether the CPT sites ever experienced an earthquake. For each combination of *PGA* and M_w, the Idriss and Boulanger (2008) CPT liquefaction model was used to predict FS_{liq} versus depth. These predictions were then input to the *LPI* manifestation model, as defined in Eq. 1. All CPT processing and calculations were performed using the software *Horizon* (Geyin and Maurer, 2020b). While different, or additional, triggering and/or manifestation models could be used, the Idriss and Boulanger (2008) triggering model - when used in conjunction with *LPI* - demonstrated an efficiency that was never bested, to a statistically significant degree, by any other model when tested on global case-history data (Geyin et al., 2020). In addition, the magnitude-scaling factor (MSF) inherent to Idriss and Boulanger (2008) is soil-independent, whereas other triggering models (Boulanger and Idriss, 2014; Green et al., 2019)

have MSF's that vary with depth depending on the inferred relative density. The more predictable scaling of computed *LPI* with increasing M_w was deemed advantageous for modelling, given the limited training data utilized herein. A subsequent study could explore the use of other CPT models, although prior testing of such models (Geyin et al., 2020) suggests the efficacy of the resulting product would be very similar.

Given the 1,712 CPTs and 152 combinations of seismic loading, a total of 260,224 *LPI* values were computed. These values are plotted in Fig. 3 as a function of magnitude-scaled *PGA* ($PGA_{M7.5}$), as computed by Idriss and Boulanger (2008), and form the primary prediction target of the proposed modelling approach. At sites of high liquefaction hazard (i.e., thick deposits of saturated, loose sand), *LPI* increases rapidly with $PGA_{M7.5}$, whereas at sites of low hazard (i.e., sites devoid of soil susceptible to liquefaction), *LPI* may remain near zero for all $PGA_{M7.5}$. The goal of geospatial modelling, in effect, is to predict the relationship between *LPI* and $PGA_{M7.5}$ without subsurface data.

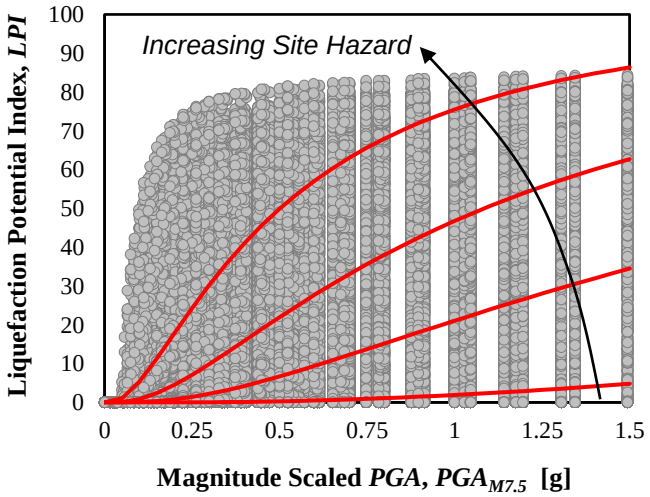


Fig. 3. *LPI* versus $PGA_{M7.5}$. Plotted are 260,224 *LPI* values computed from 1,712 CPTs subjected to 152 different levels of seismic loading.

Ten geospatial predictor variables were next compiled at the coordinates of each CPT. These consisted of: predicted V_{S30} (Heath et al., 2020); predicted ground water depth (Fan and Miguez-Macho, 2020); measured distance to river (Lehner et al., 2006) and measured distance to coast (NASA, 2012); predicted depth to bedrock (Shangguan et al. 2017); measured annual precipitation (Fick and Hijmans, 2017); and the predicted (binomial) presence of unconsolidated soil, sandy soil, clayey soil, and silty soil, as obtained from the USGS National Geologic Map compilation (Horton et al. 2017). The goal of these ten variables is to correlate to the subsurface conditions which give rise to low or high *LPI*.

Using the training set (80% of CPTs), models were next developed to remotely predict *LPI* as a function of *PGA* and M_w, which may be viewed as "demand" variables, and of the ten geospatial variables, which may be viewed as "capacity" variables. The latter can be compiled at national scale in advance of model application. The former are available at regional scale minutes after an earthquake (e.g., via a single "ShakeMap" file (Wald et al., 2005)) or for various future earthquake scenarios. Like other geospatial models, this gives the model near-real-time functionality, such that ground failure can be predicted at regional scale minutes after an event. Various ML/DL techniques were explored, including Gaussian process models (e.g., Rasmussen and Williams, 2006), support vector machines (SVM) (e.g., Vapnik, 1995), decision trees (e.g., Rokach and Maimon, 2008), model ensembles with bagging, gradient boosting, or random forests (e.g., Breiman, 1996; Piryonesi et al., 2021; Ho, 1998) and neural networks (e.g., Glorot et al., 2010). In general, modelling approaches that are easier to interpret tend to have lower predictive accuracy (e.g., single decision trees, support vector machines), while those with higher accuracy (e.g., neural networks, or ensembles of decision trees) are typically very complex to interpret. Each approach has numerous options and internal parameters (i.e., "hyperparameters") (e.g., neural net optimization algorithm, activation function,

layer quantity and size; regression tree leaf size; Gaussian basis and kernel functions; SVM kernel scale and box constraint). Once promising models were identified, hyperparameter optimization was employed, such that the hyperparameter values that minimized the model error were identified via an automated optimization scheme. 5-fold cross validation was used to control overfitting, as is common in model development. Additionally, training and test performance metrics were compared for signs of overfitting (i.e., significantly better training performance than test performance). In general, models with slightly lower accuracy but little-to-no sign of overfitting were favored over models that achieved the highest training accuracy but with suspicion of overfitting. Because many ML/DL algorithms either require or perform better when variables have a Gaussian distribution, all predictors were BoxCox transformed (Box and Cox, 1964) and normalized to have values between 0 and 1.

Results and Discussion

Using the aforementioned methodology with relatively modest sets of training data and predictor variables, several dozen preliminary models were trained. Of these, two were ultimately adopted for further implementation and testing. The first model is a boosted ensemble of decision trees, wherein numerous relatively weak models are coalesced to form one high-quality model. For brevity, we henceforth refer to it as the "ML model." The theory and algorithm underlying this approach – which is commonly included in machine learning toolkits (e.g., Scipy, TensorFlow) – is explained in detail by Friedman (2001). An excellent overview of its practical implementation is provided by Elith et al. (2008). The growth of a decision tree involves establishment of recursive binary splits, such that specific combinations of model inputs map to a predicted output. However, because a single tree is prone to overfitting and tends not to be very accurate, models that ensemble many decision trees are preferred. In "gradient boosting", a strong learner is sequentially built from weak learners, wherein each tree attempts to diminish the errors of the previous tree by gradually increasing emphasis on

observations poorly predicted by the ensemble. While gradient boosting is slow, it generally produces a more accurate model compared to other assembling algorithms (e.g., bagging or random forests) (Piryonesi et al., 2021). With respect to performance, the ML model achieved a mean absolute error (MAE) (*LPI* units) of 3.58 and 3.72 on the training and test sets, respectively, as summarized in Table 2. As discussed previously, the unbiased test set consists of *LPI* data from CPT sites unknown to the model during training.

The second model is a deep (7-layer) artificial neural network (DNN), which we henceforth refer to as the "DL model." With roots in the 1980's (e.g., Hopfield, 1982), this now ubiquitous approach mimics the perceived structure of the human brain, with layers of interconnected nodes. At the most basic level, DNNs have four components: inputs, weights, a threshold, and an output. Connections between nodes are modelled as weights, such that positive and negative weights indicate excitatory and inhibitory connections, respectively. If the output from an individual node is above a specified threshold, the node is activated, sending data to the next layer of the network. An activation function then controls the amplitude of the output at each node. As DNNs have multiple layers, the above process is repeated multiple times, with each layer potentially passing information from the previous layer to the next. During training, the weights are iteratively adjusted to optimize model performance. Similar to the ML model, DNNs are quite convoluted, rendering simple interpretations of the inner workings infeasible, since single node weights have little physical meaning, and since millions of connections may be present in a model. As shown in Table 2, the DL model achieved a MAE of 4.13 and 4.20 on the same respective datasets (i.e., it performed slightly worse than the ML model). Given the limited training set and preliminary nature of the ML and DL models, we also create a third "Ensemble model" by averaging the outputs of the ML and DL models. The merging of two models with very different structures could have the effect of "stabilizing" predictions and, conceivably,

provide benefits unrealized during testing. As shown in Table 2, the ensemble performs better than the DL model and worse than the ML model, although all are relatively similar when considering the range of the *LPI* domain.

Table 2. Summary of model performance (mean absolute error) on the training, test, and overall datasets.

Model	Mean Absolute Error (*LPI* Units)			Mean Absolute Error (Probability Units)		
	Training	Test	Overall	Training	Test	Overall
ML	3.5814	3.7237	3.6642	7.1691	7.6482	7.3698
DL	4.1329	4.2097	4.175	8.6918	8.9609	8.7894
Ensemble	3.743	3.8499	3.8039	7.6491	8.0338	7.7967

Following prediction of *LPI* via the ML, DL, or Ensemble models, probabilities of ground failure can be computed using the Geyin and Maurer (2020a) fragility functions, which are conditioned on *LPI*. As an example, the test and training set performance is shown in Fig. 4 for the ML model. Here, the "predicted probability" is the output when *LPI* is predicted via the geospatial ML model, whereas the "actual probability" is that when *LPI* is computed from the CPT data. The ML model's MAEs of 3.58 and 3.72 (*LPI* units) on the training and test sets translate to MAEs of 7.17 and 7.65 in probability units (Table 2). As seen in Fig. 4, the model is unbiased on the training set, but exhibits some bias on the test set, such that the predicted probability has an average tendency to be greater than actual (e.g., by 5% when the actual probability is 50%). This may be attributable to the dataset's modest size, such that the test set has features unrepresented in the training set.

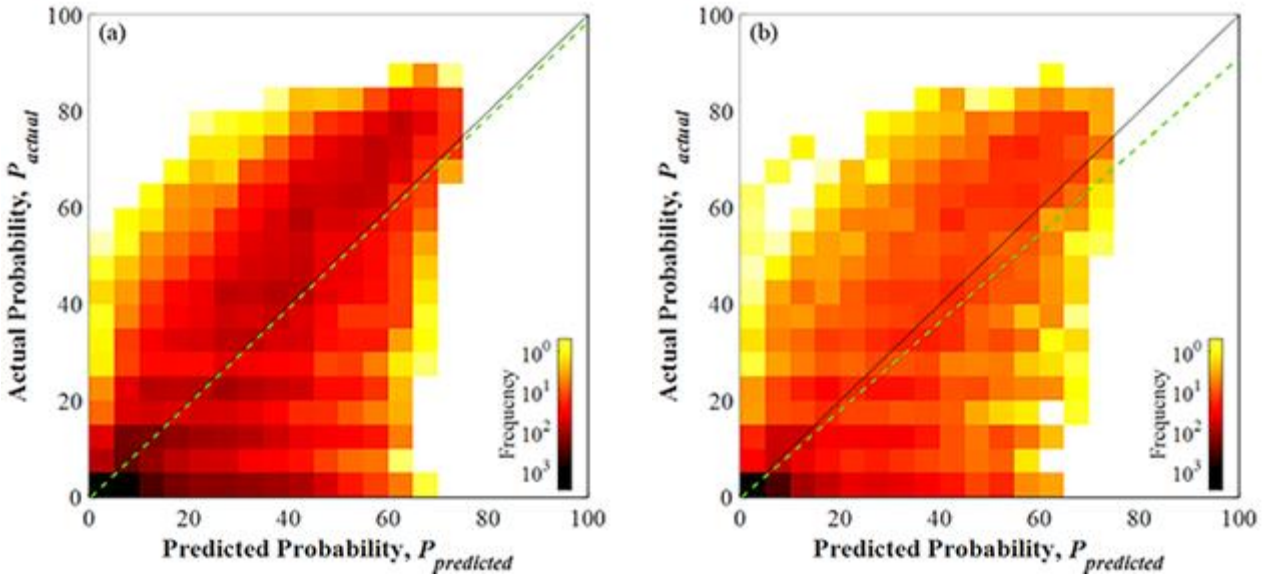


Fig. 4. Probability of ground failure: ML prediction vs. actual for the (a) training dataset; and (b) test dataset.

Model Tests

To demonstrate and test forward predictions at regional scale, the ML, DL, and Ensemble models were next used to predict ground failure in eleven U.S. earthquakes across two types of datasets. Performance was assessed using field observations and compared against the Rashidan and Baise (2020) model in all events. Due to the paucity of recent, well documented U.S. earthquakes outside of California, these tests cover a relatively narrow geographic range. Further testing on future events in other U.S. regions is thus needed. In the first series of tests, regional scale predictions are compared to mapped observations of ground failure in six events: (i) 1989 M_w6.9 Loma Prieta, California; (ii) 1994 M_w6.7 Northridge, California; (iii) 2001 M_w6.8 Nisqually, Washington; (iv) 2003 M_w6.5 San Simeon, California; (v) 2011 M_w5.8 Mineral, Virginia and (vi) 2016 M_w7.1 Ridgecrest, California. In these events, mapped observations of liquefaction-induced ground failure were obtained from the

USGS Ground Failure Database (Schmitt et al., 2017), except for observations in the 2011 Mineral and 2016 Ridgecrest events, which were respectively obtained from Green et al. (2015) and Zimmaro et al. (2020). In these datasets, mapped observations are exclusively "positive" (i.e., a lack of liquefaction is not explicitly mapped). It was therefore assumed that liquefaction did not manifest if none was documented, as has been previously assumed in the development of geospatial models (e.g., Zhu et al., 2017). While this assumption may at times be invalid and inevitably introduces uncertainty, it facilitates rapid, regional-scale testing across a variety of topographic and geomorphic environments. In this regard, we view performance in the context of model comparisons and not as an absolute measure of efficacy. Later, a second dataset containing positive *and* negative observations at discrete sites will be studied.

For each earthquake, a USGS ShakeMap file with all requisite seismic data (i.e., M_w and mapped *PGA* and *PGV*) was obtained in .xml format. The adopted geospatial predictor variables were then compiled across the ShakeMap extents (i.e., the area of perceptible shaking). As an example, these inputs are mapped in Fig. 5 for the 1989 Loma Prieta earthquake. Next, probabilities of ground failure were computed using the three models developed herein and RB20. These probabilities are mapped in Fig. 6 for the 1989 Loma Prieta earthquake, along with observations of ground failure.

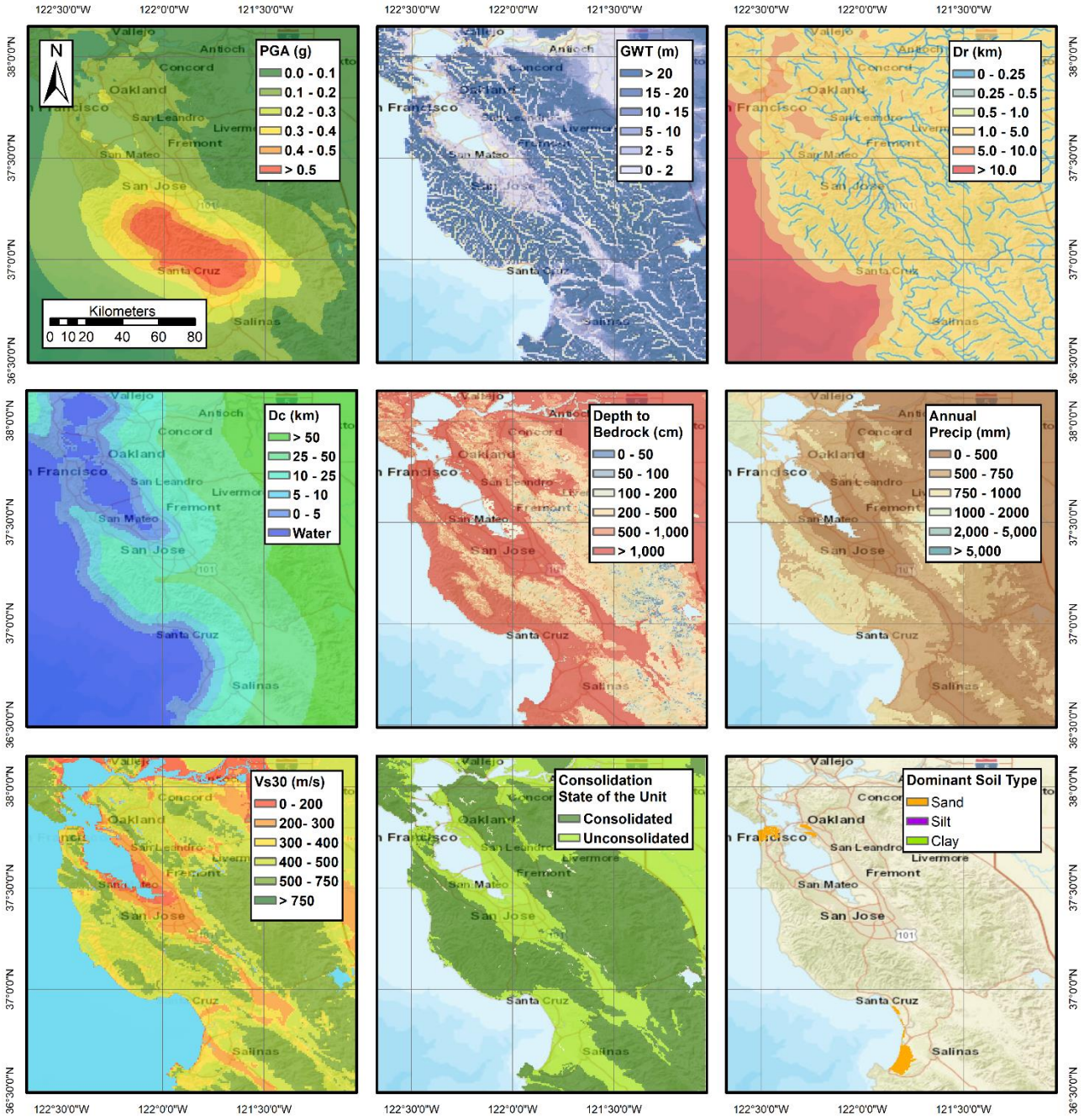


Fig. 5. ML/DL model predictor variables mapped across the area effected by the 1989 Loma Prieta earthquake.

