



Evaluating simplified methods for liquefaction assessment for loss estimation

Indranil Kongar¹, Tiziana Rossetto¹, and Sonia Giovinazzi²

¹Earthquake and People Interaction Centre (EPICentre), Department of Civil, Environmental and Geomatic Engineering, University College London, London, WC1E 6BT, UK

²Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, 8140, New Zealand

Correspondence to: Indranil Kongar (ucfbiko@ucl.ac.uk)

Received: 23 August 2016 – Discussion started: 15 September 2016

Revised: 1 March 2017 – Accepted: 27 March 2017 – Published: 1 June 2017

Abstract. Currently, some catastrophe models used by the insurance industry account for liquefaction by applying a simple factor to shaking-induced losses. The factor is based only on local liquefaction susceptibility and this highlights the need for a more sophisticated approach to incorporating the effects of liquefaction in loss models. This study compares 11 unique models, each based on one of three principal simplified liquefaction assessment methods: liquefaction potential index (LPI) calculated from shear-wave velocity, the HAZUS software method and a method created specifically to make use of USGS remote sensing data. Data from the September 2010 Darfield and February 2011 Christchurch earthquakes in New Zealand are used to compare observed liquefaction occurrences to forecasts from these models using binary classification performance measures. The analysis shows that the best-performing model is the LPI calculated using known shear-wave velocity profiles, which correctly forecasts 78 % of sites where liquefaction occurred and 80 % of sites where liquefaction did not occur, when the threshold is set at 7. However, these data may not always be available to insurers. The next best model is also based on LPI but uses shear-wave velocity profiles simulated from the combination of USGS V_{S30} data and empirical functions that relate V_{S30} to average shear-wave velocities at shallower depths. This model correctly forecasts 58 % of sites where liquefaction occurred and 84 % of sites where liquefaction did not occur, when the threshold is set at 4. These scores increase to 78 and 86 %, respectively, when forecasts are based on liquefaction probabilities that are empirically related to the same values of LPI. This model is potentially more useful for insurance since the input data are publicly available. HAZUS

models, which are commonly used in studies where no local model is available, perform poorly and incorrectly forecast 87 % of sites where liquefaction occurred, even at optimal thresholds. This paper also considers two models (HAZUS and EPOLLS) for estimation of the scale of liquefaction in terms of permanent ground deformation but finds that both models perform poorly, with correlations between observations and forecasts lower than 0.4 in all cases. Therefore these models potentially provide negligible additional value to loss estimation analysis outside of the regions for which they have been developed.

1 Introduction

The recent earthquakes in Haiti (2010), Canterbury, New Zealand (2010–2011), and Tohoku, Japan (2011), highlighted the significance of liquefaction as a secondary hazard of seismic events and the significant damage that it can cause to buildings and infrastructure. However, the insurance sector was caught out by these events, with catastrophe models underestimating the extent and severity of liquefaction that occurred (Drayton and Verdon, 2013). A contributing factor to this is that the method used by some catastrophe models to account for liquefaction is based only on liquefaction susceptibility, a qualitative parameter that considers only surficial geology characteristics. Furthermore, losses arising from liquefaction are estimated by adding an amplifier to losses estimated due to building damage caused by ground shaking (Drayton and Vernon, 2013). There is a paucity of past event data on which to calibrate an amplifier and, consequently,

significant losses from liquefaction damage will only be estimated if significant losses are already estimated from ground shaking, whereas it is known that liquefaction can be triggered at relatively low ground shaking intensities (Quigley et al., 2013).

Therefore there is scope within the insurance and risk management sectors to adopt more sophisticated approaches for forecasting liquefaction for both future risk assessments and post-event rapid response analyses. It is also important to develop a better understanding of the correlation between liquefaction effects and physical damage of the built environment, similar to the fragility functions that are used to estimate damage associated with ground shaking. This is particularly the case for critical infrastructure systems since, whilst liquefaction is less likely than ground shaking to be responsible for major building failures (Bird and Bommer, 2004), it can have a major impact on lifelines such as roads, pipelines and buried cables. Loss of power and reduction in transport connectivity are major factors affecting the resilience of business organizations in response to earthquakes as they can delay the recommencement of normal operations. Evaluating the seismic performance of infrastructure is therefore critical to understanding indirect economic losses caused by business interruption and to achieve this it is necessary to assess the liquefaction risk in addition to that posed by ground shaking. Therefore in this paper we investigate the performance of a range of models that can be applied to forecast the occurrence and scale of liquefaction based on simple and accessible input datasets. The performances are evaluated by comparing model forecasts to observations from the 2010–2011 Canterbury earthquake sequence.

Bird and Bommer (2004) surmised that there are three options that loss estimators can select to deal with ground failure hazards. They can either ignore them, use a simplified approach or conduct a detailed geotechnical assessment. The first of these options will likely lead to underestimation of losses in earthquakes where liquefaction is a major hazard and lead to recurrence of the problems faced by insurers following the 2010–2011 Canterbury earthquakes in particular. The last option, detailed assessment, is appropriate for single-site risk analysis but is impractical for insurance loss estimation purposes because (1) insurers are unlikely to have access to much of the detailed geotechnical data required as inputs to these methods; (2) they may not have the in-house expertise to correctly apply such methods and engaging consultants may not be a viable option; and (3) loss estimation studies are often conducted on a regional, national or supra-national scale for which detailed assessment would be too expensive and time consuming.

There are three stages to forecasting the occurrence of liquefaction and its scale (Bird et al., 2006). First it is necessary to determine whether soils are susceptible to liquefaction. Liquefaction susceptibility is based solely on ground conditions with no earthquake-specific information. This is often done qualitatively and currently this is also the full ex-

tent to which liquefaction risk is considered in some catastrophe models (Drayton and Verdon, 2013). The next step is to determine liquefaction triggering, which determines the likelihood of liquefaction for a given earthquake based on the susceptibility and other earthquake-specific parameters. Finally the scale of liquefaction can be estimated as a permanent ground deformation (PGDf). Since current catastrophe modelling practice is to consider only the first stage, liquefaction susceptibility, this paper focuses primarily on the extension of this practice to include liquefaction triggering.

The models assessed in this paper have been selected because their input requirements are limited to data that are in the public domain or could be easily obtained without significant time or cost implications, arising for example from detailed site investigation. Furthermore, the models are appropriate for regional-scale analysis and although some engineering judgment is required in their application, they do not require specialist geotechnical expertise. In Sect. 2, each of the models assessed in this paper are described and Sect. 3 presents a summary of the liquefaction observations from the Canterbury earthquake sequence and the method used to compare the model forecasts against observations. The results and statistical analysis of the model assessment are presented in Sect. 4, in relation to deterministic forecasts, and in Sect. 5, in relation to probabilistic forecasts. Finally, Sect. 6 briefly considers the performance of simplified models for quantifying PGDf.

2 Liquefaction assessment models

Nine liquefaction forecasting models are compared in this paper, including three alternative implementations of the liquefaction potential index (LPI) method proposed by Iwasaki et al. (1984), three versions of the liquefaction models included in the HAZUS[®]MR4 software (NIBS, 2003) and three distinct models proposed by Zhu et al. (2015). This section summarizes how each of the models are applied to make site-specific liquefaction forecasts. This paper presents a large number of acronyms and variables. For clear reference, Table 1 lists the acronyms used in this paper and Table 2 lists the variables used.

2.1 Liquefaction potential index

The most common approach used to forecast liquefaction triggering is the factor of safety against liquefaction (FS), which is defined as the cyclic resistance to cyclic stress ratio for a layer of soil at depth, z (Seed and Idriss, 1971). The cyclic stress ratio (CSR) can be expressed by Eq. (1), where a_{\max} is the peak horizontal ground acceleration; g is the acceleration of gravity; σ_v is the total overburden stress at depth z ; σ'_v is the effective overburden stress at depth z ; and r_d is a shear stress reduction coefficient given by Eq. (2).

Table 1. Reference list of acronyms used in this paper.

Acronym	Description
AUC	Area under ROC curve
CGD	Canterbury Geotechnical Database
CPT	Cone penetration test
CRR	Cyclic resistance ratio
CSR	Cyclic stress ratio
CTI	Compound topographic index
EQC	Earthquake Commission
FN	False negative model forecasts (no.)
FP	False positive model forecasts (no.)
FPR	False positive rate (FP/observed negatives)
FS	Factor of safety against liquefaction
LPI	Liquefaction potential index
MCC	Matthew’s correlation coefficient
MSF	Magnitude scaling factor
MWF	Magnitude weighting factor
NEHRP	National Earthquake Hazards Reduction Program
PGA	Peak ground acceleration
PGDf	Permanent ground deformation
PGDf _H	Horizontal permanent ground deformation
PGDf _V	Vertical permanent ground deformation
RMSE	Root-mean-square error
ROC	Receiver operating characteristic
SPT	Standard penetration test
TN	True negative model forecasts (no.)
TNR	True negative rate (TN/observed negatives)
TP	True positive model forecasts (no.)
TPR	True positive rate (TP/observed positives)
USGS	United States Geological Survey