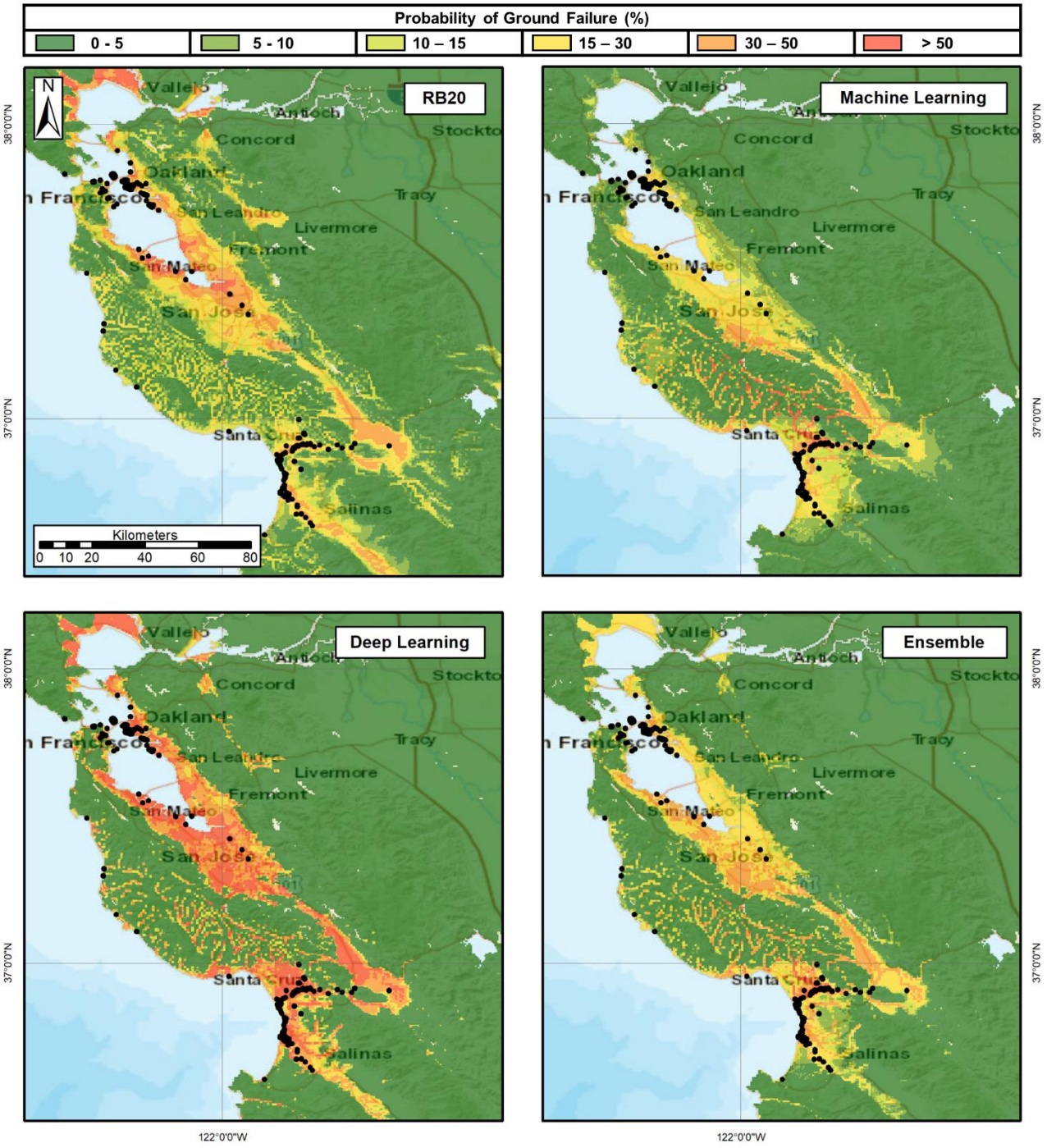


Fig. 6. Probabilities of ground failure in the 1989 Loma Prieta earthquake, as computed by the RB20 and ML, DL, and Ensemble models developed herein. Black dots are observed ground failures.

Model performance was quantified using ROC *AUC* values, as is common for binomial classifiers. Samples were collected on a 100 m by 100 m grid across the ShakeMap extents. Grids wherein ground failures were observed were classified as "positive" and those without any documented evidence as "negative." This resulted in several million data points per event, although the exact quantity depended on the event's area of influence. While the method of geospatial sampling has been shown to influence computed *AUC* values (Lin et al., 2021) (e.g., if an equal number of positive and negative points were sampled instead), we found that *relative* performance was insensitive to this decision (i.e., the best and worst models were the same in each event across a range of sampling techniques). Plotted in Fig. 7 are ROC curves for each model in the 1989 Loma Prieta earthquake from which *AUC* values were computed. Arranged by *AUC*, the best performing models were RB20 (*AUC* = 0.949), Ensemble (*AUC* = 0.945), DL (*AUC* = 0.944), and ML (*AUC* = 0.931). The four models thus exhibited very similar efficiencies, with the Ensemble model slightly outperforming the individual ML and DL models. Following the same methodology, analyses were performed for the 1994 Northridge and 2001 Nisqually events, as mapped in Fig. 8, and for the 2011 Mineral and 2016 Northridge events, as mapped in Fig. 9. A summary of model performance – as quantified by *AUC* – is presented in Table 3 for these events and others yet to be discussed. It can be seen that RB20 outperformed the Ensemble model in three of the six events. Specifically, in Loma Prieta by 0.4%, in Northridge by 3.5%, and in Nisqually by 1.3%. Conversely, the Ensemble model outperformed RB20 in San Simeon by 0.3%, in Mineral by 2.9%, and in Ridgecrest by 1.4%. The models proposed herein thus demonstrate efficacies similar to RB20 for these six events. While these measured differences in performance fluctuate with different sampling techniques, the overall conclusion of similar performance remains the same.

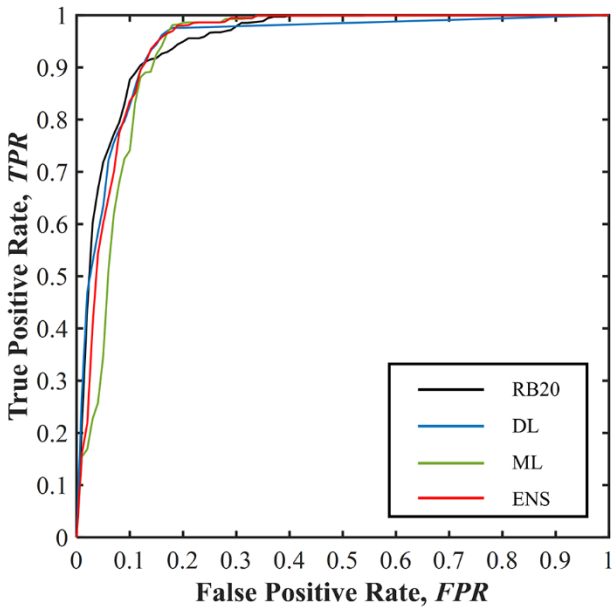


Fig. 7. Receiver Operating Characteristic (ROC) curves for the RB20 ML, DL, and Ensemble models.

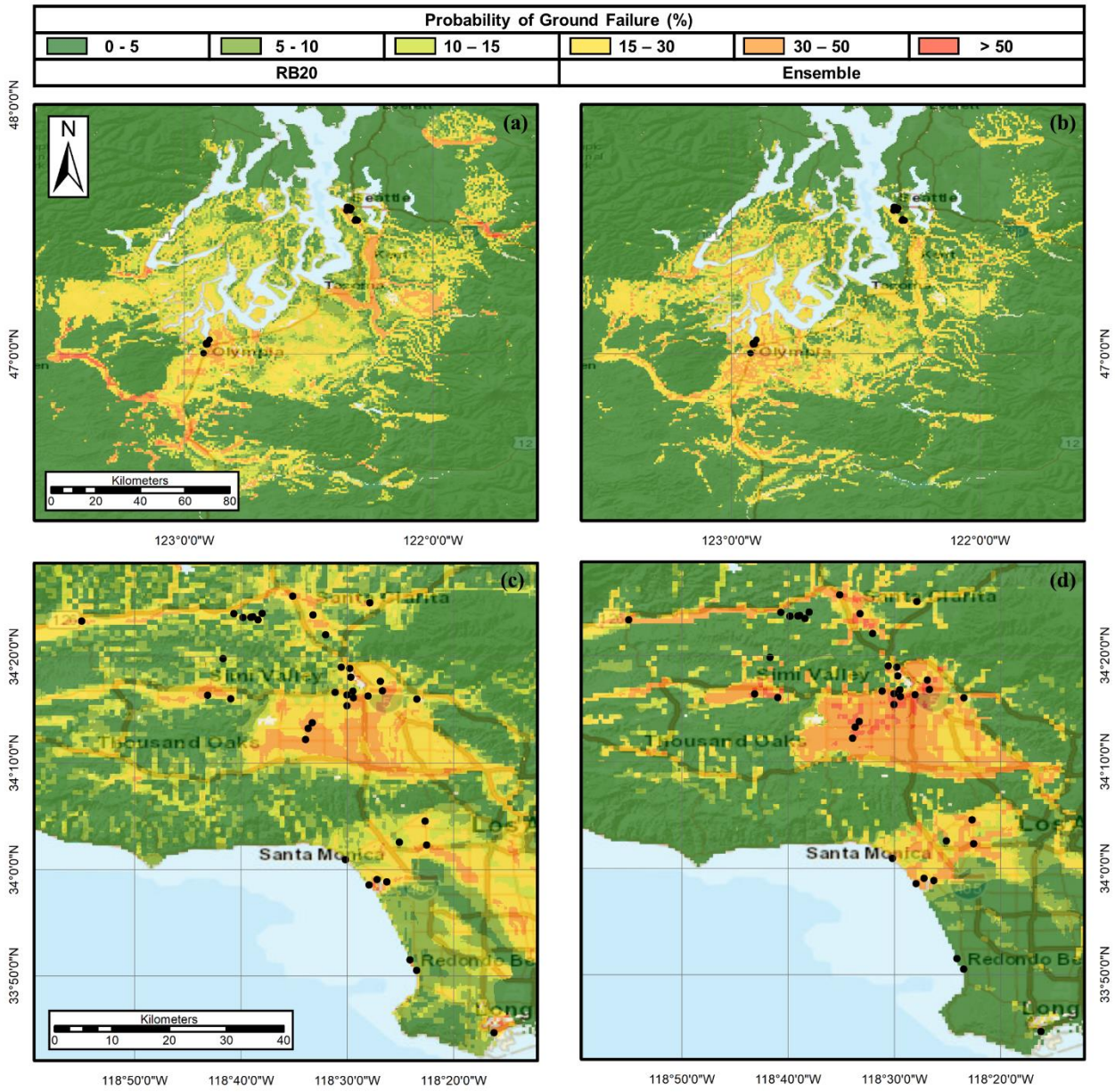


Fig. 8. Probabilities of ground failure in the 2001 Nisqually earthquake, as computed by (a) RB20; and (b) the Ensemble model; and in the 1994 Northridge earthquake, as computed by (c) RB20; and (d) the Ensemble model. Black dots are observed ground failures.

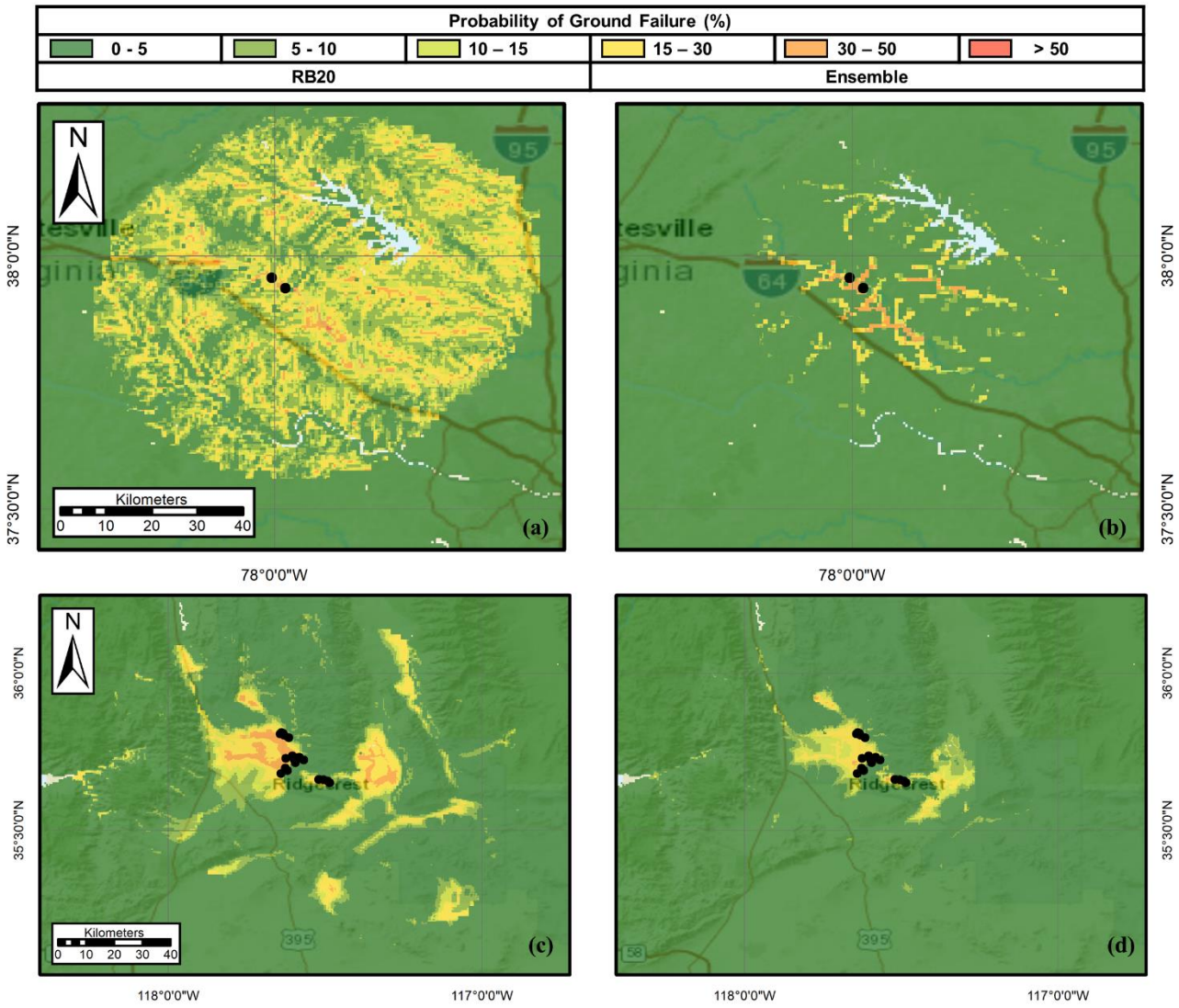


Fig. 9. Probabilities of ground failure in the 2011 Mineral earthquake, as computed by (a) RB20 and (b) the Ensemble model; and in the 2016 Ridgecrest earthquake, as computed by (c) RB20; and (d) the Ensemble model. Black dots are observed ground failures.

Table 3. Summary of *AUC* values for events and datasets described in the text, as computed for the ML, DL, and Ensemble models developed herein, and for the RB20 model.

Model	Dataset 1						Dataset 2
	Loma Prieta	Northridge	Nisqually	San Simeon	Mineral	Ridgecrest	109 Case Histories
DL	0.944	0.803	0.931	0.665	0.655	0.992	0.682
ML	0.931	0.812	0.92	0.98	0.733	0.945	0.765
Ensemble	0.945	0.813	0.933	0.979	0.732	0.992	0.734
RB20	0.949	0.848	0.946	0.976	0.703	0.978	0.504

Additional observations from Figs. 6-9 and Table 3 are as follows. *First*, with respect to model bias, the RB20 model was either originally trained (by Zhu et al., 2017) or later calibrated (by Rashidian and Baise, 2020) using the same observational data adopted herein for testing, excluding the Ridgecrest earthquake data which postdated Rashidan and Baise (2020). In contrast, the ML/DL models were trained on CPTs from areas affected by the Loma Prieta, Northridge, and San Simeon events, but not directly on the field observations adopted for testing. Moreover, the Nisqually, Mineral, and Ridgecrest earthquakes provide completely blind tests of the ML/DL models since no data from these events/regions were included in training. While a more rigorous analysis of bias is not undertaken, nor critical to the thesis of this study, we nonetheless note that the preceding tests were arguably biased in favor of RB20.

Second, it was observed that the DL model is relatively sensitive to predicted water table depth, as compared to the ML and RB20 models. In this regard, erroneous predictions by the DL and Ensemble models were often associated with erroneous expectations of the ground water depth. As an example, predictions in the 2003 San Simeon earthquake by the DL and RB20 models are mapped in Fig. 10. Specifically, an area near the towns of Oceano and Grover Beach, California is shown, where numerous ground failures were observed as mapped in Fig. 10. Because the Fan and Miguez-

Macho (2020) groundwater model predicts a depth of ~20 m beneath the northernmost features, the DL model predicts near-zero probability of ground failure, whereas RB20 generally predicts a probability of 5-15%. Due largely to this behavior, the DL and RB20 models had respective *AUC*s of 0.665 and 0.976 for this event. To assess the influence of more accurate inputs, nearby well measurements were obtained from the California Department of Water Resources (DWR, 2020), indicating that groundwater is shallower in this area than expected by Fan and Miguez-Macho (2020). Using this input, the models were rerun, as mapped in Fig. 10. While the RB20 and ML models correspondingly displayed slight improvements (~1% increase in *AUC*), the DL model's *AUC* increased nearly 30% to 0.990. Similar behavior could be observed in other events at lesser scale, from which we conclude that the performance of the DL model would likely improve with more accurate groundwater maps.

Third, considering the three models developed herein, the Ensemble model outperformed both the ML and DL models in three events. The ML model performed best in two other events, and in the last event the DL and Ensemble models tied for best performance. Considering all tests (i.e., both the blind prediction of *LPI* and now the regional scale prediction of ground failure), the DL model lacks statistical support for individual use. Accordingly, and in conjunction with the DL model's sensitivity to groundwater data, we recommend adoption of the ML or Ensemble models. Ultimately, additional tests in other events, and ideally additional model training and improvements, are needed before one model is recommended over another.

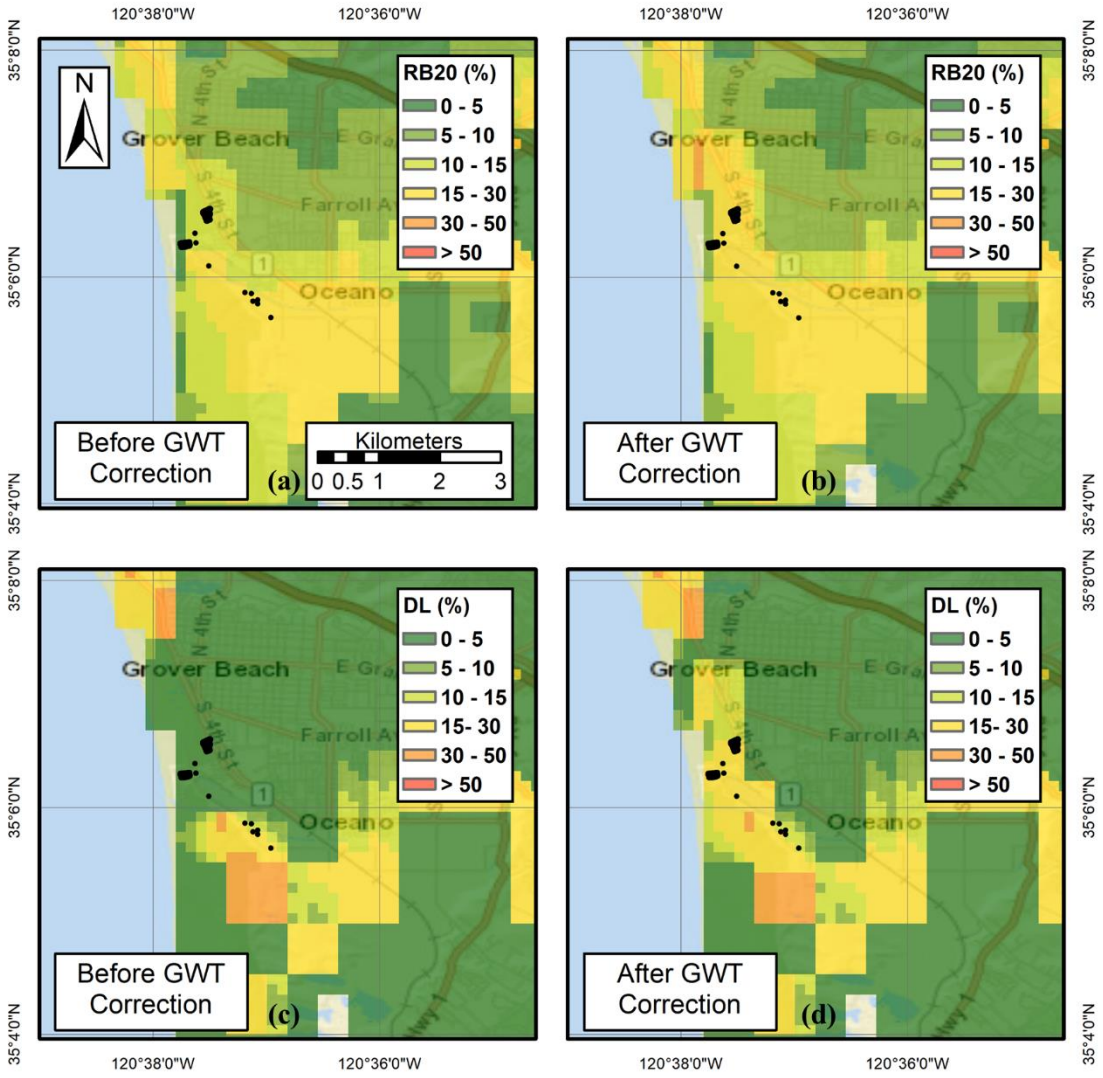


Fig. 10. Probabilities of ground failure in the 2003 San Simeon earthquake, with and without correction of measured ground water table (GWT) depths: (a) RB20 before GWT correction; (b) RB20 after GWT correction; (c) DL model before GWT correction; and (d) DL model after GWT correction. Black dots are observed ground failures.

Towards that end, a second, smaller dataset of 109 well-documented liquefaction case history sites was also used to test performance. These cases, which consist of both positive *and* negative observations, were sourced from Geyin and Maurer (2021a), who compiled CPT-based case histories

from eight U.S. events: (i); 1971 M_w7.6 San Fernando; (ii) 1979 M_w7.6 Imperial Valley; (iii) 1981 M_w5.9 Westmoreland; (iv) 1983 M_w6.9 Borah Peak; (v) 1987 M_w6.2 Elmore Ranch; (vi) 1987 M_w6.5 Superstition Hills; (vii) 1989 M_w7.6 Imperial Valley; and (viii) 1994 M_w6.9 Northridge. The number of cases per event varies from two (San Fernando and Elmore Ranch) to sixty-seven (Loma Prieta) with a median of seven. The four models were applied to each event and *AUC* values were computed for the composite dataset of 109 cases, as given in Table 3. In this analysis, the ML/DL models performed much better than RB20, which had an *AUC* near 0.5, an efficiency akin to random guessing. Of the models developed in this study, the ML model performed best (*AUC* = 0.765), and the DL model performed worst (*AUC* = 0.682). In contrast to the initial series of tests, however, those using this second dataset might be biased in favor of the ML/DL models, since a majority of the 109 sites held CPTs included in the dataset of 1,712 used in model development. While these tests provide another datapoint for consideration, wherein it is known with confidence that the field observations are correctly classified, we prefer not to glean new conclusions, given the possibility of bias and small size of the dataset.

Lastly, to assess whether the findings presented thus far might change with consideration of finite-sample uncertainty, P-values were computed using the method of DeLong et al. (1988) to assess whether differences in *AUC* could result by chance (i.e., due to limited field data) and not because one model is more efficient than another. The P-values computed by this approach are probabilities that two *AUC* samples could have come from the same distribution. Since this approach requires *AUC* normality, Anderson-Darling and Lilliefors tests (Anderson and Darling 1952; Lilliefors 1967) were used to confirm that all samples came from a normal distribution. P-values were computed to compare each model to all others in the six regional analyses, and in the dataset of CPT case histories. These values, which indicate whether differences in model performance are statistically significant, are

presented in Table 3. A significance level of 0.05 was adopted, such that P-values below 0.05 were deemed significant. Small P-values can be expected when: (i) differences between two *AUC* values are large; or (ii) the uncertainties of *AUC* values are small. Using this criterion, Table 3 compares all model pairs and identifies which, if any, is significantly better. Notable observations from Table 3 are: (i) in the initial series of six events at regional scale, the Ensemble model was significantly better than RB20 in two events (Mineral and Ridgecrest), RB20 was significantly better than the Ensemble model in one event (Northridge), and the two models were statistically indifferent in the remaining three (Loma Prieta, Nisqually, and San Simeon); and (ii) in the analysis of CPT case history sites, the Ensemble, ML, and DL models were significantly better than RB20. Collectively, these results suggest that the ML/DL models, which were trained on a modest dataset, predict ground failure with similar efficiency as RB20 and thus warrant application and further evaluation.

Table 4. [Color] P-value matrix to compare model performance.

Statistically Better	↑	Nisqually			Loma Prieta			San Simeon			Northridge			Ridgecrest			Mineral			CPT Case Histories		
←		DL	RB	ENS	DL	RB	ENS	DL	RB	ENS	DL	RB	ENS	DL	RB	ENS	DL	RB	ENS	DL	RB	ENS
Nisqually	ML	0.484	0.285	0.405																		
	DL		0.221	0.473																		
	RB			0.334																		
Loma Prieta	ML				**0.004**	**0.002**	**0.000**															
	DL					0.138	0.596															
	RB						0.278															
San Simeon	ML							**0.000**	0.370	0.341												
	DL								**0.000**	**0.000**												
	RB									0.321												
Northridge	ML										0.440	0.055	0.613									
	DL											**0.003**	0.279									
	RB												**0.036**									
Ridgecrest	ML													**0.001**	**0.017**	**0.002**						
	DL														**0.015**	0.341						
	RB															**0.016**						
Mineral	ML																0.124	**0.001**	0.666			
	DL																	0.396	0.121			
	RB																		**0.002**			
CPT Case Histories	ML																			**0.009**	**0.000**	0.057
	DL																				**0.000**	**0.004**
	RB																					**0.000**

*Cell values are the P-values (i.e., probabilities) that *AUC* samples from two prediction models could have come from the same parent distribution (i.e., be statistically indifferent). The model with better *AUC*, as reported in Table 3, is indicated via the cell shading. Values less than 0.05 are deemed "significant" and are highlighted via bold font and a red border.

Software Implementation

Arguably, a limitation of any ML/DL model is the lack of a defined analytical expression easily ported and executed via hard copy. By corollary, simple interpretations of model structure and form are also generally lacking. While these detractions may be significant to traditionalists, it is clear the use of algorithmic learning will only grow, given its demonstrated capabilities when provided with large datasets. It is critical, however, that trained ML/DL models be provided as code, ideally in a format accessible to a broad userbase. Despite this necessity, enumerable ML/DL models in the field of geotechnics and geohazards have been published without code, meaning – whether accidentally or intentionally – that the respective authors describe the development of a model that will never be used.

To facilitate user adoption and evaluation, the ML, DL, Ensemble, and RB20 models were programmed into *RapidLiq* (Geyin and Maurer, 2021b), a new Windows software program with a simple-to-use interface (Fig. 11). While the Rashidian and Baise (2020) model is widely referenced, it is not commonly implemented by individual users, owing to the predictor variables that must first be compiled. These variables, and those of the proposed models, are compiled within *RapidLiq*, making user implementation trivial. The only required input is a ShakeMap of ground-motion intensity measures, either in Extensible Markup Language (.xml) or Geotagged Image File (.tiff) format. The first is easily downloaded from the USGS earthquake catalog (https://earthquake.usgs.gov/earthquakes/search/) minutes after an earthquake, or for numerous future scenario events. The second is a more general, flexible format, allowing for motions from various sources to be analyzed. The software then extracts predictor variables across the ShakeMap extents and outputs geotiff files mapping the probabilities of liquefaction-induced ground failure. These files may be viewed within the software or explored in greater detail using GIS or one of

many free geotiff web explorers (e.g., http://app.geotiff.io/). The software also allows for tabular input, should a user wish to enter specific sites of interest and ground-motion parameters at those sites, rather than study the regional effects of an earthquake. At present, *RapidLiq* operates in the contiguous U.S. and completes predictions within 10 s for most events.

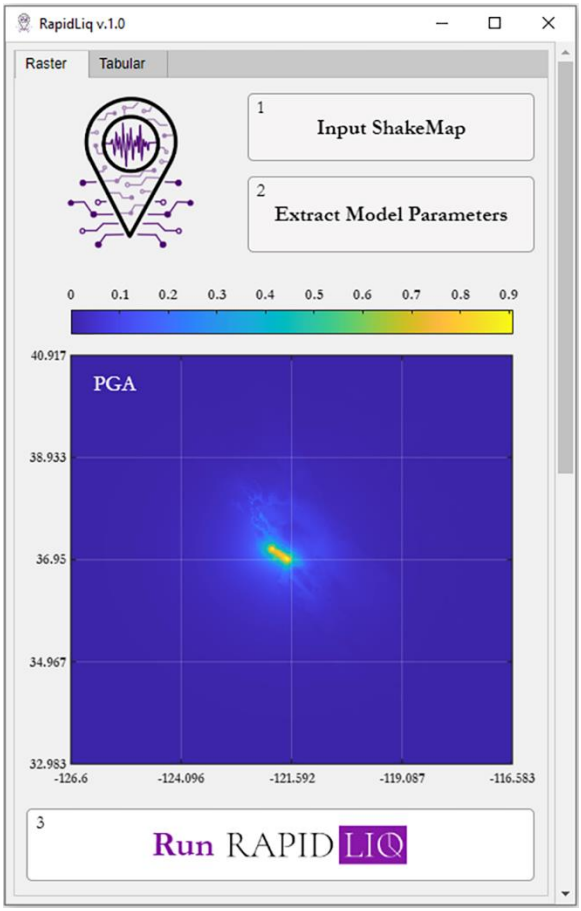


Fig. 11. User interface of *RapidLiq* (Geyin and Maurer, 2021b), which runs the ML, DL, Ensemble, and RB20 models. *RapidLiq* may be downloaded from: https://doi.org/10.17603/ds2-4bka-y039.

Limitations and Uncertainties

Inherently, the presented findings are tied to the data analyzed herein. The applicability of these findings to other earthquakes elsewhere – particularly in regions underrepresented in model training – is unknown. In addition, it should be emphasized that "ground failure," the ultimate prediction target, refers to free field liquefaction-induced surface settlement, cracking, and ejecta on ground that is generally level. Users should understand limitations of the *LPI* manifestation model to predict lateral spreading, which is a distinctly damaging expression influenced by complex subsurface and topographic features. Given that *LPI* and other similar manifestation models may be poor predictors of lateral spreading (e.g., Maurer et al. 2015b; Rashidian and Gillins 2018), the proposed models may likewise predict it poorly. In addition, these models do not explicitly predict damage to specific infrastructure types or assets, which would require detailed site and asset-specific modelling. In this respect, liquefaction could trigger at depth and damage infrastructure without otherwise manifesting, or could manifest without otherwise causing damage. In the present effort, measurement and modelling uncertainties were not considered, and as such, the model outputs should be considered as median probabilities of ground failure. This should not be interpreted to mean that uncertainties do not exist. Among other uncertainties that could be considered in the future, ShakeMap IMs are uncertain; the prediction of *LPI* via geospatial variables is uncertain; and *LPI* is an uncertain predictor of ground failure. Ultimately, additional tests in past or future events are needed to confirm the findings presented herein and summarized below.

Conclusions

This paper presented a new approach to geospatial modelling of soil liquefaction that is driven by algorithmic learning but pinned to a mechanistic framework. In effect, subsurface measurements are predicted remotely within the framing of a popular model for probabilistically predicting ground failure. This merges a body of knowledge built over the last 50 years with the potential of machine and deep learning to predict subsurface conditions remotely. As hypothesized herein, this modelling approach has potential advantages over others used to-date. Using this approach, three models termed ML, DL, and Ensemble were trained on a modest dataset of 1,712 CPTs distributed across the US using a similarly modest set of ten geospatial predictor variables. These models were shown to provide efficient predictions in unbiased, forward application and were tested against the RB20 geospatial model. Collectively, these tests indicate that the proposed models predict ground failure with similar efficiency as RB20 and thus warrant application and further evaluation. The proposed and RB20 models are available in *RapidLiq*, a free Windows program.

Ultimately, significantly more in-situ geotechnical tests are available for model training, both in the U.S. and globally. Whereas ground-failure inventories are likely to grow slowly, with data from perhaps several events per year impacting a small fraction of earth, the subsurface data needed to train the proposed approach exist in massive quantities. These data require compilation across different formats (some requiring digitization) and test types (e.g., CPTs and SPTs), as well as access from various entities, both public and private. However, community geotechnical datasets in New Zealand, Austria and Germany, Italy, and Washington State, for example, currently contain more than 40,000 CPTs. Similar datasets are likely to be created elsewhere. Likewise, the quantity of prospective geospatial predictor variables far exceeds that utilized in this

study. Expanding upon the general approach proposed herein, improved geospatial liquefaction models could thus be developed for regional, national, or global application.

Acknowledgements

This study is based on work supported by the National Science Foundation (NSF), US Geological Survey (USGS), Pacific Northwest Transportation Consortium (PacTrans), and Washington State Dept. of Transportation (WSDOT) under Grant Nos. CMMI-1751216, G18AP-00006, 69A3551747110, and T1461-74, respectively. However, any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and may not reflect the views of NSF, USGS, PacTrans, or WSDOT. In addition, we thank Janset Geyin of the Hacettepe University Interior Architecture and Environmental Design Program for their contributions toward design of the RapidLiq logos and user interface.

Data Availability

All data analyzed in this study is publicly available, as described and referenced in the text. *RapidLiq*, the software program developed to implement the models proposed herein, is available from the DesignSafe data depot at: https://doi.org/10.17603/ds2-4bka-y039.

References

Anderson TW and Darling DA (1952) Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics* 23(2): 193-212.

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.

Box, GEP and Cox, DR (1964) An analysis of transformations. Journal of the Royal Statistical Society, Series B. 26 (2): 211–252.

Breiman, Leo (1996) Bagging predictors. Machine Learning 24 (2): 123–140.

Cubrinovski M and Ishihara K (1998). State concept and modified elastoplasticity for sand modeling, *Soils and Foundations* 38(4): 213-225.

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.

DWR (2020) Periodic groundwater measurements. *California Department of Water Resources*, <https://sgma.water.ca.gov/webgis/?appid=SGMADataViewer#gwlevels> last accessed June 2020.

Elith J, Leathwick JR, & Hastie T (2008) A working guide to boosted regression trees. *Journal of animal ecology 77*(4): 802-813.

Fan Y, Li H, Miguez-Macho G (2013) Global patterns of groundwater table depth. *Science* 339 (6122): 940-943, doi:10.1126/science.1229881

Fan Y, Li H, Miguez-Macho G (2020) Updated dataset for global patterns of groundwater table depth. <http://thredds-gfnl.usc.es/thredds/catalog/GLOBALWTDFTP/catalog.html> last accessed July 2021.

Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874.

Federal Emergency Management Agency (FEMA) (2013). *Earthquake Model HAZUS-MH 2.1 Technical Manual*, Washington, D.C.

Fick SE and Hijmans RJ (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37(12): 4302-4315.

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Annals of statistics*: 1189-1232.

Geyin M and Maurer BW (2020a). Fragility Functions for Liquefaction-Induced Ground Failure. *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.

Geyin M and Maurer BW (2020b) Horizon: CPT-based liquefaction risk assessment and decision software. DesignSafe-CI, doi: 10.17603/ds2-2fky-tm46.

Geyin M and Maurer BW (2021a) CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1). DesignSafe-CI. https://doi.org/10.17603/ds2-wftt-mv37.

Geyin M and Maurer BW (2021b) "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." DesignSafe-CI. https://doi.org/10.17603/ds2-4bka-y039.

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, 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*: 10.1177/8755293021996367,

Glorot, X. and Yoshua B. (2010) Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the thirteenth international conference on artificial intelligence and statistics*: 249–256.

Green RA, Lasley S, Carter MW, Munsey JW, Maurer BW, & Tuttle MP (2015) Geotechnical aspects in the epicentral region of the 2011 Mw 5.8 Mineral, Virginia, earthquake. *Geological Society of America, Special Paper 509*: 151-172.

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.

Heath DC, Wald DJ, Worden CB, Thompson EM, & Smoczyk GM (2020) A global hybrid VS30 map with a topographic slope–based default and regional map insets. *Earthquake Spectra 36*(3): 1570-1584.

Ho TK (1998) The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(8): 832–844.

Holzer TL, Noce TE, & Bennett MJ (2011). Liquefaction probability curves for surficial geologic deposits. *Environmental & Engineering Geoscience 17*(1): 1-21.

Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A. 79(8): 2554–2558.*

Horton JD, San Juan CA, and Stoeser DB (2017) The State Geologic Map Compilation (SGMC) geodatabase of the conterminous United States (ver. 1.1). *USGS Data Series 1052*, 46 p.

Idriss IM, and Boulanger RW (2008) Soil liquefaction during earthquakes. *Monograph MNO-12* 2008; Earthquake Engineering Research Institute, Oakland, CA, 261 pp.

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

Kramer SL (2008). *Evaluation of Liquefaction Hazards in Washington State, Report No. WA-RD 668.1*, Washington State Transportation Center, Seattle, Washington.

Lehner B, Verdin K, and Jarvis A (2006) *HydroSHEDS Technical Documentation*. World Wildlife

Fund US, Washington, D.C

Lin A, Wotherspoon L, and Motha J (2021) Evaluation of a geospatial liquefaction model using land damage data from the 2016 Kaikoura earthquake. *Bulletin of the New Zealand Society for Earthquake Engineering. In Review.*

Maurer BW, Green RA, Cubrinovski M, and Bradley BA (2014). Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand. *Journal of Geotechnical and Geoenvironmental Engineering* 140(7): 04014032Maurer BW, Green RA and Taylor ODS (2015a) 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, and Bradley B (2015b) Assessment of CPT-based methods for liquefaction evaluation in a liquefaction potential index framework. *Géotechnique* 65(5): 328-336**.**

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.

NASA (2020) Distance to nearest coastline. *NASA Ocean Biology Processing Group (OBPG).* < https://oceancolor.gsfc.nasa.gov/docs/distfromcoast/> *last ccessed July 2021*.

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.

Piryonesi SM & El-Diraby TE (2021) Using machine learning to examine impact of type of performance indicator on flexible pavement deterioration modeling. *Journal of Infrastructure Systems* 27(2): 04021005.

Rashidian V and Gillins DT (2018) Modification of the liquefaction potential index to consider the topography in Christchurch, New Zealand. *Engineering Geology* 232: 68-81.

Rashidian V and Baise LG (2020) Regional efficacy of a global geospatial liquefaction model. *Engineering Geology*, 105644.

Rasmussen CE and Williams CKI (2006) *Gaussian Processes for Machine Learning*. MIT Press. Cambridge, Massachusetts, 2006.

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

Rokach L and Maimon O (2008) *Data mining with decision trees: theory and applications*. World Scientific Pub Co Inc. ISBN 978-9812771711.

Schmitt RG, Tanyas H, Jessee MAN, Zhu J, Biegel KM, Allstadt KE, ... & Knudsen KL (2017) *An open repository of earthquake-triggered ground-failure inventories* (No. 1064). US Geological Survey.

Shangguan W, Hengl T, de Jesus JM, Yuan H, & Dai Y (2017) Mapping the global depth to bedrock for land surface modeling. *Journal of Advances in Modeling Earth Systems 9*(1): 65-88.

USGS (2021). "Map of CPT Data." *United States Geological Survey. < https://earthquake.usgs.gov/research/cpt/data/> Accessed July 2021.*

Van Ballegooy S, Malan P, Lacrosse V, Jacka ME, Cubrinovski M, Bray JD, O'Rourke TD, Crawford SA, and Cowan H (2014) Assessment of liquefaction-induced land damage for residential Christchurch. *Earthquake Spectra* 30(1): 31-55.

Vapnik V (1995) *The Nature of Statistical Learning Theory.* Springer, New York.

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

Wald DJ and Allen TI (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America* 97: 1379–1395.

Wald DJ, Earle PS, Allen TI, Jaiswal K, Porter K, and Hearne M (2008) Development of the U.S. Geological Survey's PAGER system (Prompt Assessment of Global Earthquakes for Response). *Proc. 14th World Conf. Earthq. Eng*., Beijing, China, 8 pp.

Youd TL and Hoose SN (1977) Liquefaction Susceptibility and Geologic Setting. *Proceedings of the Sixth World Conference on Earthquake Engineering* 3, 2189-2194.

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.

Zimmaro P, Nweke CC, Hernandez JL, Hudson KS, Hudson MB, Ahdi SK, ... & Stewart JP (2020) Liquefaction and related ground failure from July 2019 Ridgecrest earthquake sequence. *Bulletin of the Seismological Society of America 110*(4): 1549-1566.

Ziotopoulou K and Boulanger RW (2016) Plasticity modeling of liquefaction effects under sloping ground and irregular cyclic loading conditions. *Soil Dynamics and Earthquake Eng* 84: 269-283.

APPENDIX H

Publication 8 (Software 1), highlighted below, is provided on the following pages in full. Specifically, a "User Manual" is provided.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology*, In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

***RAPIDLIQ* USER MANUAL**

v.1.0.

Summary

RapidLiq is a Windows software program for predicting liquefaction-induced ground failure using geospatial models, which are particularly suited for regional scale applications such as: (i) loss estimation and disaster simulation; (ii) city planning and policy development; (iii) emergency response; and (d) post-event reconnaissance (e.g., to remotely identify sites of interest). *RapidLiq* v1.0 includes four such models. One is a logistic regression model developed by Rashidian and Baise (2020), which has been adopted into United States Geological Survey (USGS) post-earthquake data products, but which is not often implemented by individuals owing to the geospatial variables that must be compiled. The other three models are machine and deep learning models proposed by Geyin et al. (2021). All necessary predictor variables are compiled within *RapidLiq*, making user implementation trivial. The only required input is a ground motion raster easily downloaded within minutes of an earthquake, or available for enumerable future earthquake scenarios. This gives the software near-real-time capabilities, such that ground failure can be

predicted at regional scale within minutes of an earthquake. The software outputs geotiff files mapping the probabilities of liquefaction-induced ground failure. These files may be viewed within the software or explored in greater detail using GIS or one of many free geotiff web explorers (e.g., http://app.geotiff.io/). The software also allows for tabular input, should a user wish to enter specific sites of interest and ground-motion parameters at those sites, rather than study the regional effects of an earthquake. *RapidLiq* v.1.0 operates in the contiguous U.S. and completes predictions within 10 seconds for most events.

Installation and Running RapidLiq

Install *RapidLiq* using "RapidLiqv1_Installer.exe." If prompted to do so, extract all zipped contents prior to continuing the installation. In addition, you may be warned by the Windows OS that .exe files from unknown sources could be harmful. Please ignore any such message and click "run anyways." There are no specific system requirements. 4GB of RAM will be enough for most scenarios, but if high resolution rasters with a large land domain will be used (2 million + pixels), at least 8 GB of RAM is recommended. Please note that the software will initiate slowly the first time it is used, since some of the installation process occurs during first use of the software. The software will thereafter open more quickly.

Using RapidLiq

The *RapidLiq* user interface has 2 main tabs that perform the same analyses with different types of inputs. Both use a simple 3-step process wherein: (1) an input file is loaded; (2) model parameters are extracted; (3) analyses are performed.

The first tab (Raster) requires a raster file in either .xml (Extensible Markup Language) or .tif (GeoTagged Image File) format. This option is best for studying the regional effects of an

earthquake. The first format can be downloaded from the USGS earthquake catalog (https://earthquake.usgs.gov/earthquakes/search/) minutes after an event, or from the USGS scenario catalog (https://earthquake.usgs.gov/scenarios/catalog/) for future scenario events. The second format (.tiff) is a more general, flexible format, allowing for motions from various sources to be analyzed. In either case, the software extracts predictor variables across the ShakeMap extents and outputs geotiff files mapping the probabilities of liquefaction-induced ground failure. These output files (saved in the same directory as the input file) may be viewed within the software or explored in greater detail using GIS or one of many free geotiff web explorers (e.g., http://app.geotiff.io/).

The second tab (tabular) allows for tabular input, should a user wish to enter specific sites of interest and ground-motion parameters at those sites, rather than study the regional effects of an earthquake. With this option, an .xlsx input file is required with the following ordered columns: (1) Latitude (WGS84); (2) Longitude (WGS84); (3) Moment Magnitude (*Mw*); (4) Peak Ground Acceleration, *PGA*, in g; and (5) Peak Ground Velocity, *PGV*, in cm/sec. With this option, the software creates an .xlsx output file within the same directory where the input file is located.

Use of the software is next demonstrated via two examples.

Example 1. Input type: .xml Raster

Extensible Markup Language (.xml) ground motion rasters are available from the USGS for past earthquakes (https://earthquake.usgs.gov/earthquakes/search/) and future scenario events (https://earthquake.usgs.gov/scenarios/catalog/). These files contain all necessary seismic data, including earthquake magnitude and the regional distribution of shaking intensities. Once on an event page, navigate to the event's "ShakeMap."

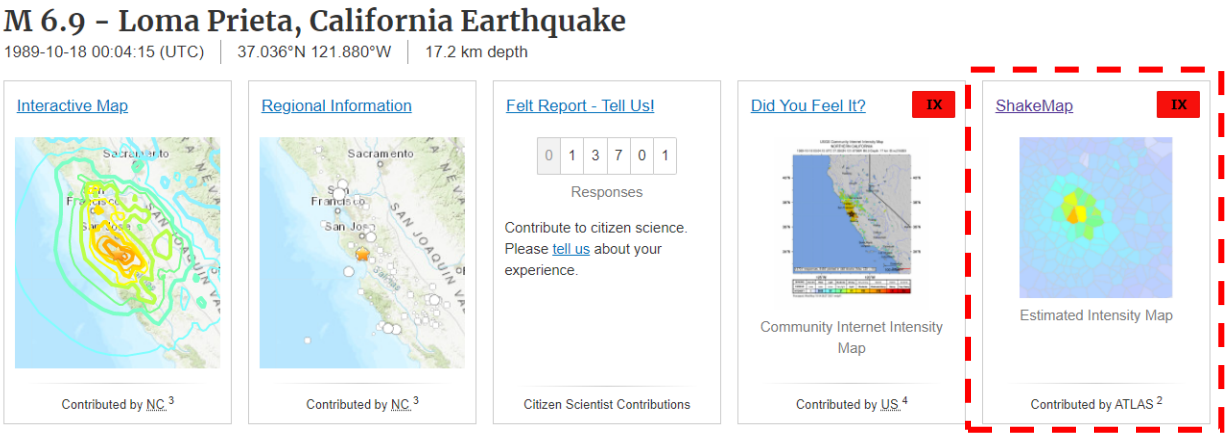


Figure 1. Navigating to ShakeMaps within the USGS event catalogs.

Next, below the map of ShakeMap information, select "Downloads."

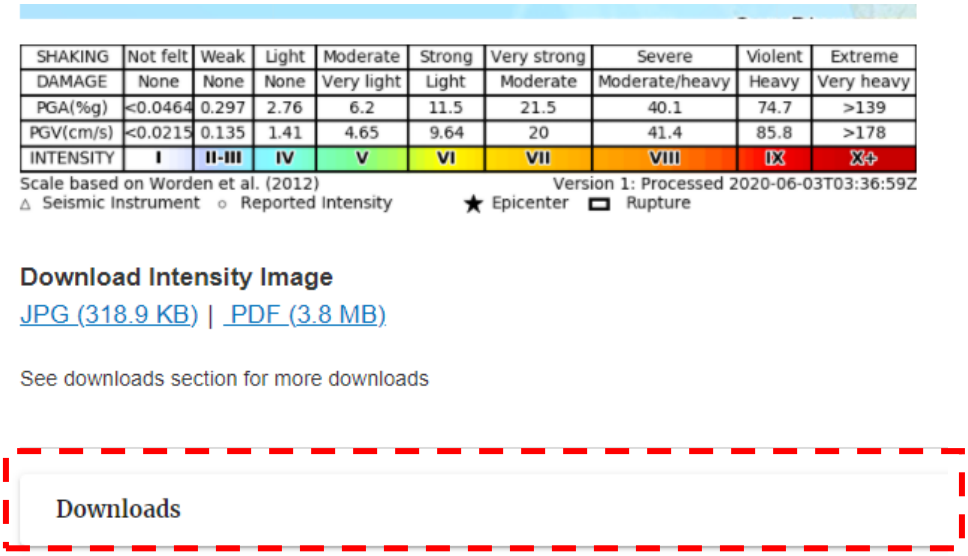


Figure 2. Select the Downloads dropdown menu within a given ShakeMap.

From the dropdown menu, a multitude of download options will appear. The exact position and title of the .xml file may change with time. For the 1989 Loma Prieta earthquake, for example, it appears as:

Figure 3. Location and name of XML grid file for 1989 Loma Prieta, CA, earthquake.

For the 2020 Stanley, Idaho earthquake, it appears as:

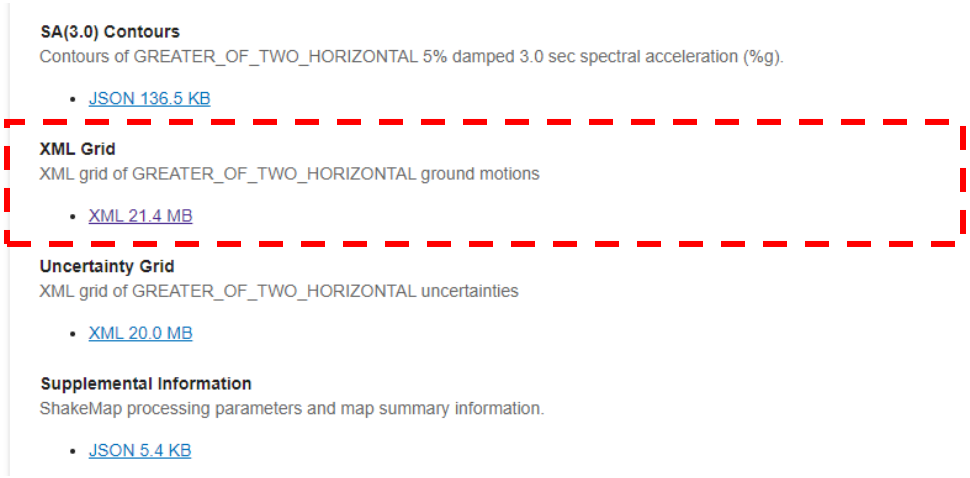


Figure 4. Location and name of XML grid file for 2020 Stanley, ID, earthquake.

Open and save the .xml file to any desired file directory. Please note that an “OutFiles” subfolder will be created in the same directory. Here, the program outputs will be saved.

In *RapidLiq*, under the “Raster” tab, select the “Input ShakeMap” button. Specify the input type as .xml and direct *RapidLiq* to the .xml file of your choosing. Please wait several seconds as the

file is loaded. *RapidLiq* will inform the user whether the input was successful. Press OK to continue.

An interactive map of the *PGA* raster with a scalebar will appear on the screen, as shown in Figure 5.

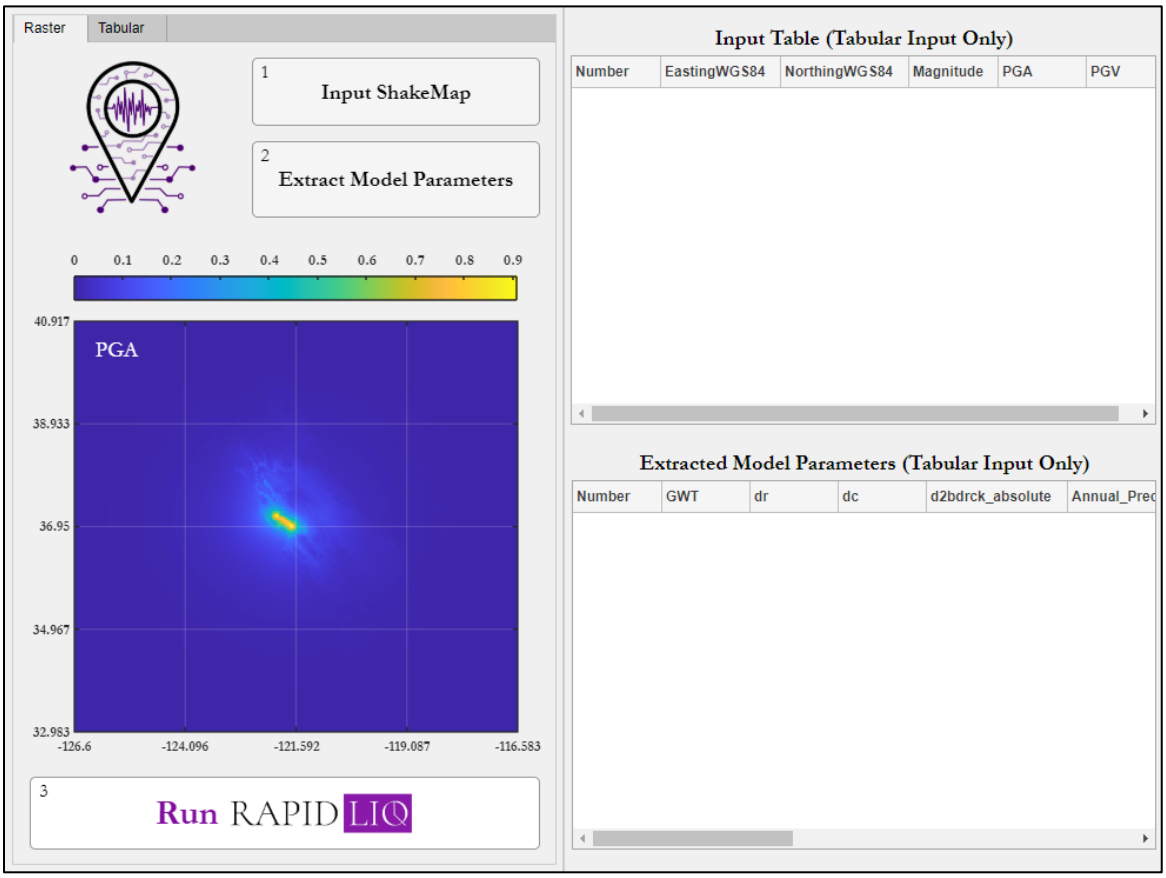


Figure 5. RapidLiq interface after Raster input

Next, select the "Extract Model Parameters" button to gather all model variables over the ShakeMap extents. In *RapidLiq* v1.0, these include predicted V_{S30} (Heath et al., 2020); predicted ground water depth (Fan and Miguez-Macho, 2020); measured distance to river (Lehner et al., 2006) and measured distance to coast (NASA, 2012); predicted depth to bedrock (Shangguan et al. 2017); measured annual precipitation (Fick and Hijmans, 2017); and the predicted (binomial)

presence of unconsolidated soil, sandy soil, clayey soil, and silty soil, as obtained from the USGS National Geologic Map compilation (Horton et al. 2017).

Please wait several seconds until extraction is complete. After the process finishes, a pop-up window should inform the user whether the parameter extraction was successful. Press OK to continue.

To compute predictions of liquefaction-induced ground failure using all available models, select the "Run *RapidLiq*" button. Rapidliq v.1.0 contains four models. These include the logistic regression model of Rashidian and Baise (2020) and the Deep Learning (DL), Machine Learning (ML), and Ensemble (ENS) models proposed by Geyin et al. (2021). All four models are described in detail by Geyin et al. (2021).

All calculations should be performed within several seconds After the predictions are made, a pop-up window informs the user whether the calculation process was successful. Press OK to plot the model predictions on the screen, as shown below in Figure 6. Results from each model, by default, are also saved to the aforementioned "OutFiles" folder as separate geotiff (.tiff) files.

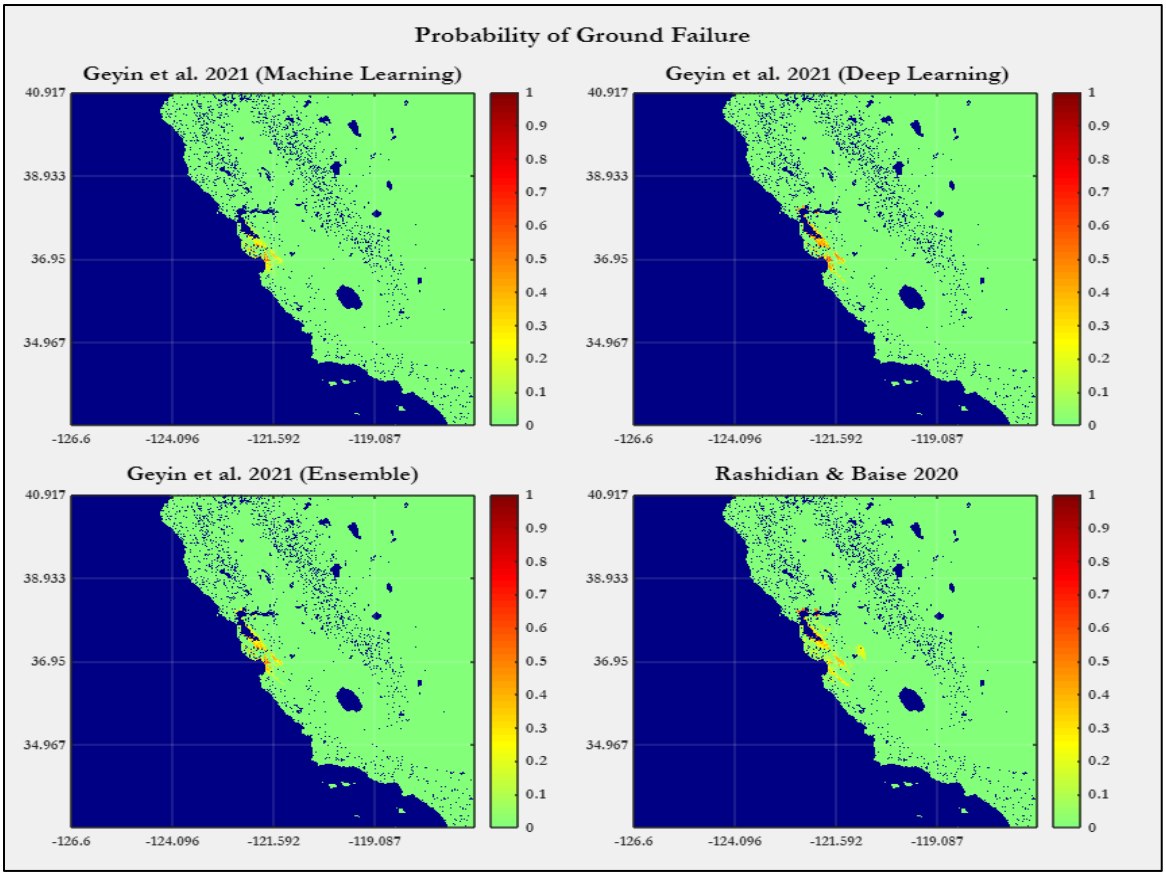


Figure 6. RapidLiq output from all 4 models.

To explore the outputs in greater detail (e.g., to make custom maps, or to further explore and study the predictions), the output files are easily opened in GIS or one of many free geotiff web explorers (e.g., http://app.geotiff.io/). As an example, results for the 2011 Mineral, VA, earthquake are shown in Figure 7 using the geotiff.io web explorer, which has several simple statistical analysis tools built in.

Alternatively, but following a very similar process, *RapidLiq* allows for geotiff (.tiff) input files, which might be obtained from various sources (e.g., ground motion simulations). Using this approach, .tiff files containing *PGA* and *PGV* are separately input, and the earthquake magnitude is entered in a text window.

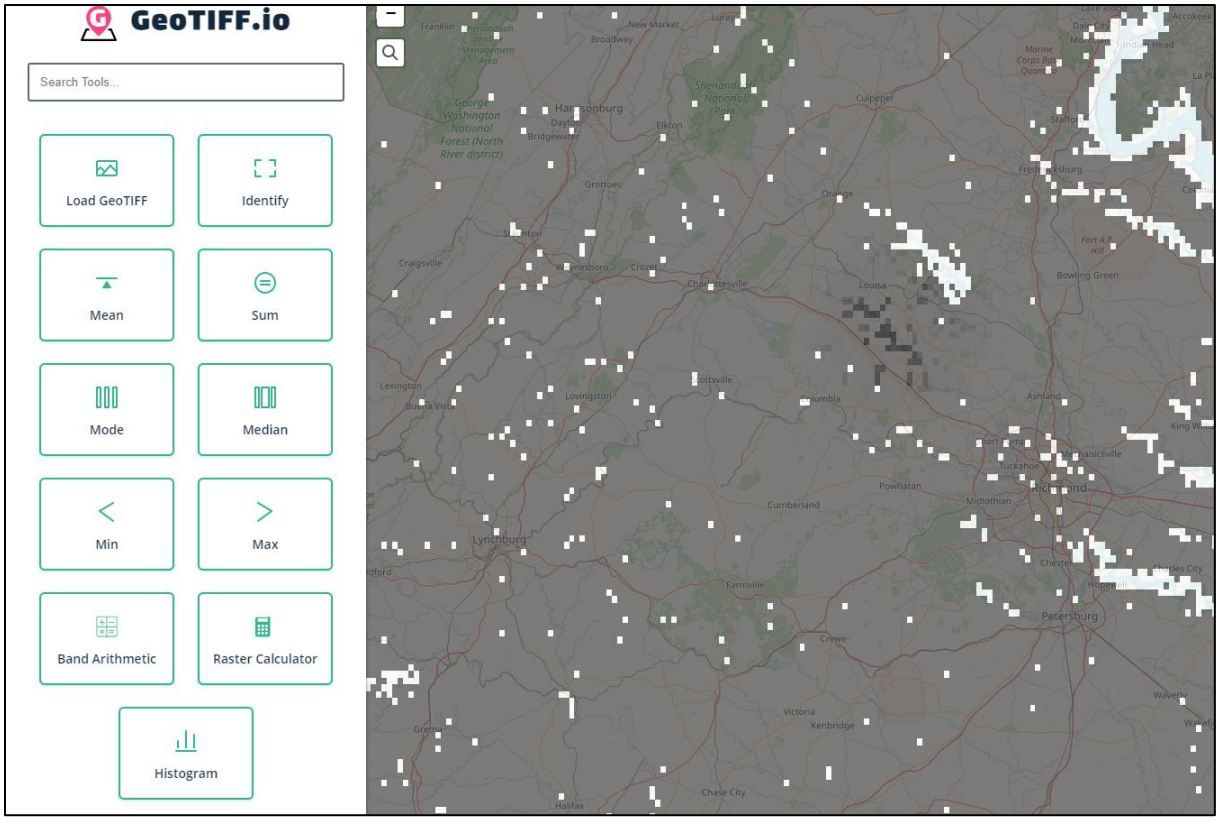
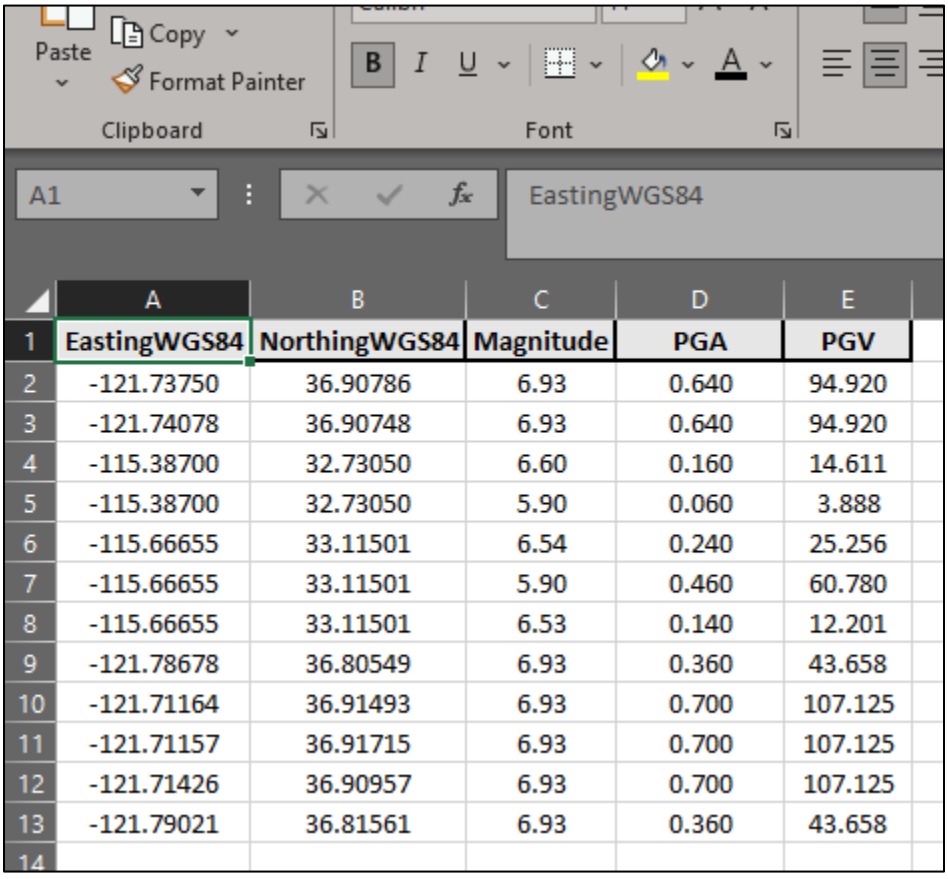


Figure 7. Exploring RapidLiq predictions with the http://app.geotiff.io/) web explorer.

Example 2. Input type: Tabular

The software also allows for tabular input, should a user wish to enter specific sites of interest and ground-motion parameters at those sites, rather than study the regional effects of an earthquake. To use this option, select the "Tabular" tab from the user interface and create a .xlsx file with columnar data in the order of: (1) Latitude (WGS84); (2) Longitude (WGS84); (3) Moment Magnitude (*Mw*); (4) Peak Ground Acceleration, *PGA*, in g; and (5) Peak Ground Velocity, *PGV*, in cm/sec. An example input file is shown below in Figure 8. Import this table to *RapidLiq* by selecting the "Input Table" button. *RapidLiq* will inform the user whether the import is successful and display the tabular data on the upper right side of the user interface. Next, select the "Extract

Model Parameters" button to compile necessary model inputs. This may take several seconds depending on the size of the dataset. After extraction is complete, parameters at each site of interest are displayed on the lower right side of the user interface. Press "Run *RapidLiq*" to compute the predictions using each of the four models. A Table of results will appear and a new timestamped .xlsx file will be created in the directory where the input file is located.



	A	B	C	D	E
1	EastingWGS84	NorthingWGS84	Magnitude	PGA	PGV
2	-121.73750	36.90786	6.93	0.640	94.920
3	-121.74078	36.90748	6.93	0.640	94.920
4	-115.38700	32.73050	6.60	0.160	14.611
5	-115.38700	32.73050	5.90	0.060	3.888
6	-115.66655	33.11501	6.54	0.240	25.256
7	-115.66655	33.11501	5.90	0.460	60.780
8	-115.66655	33.11501	6.53	0.140	12.201
9	-121.78678	36.80549	6.93	0.360	43.658
10	-121.71164	36.91493	6.93	0.700	107.125
11	-121.71157	36.91715	6.93	0.700	107.125
12	-121.71426	36.90957	6.93	0.700	107.125
13	-121.79021	36.81561	6.93	0.360	43.658
14					

Figure 8. Example input table for RapidLiq tabular analyses.

Acknowledgements

This study is based on work supported by the National Science Foundation (NSF), US Geological Survey (USGS), Pacific Northwest Transportation Consortium (PacTrans), and Washington State Dept. of Transportation (WSDOT) under Grant Nos. CMMI-1751216, G18AP-00006, 69A3551747110, and T1461-74, respectively. However, any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and may not reflect the views of NSF, USGS, PacTrans, or WSDOT. In addition, we thank Janset Geyin of the Hacettepe University Interior Architecture and Environmental Design Program for their contributions toward design of the RapidLiq logos and user interface.

References

Geyin, M., Maurer, B.W., and Christofferson, K. (2021). An AI driven, mechanistically grounded geospatial liquefaction model for rapid response and scenario planning. *Soil Dynamics and Earthquake Engineering, In Review.*

Rashidian V and Baise LG (2020) Regional efficacy of a global geospatial liquefaction model. *Engineering Geology*, 105644.

APPENDIX I

Publication 9 (Conference Paper 2), highlighted below, is provided on the following pages in full.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology,* In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

LIFECYCLE LIQUEFACTION HAZARD ASSESSMENT AND MITIGATION

Mertcan Geyin, S.M.ASCE[1]; Brett W. Maurer, M.ASCE[2]; and Sjoerd van Ballegooy[3]

[1] Graduate Research Assistant, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195; e-mail: mgeyin@uw.edu

[2]Assistant Professor, Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195; e-mail: bwmaurer@uw.edu

[3] Technical Director, Tonkin + Taylor Ltd, 105 Carlton Gore Rd, Newmarket, Auckland 1023, New Zealand; svanballegooy@tonkintaylor.co.nz

Abstract

Decisions about risk intervention, such as ground improvement to mitigate soil liquefaction, should be made based on the expected cost and lifecycle benefit (i.e., reduction in loss over the life of an infrastructure asset). In the jargon of performance-based earthquake engineering, vulnerability functions predict loss, and thus, are central to risk management strategy. However, field measurements of loss due to liquefaction have historically been poorly quantified, slowing the development of vulnerability functions requisite for studying the lifecycle economics of ground improvement. Accordingly, aided by engineering and economic data from the 2010-2011 Canterbury, New Zealand, earthquakes - the quantity and combination of which are unprecedented – this study proposes practical functions for predicting liquefaction-induced loss. These functions are: (i) applicable to lightweight structures on shallow foundation systems; (ii) developed as extensions to CPT-based liquefaction analytics already in popular use; and (iii) allow for lifecycle benefits to be quantitatively weighed for a range of intervention strategies. Ultimately, these functions can be used to decide whether, and how, liquefaction should be mitigated, or to demonstrate the economic feasibility of a new type of ground improvement (e.g., bio-inspired and

bio-mediated technologies). Towards this end, the proposed functions are demonstrated via application to a liquefiable site in the SODO district of Seattle, USA, wherein a lifecycle cost-benefit analysis of ground improvement is performed.

Introduction

Lifecycle economic analysis of liquefaction mitigation requires accurate predictions of loss, with and without intervention, for every earthquake that could impact an asset during its lifespan, considering also the probabilities of these earthquakes occurring. Yet, because field-assessments of loss due to liquefaction are historically sparsely available, losses are usually inferred in an informal, heuristic manner in current practice, leading to subjective and potentially inaccurate estimates. Accordingly, motivated by unprecedented engineering and economic data compiled during the 2010-2011 Canterbury, New Zealand, earthquakes, this study proposes practical functions for predicting liquefaction-induced loss, referred to herein as *vulnerability functions*. These functions are applicable to lightweight structures on shallow foundations and are developed as extentions to liquefaction analytics already popular in practice. Using the proposed functions, the expected costs and lifecycle benefits of various intervention strategies can be quantitatively weighed, resulting in more efficient and transparent management of liquefaction risk.

While cost-benefit analyses of seismic retrofits have transformed structural earthquake engineering practice and policy (e.g., Kappos and Dimitrakopoulos 2008; Liel and Deierlein 2013), geotechnical engineers face added challenges in performing analogous studies for liquefaction mitigation. Damage to structural elements (and thus, loss) can often be predicted directly from shaking intensity. In contrast, foundation damage and loss due to liquefaction must be related to the cumulative and complex response of the soil profile, which is related to the liquefaction response of discrete strata at different depths, which is related to the intensity and

duration of shaking, and to the soil properties. Consequently, models are needed to: (a) predict liquefaction triggering at depth; and (b) relate the results of triggering analysis to damage potential at the ground surface.

Collectively, these models may be referred to as *liquefaction demand parameters* (LDPs). In simplest terms, an LDP projects the potential for liquefaction-induced damage and loss (typically at the ground surface) by characterizing the system-response of a liquefiable soil profile. While various LDPs are available, the present study will demonstate the development and utility of the proposed vulnerability functions using one LDP: the Boulanger and Idriss (2014) liquefaction triggering model used in series with the Liquefaction Potential Index (*LPI*) manifestation model. *LPI* is defined as (Iwasaki et al. 1978):

$$LPI = \int_0^{z_{max}} F\left(FS_{liq}\right) \cdot w(z)\, \mathrm{d}z \tag{1}$$

where $F(FS_{liq})$ and $w(z)$ are functions that weight the respective influences of FS_{liq} and depth, z, on damage potential; and FS_{liq} is the factor of safety against liquefaction triggering, computed by Boulanger and Idriss (2014). In Equation 1, $F(FS_{liq}) = 1 - FS_{liq}$ for $FS_{liq} \leq 1$ and $F(FS_{liq}) = 0$ otherwise; $w(z) = 10 - 0.5z$; and $z_{max} = 20$ m. Thus, *LPI* assumes the severity of surface manifestation depends on the cumulative thickness of liquefied strata in a profile's upper 20 m, the proximity of those strata to the ground surface, and the amount by which FS_{liq} in each stratum is less than 1.0. Given this definition, *LPI* can range from zero to 100. Research is ongoing to incorporate an array of other LDPs popular in current geotechnical practice.

In the following, the engineering and economic data used in the current study is first succinctly summarized. This is followed by the development of functions for predicting foundation damage

and loss due to liquefaction. Lastly, use of the functions is demonstrated via application to a liquefiable site in Seattle, USA, wherein a lifecycle cost-benefit analysis of two ground improvement options is performed.

Data

Earthquakes occurring over the last decade in the Canterbury region of New Zealand have resulted in case-history data of unprecedented quantity and quality. A summary of these events, to include tectonic and geologic settings, seismology, and effects, is provided by Quigley et al. (2016). The present study analyzes data from the M_w7.1, 4 Sept. 2010 Darfield earthquake and the M_w6.2, 22 Feb. 2011 Christchurch earthquake. This consists of: (1) inspections of foundation damage due to liquefaction; (2) building-damage ratios, which quantify monetary loss and are defined subsequently; (3) cone-penetration test (CPT) data; and (4) peak ground accelerations (*PGA*s) and ground water depths (*GWD*s) at the location of each inspected foundation.

The unique role of the New Zealand Earthquake Commission as a land insurer, which covers losses from liquefaction, led to the assessment of damage and loss for lightweight residential buildings on shallow foundations. Post-earthquake inspections, as compiled by Tonkin + Taylor Ltd., documented both the mode and magnitude of foundation deformation, allowing for the severity of damage to be classified. Three types of shallow foundation systems were inspected: (1) concrete slab-on-grade with a thickened edge or "ring beam" footing; (2) suspended timber floors with perimeter concrete footings; and (3) suspended timber floors supported by shallow, small diameter piles (~20 cm). Between 1999 and 2015, 50.9% of residential construction in the U.S. utilized slab-on-grade foundations, consistent with type 1; an additional 16.9% utilized crawl-space designs consistent with types 2 and 3 (Bradtmueler et al., 2015). In addition, many lightweight commercial and industrial structures have similar designs. The results of this research

are thus applicable to design practices in the U.S. and elsewhere. For brevity, the current study combines the three foundation types together, though ongoing research will study their performance individually. Seven modes of foundation deformation were documented: stretching, hogging, dishing, racking, tilting, discontinuity, and global settlement. The severity of each mode, for each foundation in each earthquake, was classified based on its magnitude as either: none-to-neglible, moderate, or severe. Many foundations showed multiple modes of damage with different severities. To develop fragility and vulnerability function herein, the "worst" foundation-damage severity was recorded for each foundation in each event (i.e., the greatest severity, independent of deformation mode, if multiple modes were observed).

In parallel with the foundation inspection program, analysts assessed the repair cost and replacement value of each residentual building following each earthquake. The results of these assessments are represented via the Building Damage Ratio (*BDR*), which allows for losses to be compiled for various buildings having different values. *BDR* is defined as:

$$BDR = \frac{Repair\ Cost}{Replacement\ Value} \tag{2}$$

This study utilizes CPT soundings available in the New Zealand Geotechnical Database (NZGD, 2019) and performed at sites where the severity of liquefaction-induced foundation damage and loss were reliably documented, as described above. In compiling this data, CPTs were first removed from the study using several quality-assurance criteria. Extended coverage of CPT data selection and processing is provided in Maurer et al. (2014, 2015). *PGA*s were estimated at each CPT location using the Bradley (2013) procedure, which has been used in many previous Canterbury earthquake studies (e.g., van Ballegooy et al. 2015; Maurer et al. 2019). This procedure geostatistically merges *PGA*s recorded at strong-motion stations with *PGA*s estimated by ground-

motion prediction equations. *GWD*s at each CPT location were sourced from the event-specific regional models of van Ballegooy et al. (2014). Lastly, *LPI* was computed for each CPT in each earthquake using Equation 1, wherein the Boulanger and Idriss (2014) CPT-based liquefaction triggering model was used to compute FS_{liq}.

Prediction of Liquefaction-Induced Damage and Loss

To relate *LPI* to damage and loss, *LPI* values were spatially interpolated at foundation locations via kriging. Inherent to this process, allowances were placed on the the uncertainty of interpolated *LPI* values. Moreover, a maximum distance of 50 m from the center of a building to the nearest CPT was permitted. As a result, foundations where *LPI* values are sparse and/or highly variable are generally not considered. This process, along with the use of other quality-control criteria, resulted in the analysis of 8,700 buildings atop shallow foundations on liquefiable soils. Developed from these analyses are functions that predict liquefaction-induced damage and loss as a function of *LPI*. These tools are respectively termed *fragility* and *vulnerability* functions.

Fragility Functions for Prediction of Liquefaction-Induced Foundation Damage

The probability of foundation damage reaching or exceeding a damage-state, *DS*, given a computed *LPI* value, is herein denoted $F_{DS}(LPI)$ and idealized by a lognormal distribution:

$$F_{DS}(LPI) = \Phi\left(\frac{ln(LPI) - ln\,(\theta)}{\beta}\right) \quad (3)$$

where Φ denotes the Gaussian cumulative distribution function; β is the logarithmic standard deviation; and θ is the distribution median (i.e., the *LPI* value at which a *DS* has a 50% probability of exceedence). Best-fit parameters θ and β were determined by minimizing the squared error (ε^2), as described in Porter (2018) and defined by:

$$\varepsilon^2(\theta,\beta) = \sum_{i=1}^{m_i} n_i \cdot \left(F_{DS_i} - \frac{f_i}{n_i}\right) \quad (4)$$

Where *m* is the number of *LPI* bins; *n* is the number of foundations in each bin, of which *f* reached or exceeded a given *DS*; and *i* is the bin index. The bin size was selected to be inversely proportional to the square root of the sample size (Porter 2018). Using this approach, fragility functions were developed to predict foundation damage, as shown in Figure 1.

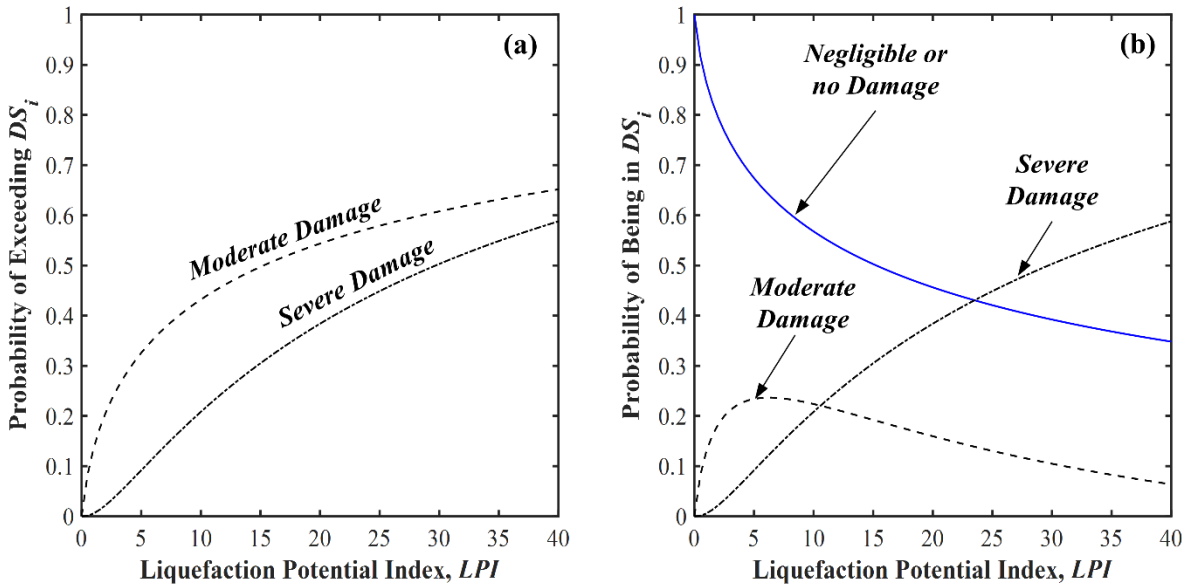


Figure 1. Fragility functions conditioned on *LPI*, as computed using Boulanger and Idriss (2014) liquefaction triggering model, for predicting the probability of **(a)** exceeding a foundation damage-state; and **(b)** being in a foundation damage-state.

Vulnerability Functions for Prediction of Liquefaction-Induced Loss

Vulnerability functions are developed using *BDR* and the formulation of Ramirez et al. (2012):

$$E[BDR|LPI] = \sum_{i=1}^{3} E[BDR|DS_i, LPI] \cdot P[DS_i|LPI] \quad (5)$$

Where *BDR* is as previously defined; *E* denotes the expected value of *BDR*; subscripts refer to the three aforementioned foundation damage-states (i.e., 1 = negligible or no damage; 2 = moderate damage; and 3 = severe damage); and $P[DS_i|LPI]$ describes the suite of previously developed fragility functions shown in Figure 1b. The expression for $E[BDR|LPI]$, shown in Figure 2a, follows a cumulative beta distribution, which is desirable in risk modeling due to its upper and lower bounds. To account for and visualize finite-sample uncertainty, the data used to develop Figure 2a was bootstrapped 1,000 times (i.e., resampled with replacement), with each simulation resulting in a vulnerability curve.

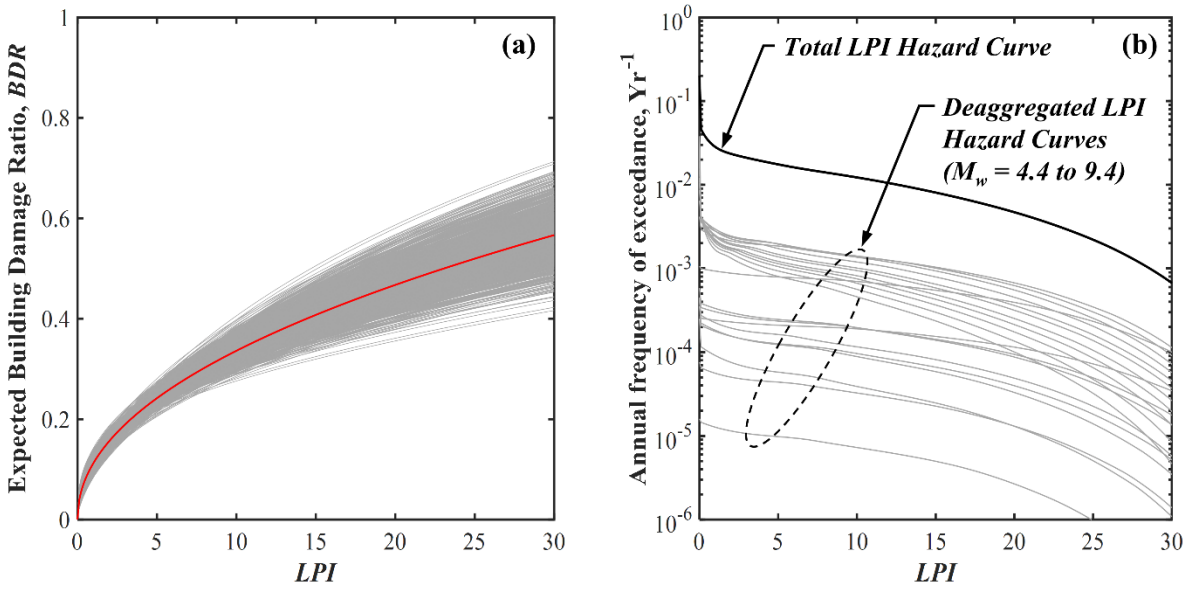


Figure 2. (a) Bootstrapped vulnerability curves (grey) for expected *BDR* as a function of *LPI*; **the** median relation is shown in red. **(b)** *LPI* hazard curves – for the Seattle location described below – deaggreaged on earthquake magnitude, such that the total *LPI* hazard curve is equal to the sum of the *LPI* hazard curves for all magnitudes.

Expected Annual Loss (*EAL*) and Damage-State (*DS*) Return Periods

A building's expected annual loss (*EAL*) due to liquefaction can be computed using: (i) the fragility and vulnerability functions developed herein; (ii) CPT data at the building's location; and (iii) the *LPI* hazard curve, which describes the mean annual rate of *LPI* exceeding a given value (λ_{LPI}) at the building's location. With this approach, *EAL* is computed by integrating over all *LPI* values:

$$EAL = V \cdot \int_{LPI=0}^{\infty} E[BDR|LPI] \cdot \left|\frac{d\lambda_{LPI}}{dLPI}\right| \cdot dLPI \quad (6)$$

where *V* is the replacement value of the building, in dollars; $E[BDR|LPI]$ is defined in Equation 5; and $|d\lambda_{LPI}/dLPI|$ is the absolute value of the derivative of the *LPI* hazard curve, which is computed from CPT data and integration with the *PGA* hazard curve, which describes the mean annual rate of *PGA* exceeding a given value (λ_{PGA}). Specifically, to compute a building site's λ_{LPI}, the *PGA* hazard-curve must first be deaggregated by earthquake magnitude. This requirement results from use of both *PGA* and magnitude as necessary inputs for predicting liquefaction triggering (Boulanger and Idriss, 2014). λ_{LPI} is then computed as:

$$\lambda_{LPI} = \sum_{j=1}^{N_{Mw}} \sum_{i=1}^{N_{PGA}} P\left(LPI > lpi \middle| pga_i, m_j\right) \Delta\lambda_{pga_i,m_j} \quad (7)$$

where N_{Mw} and N_{PGA} are respectively the number of magnitude and *PGA* increments into which the computed seismic hazard is subdivided, and $\Delta\lambda_{pgai,mj}$ is the incremental annual-exceedance rate for intensity measure, pga_i, and magnitude, m_j, which follows an established approach in performance-based liquefaction modelling (Kramer and Mayfield 2007).

To demonstrate this process, a CPT from the South-of-Downtown (SODO) district of Seattle, USA, is analyzed. Deaggregated *PGA* hazard-curve data was obtained for the CPT's location (47.587130, -122.331487) using the USGS Unified Hazard Tool. From Equations 1 and 7, in

conjunction with the Boulanger and Idriss (2014) liquefaction triggering model, the *LPI* hazard curve was computed and is shown in Figure 2b. Lastly, using the *LPI* hazard curve within Equation 6, the *EAL* due to liquefaction is computed for a lightweight structure that: (i) is located at the CPT site coordinates given above; (ii) is built atop a slab-on-grade foundation; and (iii) has a replacement value of $500,000. The computed *EAL* due to liquefaction is $6,529 ± $360. Alternatively, using an integrand similar to that in Equation 6, the return periods (T_R) of different foundation damage-states may also be computed:

$$T_R = 1/\lambda_{DS} = \int_{LDM=0}^{\infty} F_{DS}(LPI) \cdot \left|\frac{d\lambda_{LPI}}{dLPI}\right| \cdot dLPI \quad (8)$$

where F_{DS} is the fragility function for a particular foundation damage state (*DS*), as shown in Figure 1a; λ_{DS} is the annual exceedance rate of the *DS*; and λ_{LPI} is as previously defined. Using equation 8, and given the example above, the return periods of moderate and severe foundation damage are computed as 73 and 182 years, respectively. Thus, and assuming a Poisson process, the probabilities of at least one occurrence of moderate and severe foundation damage in the next 100 years are 74% and 42%, respectively.

Life-Cycle Cost-Benefit Analysis

To demonstrate the full utility of the functions developed herein (and their applicability to bio-inspried technologies), a life-cycle cost-benefit analysis will be performed for the site in SODO considering two hypothetical ground improvement strategies. For simplicity, these hypothetical strategies either double or quadruple the normalized, fines-corrected CPT tip resistance ($q_{c1N\text{-}cs1}$) at the ground surface, with both improvements decreasing linearly to a depth of 20 m, where no improvement is achieved. These strategies are henceforth referred to as Mitigation Options 1 and

2, respectively, where the former has an expected upfront cost of $30,000 and the latter (which achieves more improvement) has an expected upfront cost of $50,000. The relative efficiencies of these intervention strategies can be assessed and compared using various economic decision-modeling tools, made possible by the functions developed in this study. As an example, the cost-efficiency of a strategy may be defined by its cost-benefit difference (*CBD*), such that a *CBD* less than zero indicates an auspicious decision:

$$CBD = Cost - Benefit = Cost - (EAL_o - EAL_m)\left(\frac{1-e^{-rt}}{r}\right) \quad (9)$$

where *Cost* is the upfront expense of the strategy; *Benefit* is the reduction in loss obtained from the strategy; *EAL* is the expected annual loss from liquefaction; subscript *o* refers to the original state; subscript *m* refers to the mitigated state (e.g., with ground improvement); *t* is the planning horizon, or service life, of the infrastructure asset, in years; and *r* is an interest rate for determining the present value of future cash flows. An investment in remediation thus makes sense if the service life of the asset exceeds the time required for *CBD* to reach zero. Moreover, negative *CBD*'s convey the time-dependent net benefit resulting from an up-front investment in remediation. To compute *CBD* values for each intervention strategy, *LPI* hazard curves are first computed for each CPT following the approach taken previously, the results of which are shown in Figure 3a.

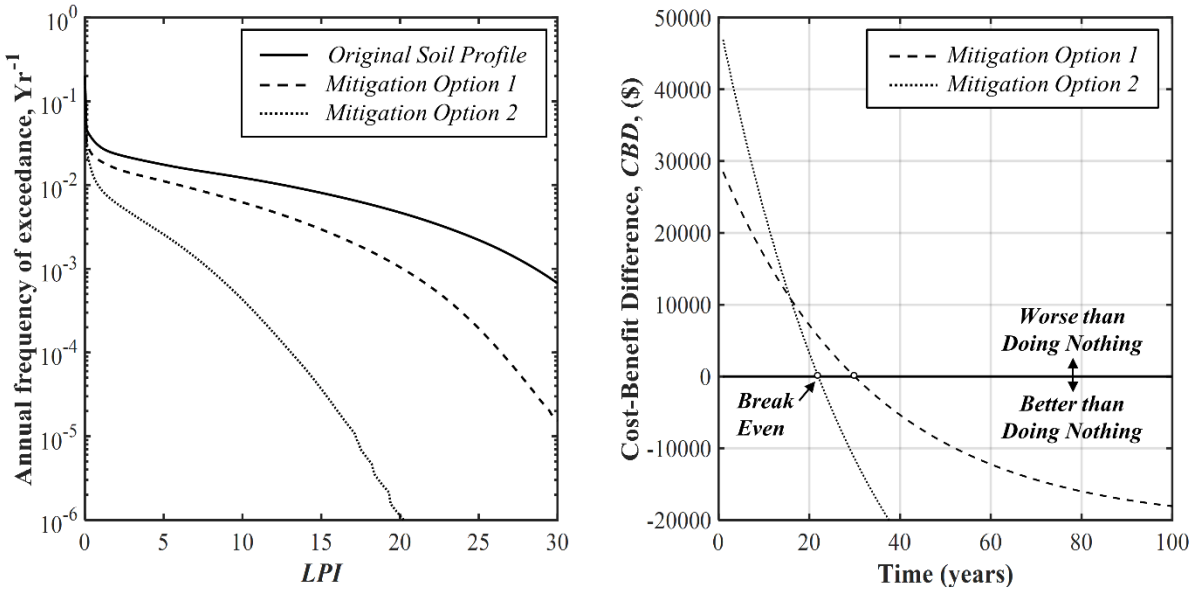


Figure 3. (a) *LPI* hazard curves for the SODO study site, with and without two different ground-improvement methods for liquefaction mitigation; **(b)** effect of planning-horizon on cost-benefit-difference analysis of ground improvement.

Thus, as illustrated in Figure 3a, mitigation strategies are detected and incorporated into the analysis through changes in λ_{LPI}, such that the annual exceedance rate of any *LPI* value is reduced due to changes to the soil profile (in this case, increased penetration resistance). By corollary, the *EAL* due to liquefaction will also diminish, per Eq. (6). In this case, the modified *EAL*s due to liquefaction are reduced from \$6,528 to \$5,010 for Option 1, and to \$3,422 for Option 2. Using these results, and assuming an interest rate of 3%, *CBD* values are computed and plotted in Figure 3b for Mitigation Options 1 and 2. It can be seen that the more costly of the two methods – Option 2 – reaches its break even point sooner than Option 1. In this case, Option 2 pays for itself through reductions in *EAL* after approximately 21 years, whereas Option 1 breaks even after approximately 30 years. Moreover, the return periods of moderate and severe damage are increased to 106 and 430 years, respectively, for Option 1, and to 187 and 1578 years, respectively, for Option 2. This

modelling approach can be used to decide whether, and how, liquefaction should be mitigated, or to demonstrate the economic feasibility of a new ground improvement technology. For example, given expected (or actual) ground-improvement costs and outcomes, the economic feasibility of any bio-inspired or bio-mediated technology can be readily demonstrated. In this regard, environmental costs and losses can and will be incorporated into the modelling as part of ongoing research, in addition to the direct costs and losses considered herein. Ultimately, the economic feasibility of any liquefaction mitigation strategy is dependent on the interactions of numerous location-, profile- and asset-specific factors. This study presented a framework for studying these factors in a consistent, practical, rational, and transparent manner. Its use could potentially lead to more efficient management of liquefaction risk worldwide.

Conclusion

Decisions about ground improvement to mitigate soil liquefaction should be made based on the expected cost and lifecycle benefit (i.e., reduction in loss over the life of an infrastructure asset). In the jargon of performance-based earthquake engineering, vulnerability functions predict loss, and thus, are critical to lifecycle cost analyses. However, field measurements of loss due to liquefaction have historically been poorly quantified, slowing the development of such functions. Accordingly, aided by new engineering and economic data from the 2010-2011 Canterbury, New Zealand, earthquakes, this study proposed practical functions for predicting liquefaction-induced loss. These functions are: (i) applicable to lightweight structures on shallow foundation systems; (ii) developed as extensions to CPT-based liquefaction analytics already in popular use; and (iii) allow for lifecycle benefits to be quantitatively weighed for a range of intervention strategies; but (iv) are not intended to predict losses incurred from lateral spreading. The proposed functions were demonstrated via application to a site in the SODO district of Seattle, USA, wherein a lifecycle

cost-benefit analysis of ground improvement was performed. Ongoing research will: (i) develop analogous functions for an array of popular liquefaction models (only one model was utilized herein); (ii) develop functions for specific types of shallow-foundation systems (all shallow foundations were grouped herein); and (iii) expand upon the economic modelling briefly demonstrated in this paper (e.g., to reflect risk-aversion behaviors or to include indirect environmental costs and losses).

References

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, Civil & Environmental Eng., UC Davis, CA.

Bradley, B.A. (2013). "Site-specific and spatially-distributed ground motion intensity estimation in the 2010-2011 Christchurch earthquakes." *Soil Dynamics & Earthquake Eng*, 48, 35-47.

Bradtmueller, J.P., McFarland, A.E., and Foley, S.P. (2015). "A comparison of house size and foundation type for U.S. residential homes." In: *Proceedings of the 51st ASC Annual International Conference*, The Associated Schools of Construction.

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." *Proc., 2nd Int. Conf. on Microzonation*, San Francisco, USA.

Kappos, A.J., Dimitrakopoulos, E.G. (2008). "Feasibility of pre-earthquake strengthening of buildings based on cost-benefit and life-cycle cost analysis, with the aid of fragility curves." Nat Hazards, 45, 33–54.

Kramer, S.L., Mayfield, R.T. (2007). "Return Period of Soil Liquefaction." *Journal of Geotechnical and Geoenvironmental Engineering*, 133(7), 802-813.

Liel, A.B., Dierlein, G.G. (2013). "Cost-Benefit Evaluation of Seismic Risk Mitigation Alternatives for Older Concrete Frame Buildings." *Earthquake Spectra,* 29(4), 1391-1411.

Maurer, B.W., Green, R.A., Cubrinovski, M., and Bradley, B.A. (2014). "Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand." *Journal of Geotechnical and Geoenvironmental Engineering* 140(7), 04014032.

Maurer, B.W., Green, R.A., Cubrinovski, M., Bradley, B.A. (2015). "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, 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.

NZGD (2019). New Zealand Geotechnical Database; New Zealand Earthquake Commision, Wellington New Zealand. Accessed via < https://www.nzgd.org.nz>

Porter, K. (2018). *A Beginner's Guide to Fragility, Vulnerability, and Risk*. University of Colorado, 92 pp.

Ramirez, C.M., Liel, A.B., Mitrani-Reiser, J., Haselton, C.B., Spear, A.D., Steiner, J., Dierlein, G.G, Miranda, E. (2012). "Expected earthquake damage and repair costs in reinforced concrete frame buildings." *Earthquake Eng and Structural Dynamics.* 41:1455-1475.

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

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.

van Ballegooy, S., Green, R.A., Lees, J., Wentz, F., and Maurer, B.W., 2015. Assessment of various CPT based liquefaction severity index frameworks relative to the Ishihara (1985) H_1-H_2 boundary curves, *Soil Dynamics and Earthquake Engineering* 79, 347-364.

APPENDIX J

Publication 10 (Software), highlighted below, is provided on the following pages in full. Specifically, a "Quick ReadMe" installation file and a "User Manual" are provided.

No.	Type	Summary in Ch.	Publication Name	Current Status
1	Journal Paper 1	2	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,* In Press.	Published
2	Dataset 1	2	Geyin M and Maurer BW (2021). "CPT-Based Liquefaction Case Histories from Global Earthquakes: A Digital Dataset (Version 1)." *DesignSafe CI.*	Published
3	Journal Paper 2	3	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.	Published
4	Journal Paper 3	3	Geyin M and Maurer BW (2020). "Fragility Functions for Liquefaction Induced Ground Failure." *Journal of Geotechnical and Geoenvironmental Engineering* 146(12): 04020142.	Published
5	Conference Paper 1	3	Geyin M and Maurer BW (2019). "An Analysis of Liquefaction-Induced Free-Field Ground Settlement Using 1,000 + Case-Histories: Observations vs. State-of-Practice Predictions." *ASCE Geotechnical Special Publication No.* 308: 489–498.	Published
6	Journal Paper 4	3	Geyin M and Maurer BW (2021). "Evaluation of a Cone Penetration Test Thin-Layer Correction Procedure in the Context of Global Liquefaction Performance." *Engineering Geology,* In Review.	Published
7	Journal Paper 5	3	Geyin M, Maurer BW, and Christofferson K (2021). "An AI-Driven, Mechanistically Grounded Geospatial Liquefaction Model for Rapid Response and Planning." *Soil Dynamics and Earthquake Engineering.*	In Review
8	Software 1	3	Geyin M and Maurer BW (2021). "RapidLiq: Software for Near-Real-Time Prediction of Soil Liquefaction." *DesignSafe CI.*	Published
9	Conference Paper 2	4	Geyin M, Maurer, BW, and van Ballegooy S (2020). "Lifecycle Liquefaction Hazard Assessment and Mitigation." *ASCE Geotechnical Special Publication No. 320*: 312-320.	Published
10	Software 2	5	Geyin M and Maurer BW (2021). "Horizon: CPT-based liquefaction risk assessment and decision software." *DesignSafe CI.*	Published

HORIZON: CPT-BASED LIQUEFACTION RISK ASSESSMENT AND DECISION SOFTWARE

Mertcan Geyin, Brett Maurer

***Horizon* User Manual**

Ver 1.0

Introduction

Horizon is a standalone software that is capable of calculating return periods of liquefaction manifestations, liquefaction induced foundation damages and building damage ratios for various shallow foundations. When associated with NSHMP-haz by USGS, *Horizon* is also capable of computing expected annual losses and perform risk integral on the computed losses. User can also assess the feasibility of ground improvements by means of cost-benefit analyses and see the return of investments of their mitigation strategies.

Input Information

Horizon needs CPT data to estimate liquefaction damage potentials. This includes depth, tip resistance (qc), sleeve resistance (fs) and pore pressure measurements (upp - Optional). Version 1: Input file should be .xlsx format.

Horizon works with SI units. MS Excel input file should be 3 columns (4 if pore pressure measurements exist), preferably with same number of rows with no headers. Columns should be ordered as: (1) depth (meters), (2) tip resistance (MPa), (3) sleeve resistance (MPa), (4) pore pressures (MPa). Excel file should have a single tab. If input data has missing rows, non-numeric characters etc. (except in depth column), Horizon will try to linearly interpolate or extrapolate each column's data to make them rectangular. If there is an error during interpolation or extrapolation process, it will let the user know. Nonetheless, user is encouraged to input a rectangular and numeric dataset (i.e. same number of rows).

File handling

File handling in Horizon is performed by Windows® Requester working within MATLAB Runtime Environment (MRE). Data is then stored and parsed in structural cell format in MRE. During NSHMP-haz based calculations, additional folders and files may be created and deleted consecutively.

Help

User is encouraged to hover over the fillable boxes, buttons, switches and drop-down menus with the cursor to read detailed information about the modifications and/or analyses. For additional info and questions, or found any bugs, please email: mgeyin@uw.edu

Flow of *Horizon*

In this section, general flow of the software will be explained in detail for each tab. User can hover over many buttons, tick boxes or input boxes to see detailed descriptions of selections/inputs. *Horizon* requires a CPT data to start the calculations. After importing the CPT, user is able to fully utilize "Data input and Processing" tab. Rest of the tabs are in the order that of the calculations should be performed.

Data input and Processing Tab

Input a compatible .xlsx file by pressing "Press to input CPT data" button in the first tab. Select file and import. *Horizon* will automatically draw the CPT profile on the "Input Plots," and put the data on the "Input Data" tabs.

At this stage, user can explore the characteristics of the CPT profile, make corrections on data that includes corrections regarding (1) pre-drill depth; (2) pore pressure; (3) noise; (4) tip-sleeve alignment; and finally (5) thin layer and interfaces.

Figure 2. Data Input and Processing Tab

After selecting necessary modifications/corrections on the CPT, press "Press to apply corrections or continue" button. This will update "Input Plots," and "Input Data" tabs.

To be able to run a liquefaction triggering model, user is required to specify the event parameters: (1) Moment magnitude, *Mw*; (2) Peak Ground Acceleration, *PGA* and (3) ground

water depth at the time of the earthquake. After selecting required liquefaction triggering model and its settings, use “Press to run liquefaction triggering and manifestation models” button.

After running the models, Horizon will fill the table at the bottom of the current tab and update; (1) Soil Behavior Index (*SBI*) Profile; (2) Cyclic Stress Ratio (*CSR*) Profile; and (3) Factor of Safety Profile (*FoS*) tabs on the right section (Figure 1). It is possible to visualize how *CSR* and *FoS* profiles behave under different combinations of *PGA* and *Mw* by simply sliding the cursors on the intensity measures.

Manifestation Fragilities Tab

To see the probability of reaching or exceeding a given surface manifestation severity, Manifestation Fragilities Tab utilizes functions developed in Geyin et al. (2020a). This tab is capable of displaying the fragilities that are derived by using Liquefaction Potential Indices (*LPI*, LPI_{ish}) or Liquefaction Severity Number (*LSN*) as metrices to predict the cumulative response of the input soil profile (Figure 2).

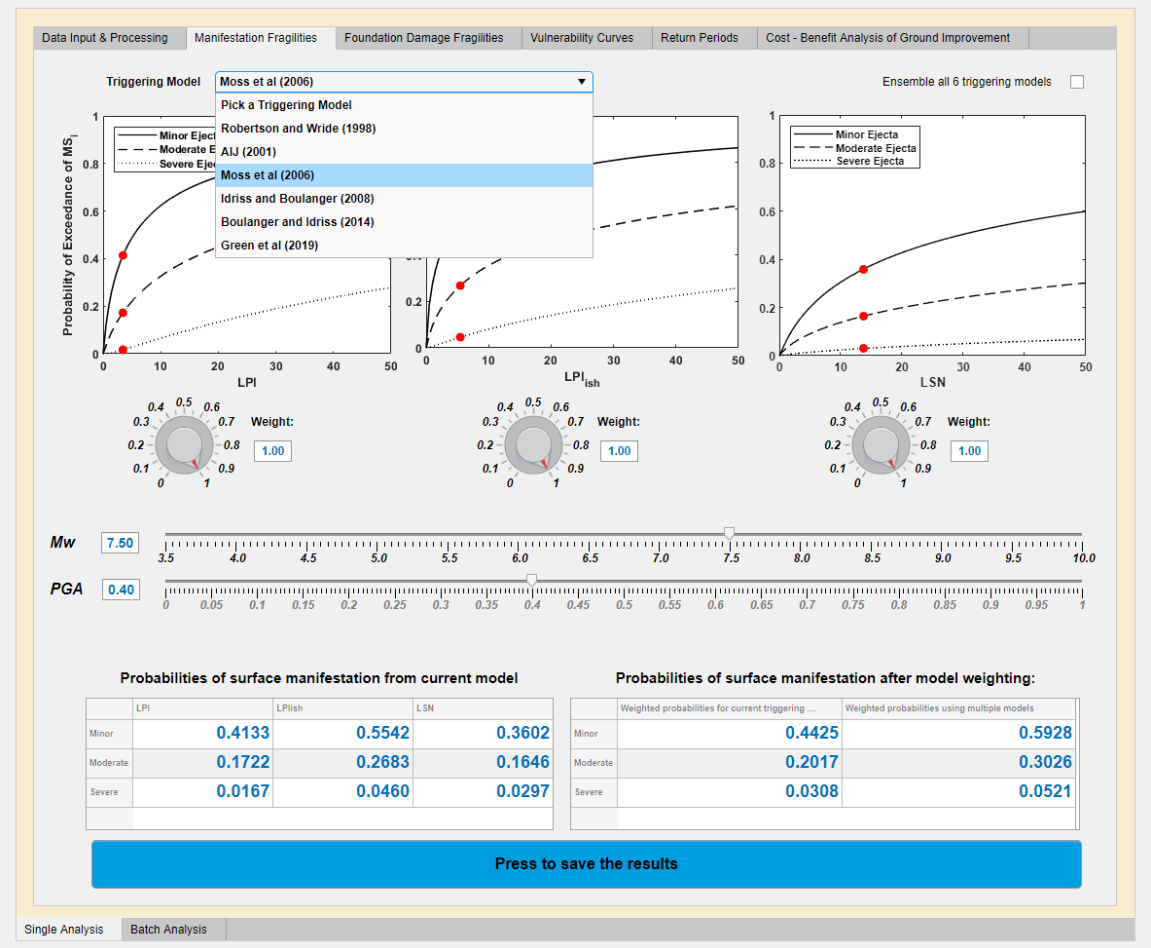


Figure 3. Manifestation Fragilities Tab

While it is possible to gather predictions from each triggering and manifestation model combination separately, user is able to ensemble models by assigning weight factors using the knobs for a particular triggering model. If user wants to ensemble all triggering models by weighting to predict the probability of a given severity, use “Ensemble all 6 triggering model” tick box.

Foundation Damage Fragilities Tab

Figure 4. Foundation Fragilities Tab

Like Manifestation Fragilities Tab, user is able to see the probabilities of reaching or exceeding a given severity of a foundation damage. It is possible to pick from 4 foundation types and akin to manifestation fragilities, user can weight the predictions from different manifestation models using the knobs. Please note that this tab uses Boulanger and Idriss (2014) triggering model to compute the liquefaction demand parameters, *LDP*s, (i.e., *LPI*, LPI_{ish}, *LSN*).

Vulnerability Curves Tab

Vulnerabilities (i.e., monetary loss due to the earthquake event) can be estimated in two ways: (1) Expected Building Damage Ratios and, (2) Building Damage Ratio Fragilities. Users are encouraged to see Geyin et al. (2020b) for detailed descriptions of data used to drive the predictions of liquefaction-induced monetary loss and the methodology utilized to create the vulnerability functions.

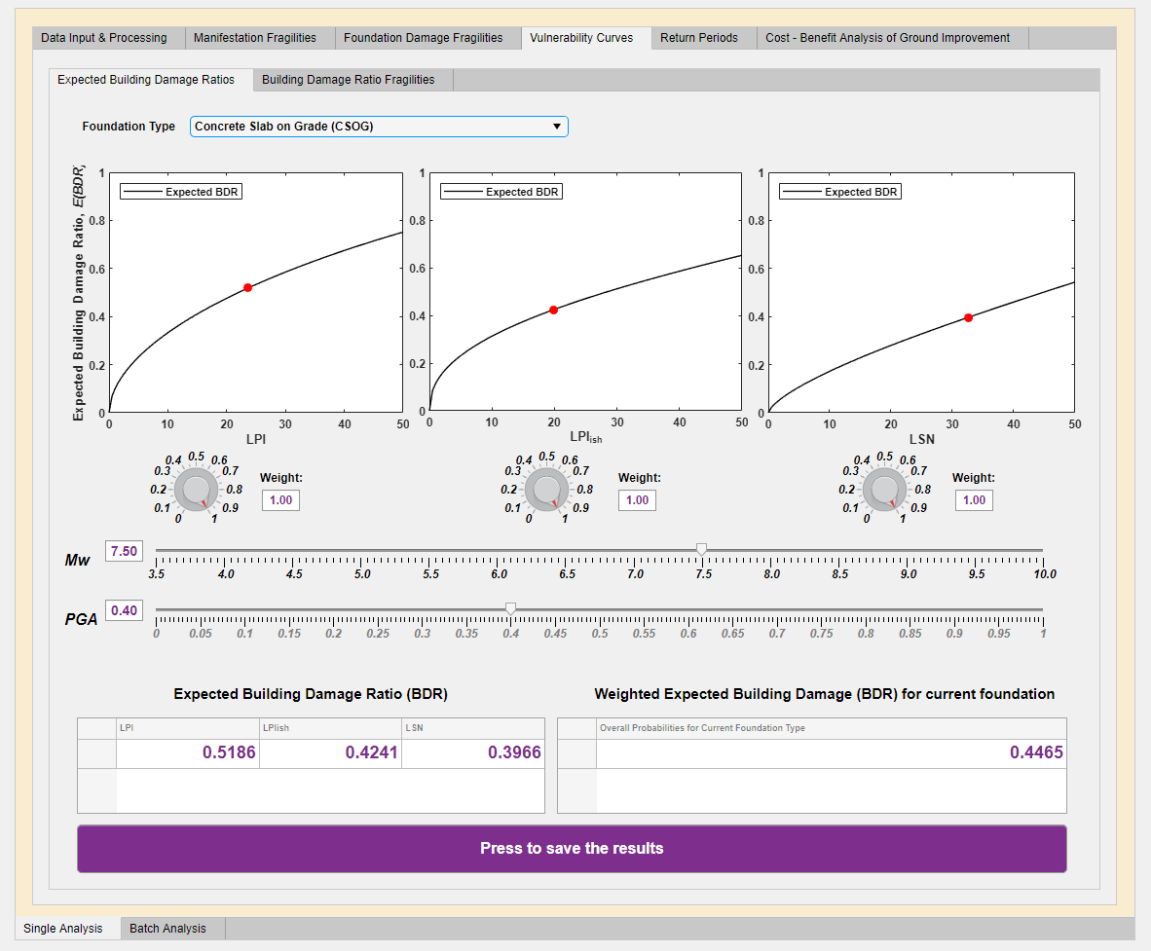


Figure 5. Vulnerability Curves Tab

Expected Building Damage Ratios tab will create vulnerability curves for the mean value of the loss for a given liquefaction demand parameter and easier to interpret (Figure 4). For more robust

analyses of vulnerability, use Building Damage Ratio Fragilities tab that will develop the fragilities of certain building damage ratios for a given *LDP*. This tab will display the probabilities of reaching or exceeding a given building damage ratio as a function of the *LDP* for the selected foundation. As in previous sections, it is possible ensemble model predictions with weight factors using the knobs.

Return Periods Tab

Return Periods tab will compute the return periods of different severities of: (1) surface manifestations (2) foundation damages (3) building damage ratios. To accomplish this, *Horizon* works with NSHMP-haz to automatically gather hazard data. User is required to enter the location of the input CPT and specify a working folder for *Horizon* to solve the hazard integrals.

Click "Press to specify a working folder for hazard calculations" button. Refer to *Horizon* Quick Setup Guide and direct *Horizon* to the Source Database (Figure 5). Select "NSHMP Database" folder. Depending on the CPT location entered, Horizon will automatically start the hazard deaggregation process while communicating with NSHMP-haz via Windows Shell. Note that all data is stored and parsed in structural cell format in MRE. During NSHMP-haz based calculations, additional folders and files may be created and deleted consecutively within the "NSHMP Database" folder. Do not alter this folder while *Horizon* is open.

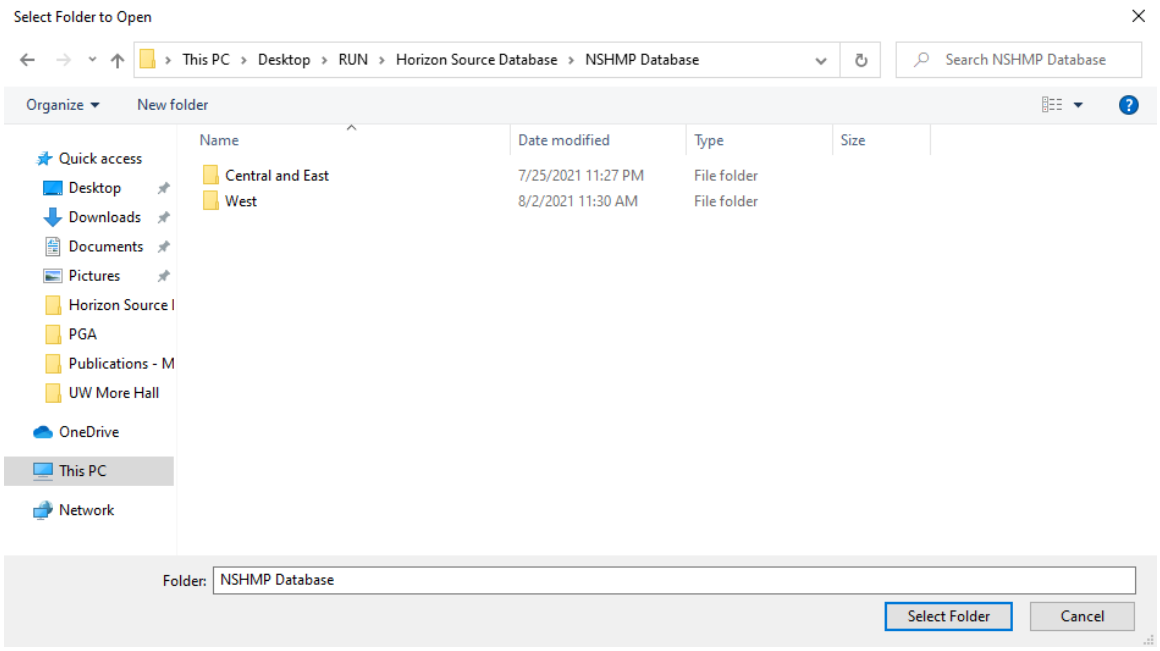


Figure 6. Specifying a working folder for Horizon

Press “Seismic Hazard Settings” button to change the soil class and the resolution of which the deaggregation on magnitude will be performed. Note that the lower the resolution the faster the calculations will be performed.

Start calculations for magnitude-based hazard deaggregation using “Press to start deaggregation” button. Calculations may take several minutes. See detailed progress of the deaggregation on the top right section of the window. After the calculations are finished, magnitude-deaggregated PGA-based hazard figure will be updated.

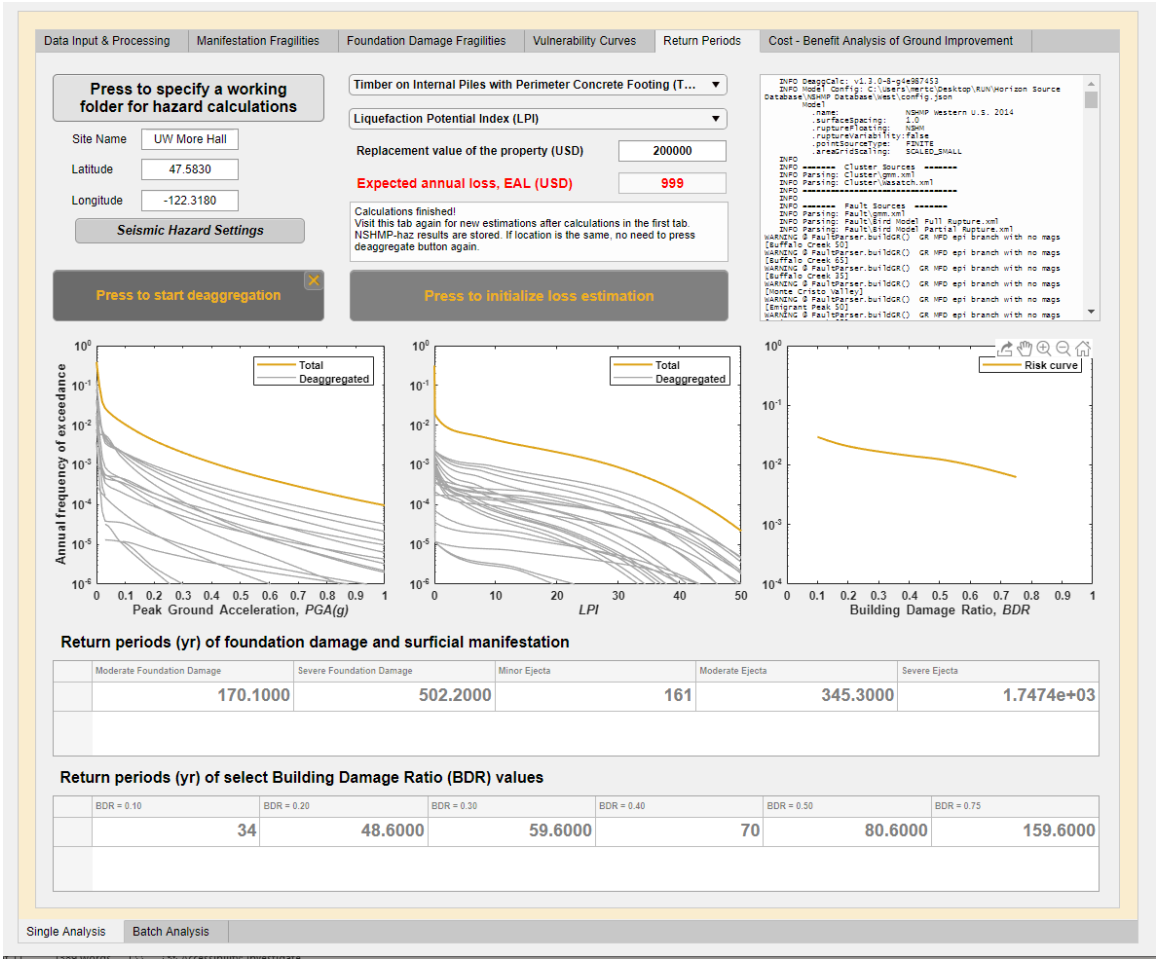


Figure 7. Return Periods Tab

Pick the foundation type and the LDP from the drop-down menus. After entering the replacement value of the property, use "Press to initialize loss estimation" button. (Figure 6). This will quickly perform hazard integrations based on selected LDP using associated fragility, vulnerability and PGA-based hazard data for the specified location for the input CPT.

Cost – Benefit Analysis of Ground Improvement

To perform cost-benefit analyses for the long-term performance of ground improvements, user can: (1) improve the current profile using a multiplier on corrected tip resistance or (2) manually

input post-improvement CPT data. If option (1) is selected, it is possible to utilize depth weighted multipliers by pressing the pushbuttons after specifying the multiplier at the surface. It is also possible to change the multipliers at each depth using the table.

Press “Apply corrections” button and check normalized fines-corrected tip resistance profiles before and after improvement on the top right section. Multiplier plot shows the ratio of improved to unimproved tip resistance profiles.

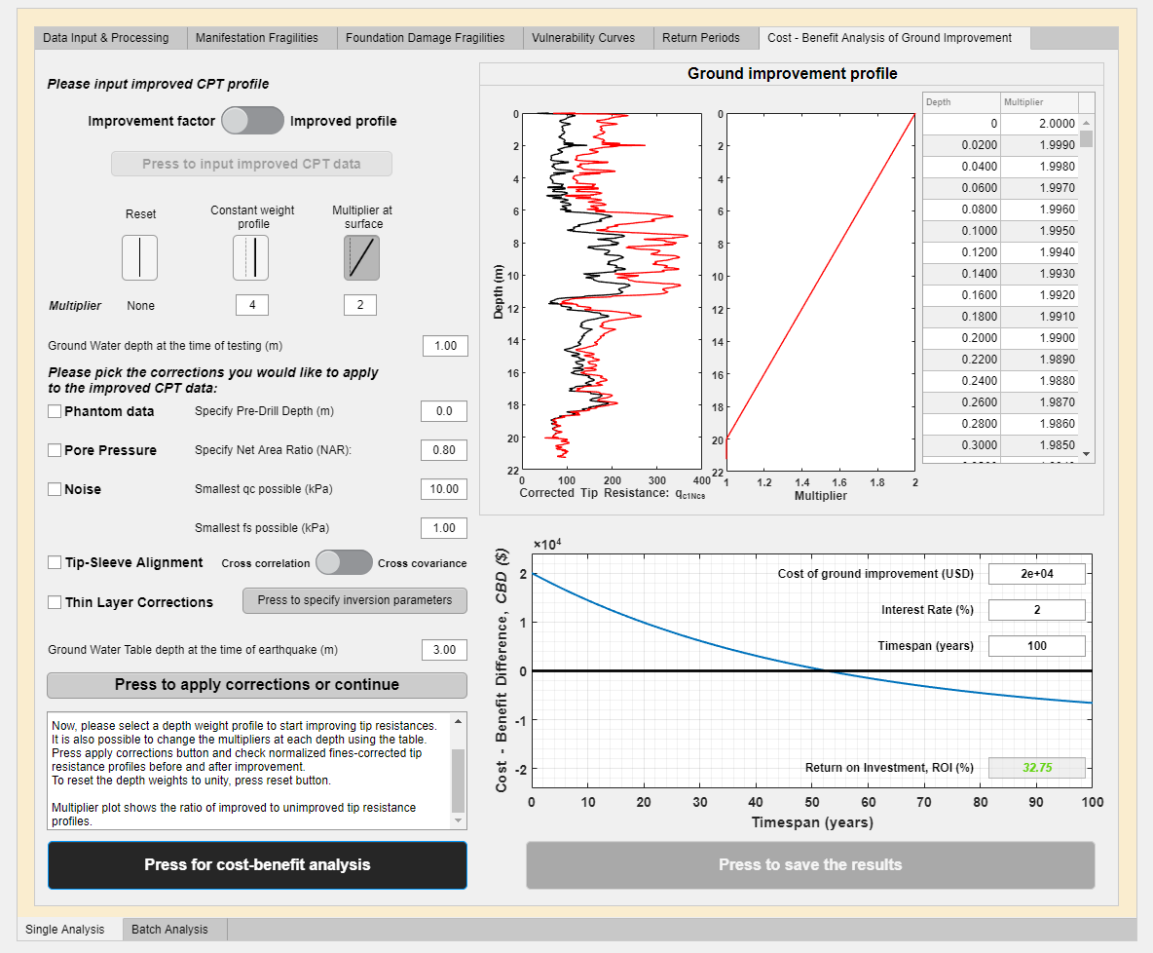


Figure 8. Cost-Benefit Analyses of Ground Improvement Tab

Finally, use “Press for cost-benefit analysis” button and see the long-term economic viability from the figure on the bottom right. When the Cost Benefit Difference, CBD is less than zero,

intervention strategy becomes an auspicious decision considering the service life of the project. Return on Investment, ROI can also be obtained by entering a planning horizon in years onto the same figure. Figure 7 shows a CPT profile before and after improvement. Input CPT in the "Data input and Processing" tab was improved by applying a linearly decreasing multiplier to the corrected tip resistance of the CPT profile, also reducing the water table depth at the earthquake.

References

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.

Geyin M and Maurer BW (2020a). Fragility functions for liquefaction induced ground failure. *Journal of Geotechnical and Geoenvironmental Engineering*, doi: 10.1061/(ASCE)GT.1943-5606.0002416.

Geyin, M., Maurer, B.W., and van Ballegooy, S. (2020b). Lifecycle Liquefaction Hazard Assessment and Mitigation. *Geocongress 2020: BioGeotechnics (E. Kavazanjian, J.P. Hambleton, R. Makhnenko, and A.S. Budge, eds.)*, Geotechnical Special Publication 320: 312-320. American Society of Civil Engineers.

HORIZON: QUICK README (BUILDING AND RUNNING HORIZON) VER 1.0

1. If location-specific probabilistic seismic-hazard calculations are not desired by the user (e.g., return periods of free-field liquefaction manifestation, foundation damage, or economic loss; cost-benefit mitigation analyses), then the standalone executable file is sufficient. This approach allows users to process and analyze CPTs (e.g., to predict liquefaction occurrence and consequence) without following the full installation procedure. Please install Horizon via "MyAppInstaller_mcr.exe" in the "Redist" folder.

2. If location-specific probabilistic seismic-hazard calculations are desired (i.e., all calculations performed within the "Return Periods" and "Cost-Benefit Analyses" tabs of *Horizon*), then *Java* and *Git* must also be installed to utilize nshmp-haz, a *Java*-based platform for conducting probabilistic seismic hazard related analyses (PSHA). nshmp-haz is developed and maintained by the National Seismic Hazard Mapping Project (NSHMP) within the U.S. Geological Survey's (USGS) earthquake hazards program (EHP).

To utilize the full capabilities of *Horizon*, please download and install:

Java (64-bit) from:

https://www.oracle.com/java/technologies/javase/javase-jdk8-downloads.html

and *Git* from:

https://git-scm.com/downloads

3. *Java* and *Git* are used to perform NSHMP-Haz calculations. *Horizon* passes the user-input coordinates and site conditions (i.e., *Vs30*) to NSHMP-Haz via Windows shell. This requires prior installation of NSHMP-Haz, which can be achieved by following the procedure below (*Java* and *Git* must first be installed).

i) Create a folder in any known location on your computer (e.g., within the "Documents" folder), having any arbitrary name (e.g., "NSHMP").

ii) Enter the following lines in the Windows command prompt (enter "CMD" in the Windows search bar to locate the command prompt):

cd /path/to/project/directory

[This changes the directory to the newly created folder. Note space after "cd" and FORWARD slashes]

[Example of correct entry:]

```
C:\Users\Brett>cd /Users/Brett/Documents/NSHMP
```

---

git clone https://github.com/usgs/nshmp-haz.git

[This copies program files into the folder created previously]

[Example of correct entry:]

```
C:\Users\Brett\Documents\NSHMP>git clone https://github.com/usgs/nshmp-haz.git
Cloning into 'nshmp-haz'...
remote: Enumerating objects: 29061, done.
Receiving objects:1 (delta 0), reused 0 (delta 0), pack-reused 29061Receiving objects:
 100% (29061/29061), 35.77 MiB | 6.95 MiB/s, done.
Resolving deltas: 100% (19426/19426), done.
Updating files: 100% (773/773), done.
```

[If successful, a new subfolder "nshmp-haz" should appear within the folder you created]

---

cd nshmp-haz

[This changes the directory to the newly created sub-folder titled "nshmp-haz"]

[Example of correct entry:]

```
C:\Users\Brett\Documents\NSHMP>cd nshmp-haz
```

---

gradlew assemble

[This assembles the program]

[Example of correct entry:]

```
C:\Users\Brett\Documents\NSHMP\nshmp-haz>gradlew assemble
Downloading https://services.gradle.org/distributions/gradle-4.10.2-bin.zip
..........................................................................

Welcome to Gradle 4.10.2!

Here are the highlights of this release:
 - Incremental Java compilation by default
 - Periodic Gradle caches cleanup
 - Gradle Kotlin DSL 1.0-RC6
 - Nested included builds
 - SNAPSHOT plugin versions in the `plugins {}` block

For more details see https://docs.gradle.org/4.10.2/release-notes.html

Starting a Gradle Daemon (subsequent builds will be faster)

> Task :compileJava
Note: Some input files use or override a deprecated API.
Note: Recompile with -Xlint:deprecation for details.

BUILD SUCCESSFUL in 40s
3 actionable tasks: 3 executed
C:\Users\Brett\Documents\NSHMP\nshmp-haz>
```

4. Create another new folder in any known location (e.g., within the "Documents" folder) titled "Horizon Source Database" (but any name may be chosen). This will be used as a working folder by *Horizon* and will contain all input files needed by NSHMP-Haz. *Horizon* will automatically create and erase files within this folder, so its location and contents should not be changed. Copy and paste the "NSHMP Database" folder (obtained when downloading *Horizon*) into the newly created folder. Remember the folder location – it will be needed when performing PSHA-related calculations in *Horizon*.

With completion of steps **1 - 4**, *Horizon* will operate with full capability.

VITA

Mertcan Geyin graduated from METU in Ankara, Turkey in 2014 with a degree of Bachelor of Science in Civil Engineering. Following, they entered Virginia Polytechnic Institute and State University and received their master's in Civil Engineering. They entered University of Washington for their PhD studies and received their PhD in Civil and Environmental Engineering in 2021.