$$CSR = 0.65 \left(\frac{a_{max}}{g} \right) \left(\frac{\sigma_v}{\sigma'_v} \right) r_d \tag{1}$$

$$r_d = 1 - 0.00765z, \text{ for } z < 9.2 \text{ m} \tag{2}$$

$$r_d = 1.174 - 0.0267z, \text{ for } z \geq 9.2 \text{ m}$$

The cyclic resistance ratio (CRR) is normally calculated from geotechnical parameters based on cone penetration test (CPT) or standard penetration test (SPT) results. However, Andrus and Stokoe (2000) propose an alternative method for calculating CRR based on shear-wave velocity, V_S , as shown in Eq. (3), where V_{S1} is the stress-corrected shear-wave velocity; V_{S1}^* is the limiting upper value of V_{S1} for cyclic liquefaction occurrence, which varies between 200 and 215 $m s^{-1}$ depending on the fines content of the soil; and MSF is a magnitude scaling factor. V_{S1} is given by Eq. (4), where P_a is a reference stress of 100 kPa. The magnitude scaling factor is given by Eq. (5), where M_w is the moment magnitude of the earthquake.

$$CRR = \left[0.022 \left(\frac{V_{S1}}{100} \right)^2 + 2.8 \left(\frac{1}{V_{S1}^* - V_{S1}} - \frac{1}{V_{S1}^*} \right) \right] \times MSF \tag{3}$$

$$V_{S1} = V_S \left(\frac{P_a}{\sigma'_v} \right)^{0.25} \tag{4}$$

$$MSF = \left(\frac{M_w}{7.5} \right)^{-2.56} \tag{5}$$

Liquefaction is forecast to occur when $FS \leq 1$ and forecast not to occur when $FS > 1$. However, Juang et al. (2005) found that Eq. (3) is conservative for calculating CRR, resulting in lower factors of safety and overestimation of the extent of liquefaction occurrence. To correct for this, they propose a multiplication factor of 1.4 to obtain an unbiased estimate of the factor of safety, FS^* , given by Eq. (6).

$$FS^* = 1.4 \times \frac{CRR}{CSR} \tag{6}$$

FS^* is an indicator of potential liquefaction at a specific depth. However, Iwasaki et al. (1984) noted that damage to structures due to liquefaction was affected by the severity of liquefaction at ground level and so propose an extension to the factor of safety method, the LPI, which estimates the likelihood of liquefaction at surface level by integrating a function of the factors of safety for each soil layer within the top 20 m of soil. They calculate LPI by Eq. (7), where $F^* = 1 - FS^*$ for a single soil layer. The soil profile can be subdivided into any number of layers (e.g. 20 1 m layers or 40 0.5 m layers), depending on the resolution of data available. Using site data from a collection of nine Japanese earthquakes between 1891 and 1978, Iwasaki et al. (1984) calibrated the LPI model and determined guideline criteria for determining liquefaction risk. These criteria propose that liquefaction risk is very low for $LPI = 0$, low for $0 < LPI \leq 5$, high for $5 < LPI \leq 15$ and very high for $LPI > 15$.

$$LPI = \int_0^{20} F^* (10 - 0.5z) dz \tag{7}$$

One of the critical considerations for insurers is availability of model input data. For post-event analysis, ground accelerations may be available from various online sources, with one example being the USGS ShakeMaps (USGS, 2014a). However, if they are not, then it would be necessary to apply engineering judgment in the selection of appropriate ground motion prediction equations (either a single equation or multiple equations applied in a logic tree). The LPI model also requires water table depth and soil unit weights. If these are not known exactly, engineering judgment needs to be applied to estimate these based on information in existing literature. For the specific case study presented in this paper, some V_S data are available from published sources. However, more generally V_S data are not in the public domain and would require ground investigation to acquire. Even in cases where V_S data are available, they may not necessarily be available

Table 2. Reference list of variables used in this paper.

Variable	Description	Units	Input variables
a_{\max}	Peak horizontal ground acceleration	m s^{-2}	–
CRR	Cyclic resistance ratio	–	MSF, V_{S1} , V_{S1}^*
CSR	Cyclic stress ratio	–	a_{\max} , σ_v , σ'_v , r_d
CTI	Compound topographic index	–	–
d_w	Depth to groundwater	m	–
FS	Factor of safety against liquefaction	–	CRR, CSR
K_m	HAZUS moment magnitude correction factor	–	M_w
K_w	HAZUS groundwater correction factor	–	d_w
K_{Δ}	Displacement correction factor	–	M_w
LPI	Liquefaction potential index	–	FS, z
MSF	Magnitude scaling factor	–	M_w
M_w	Moment magnitude	–	–
MWF	Magnitude weighting factor	–	M_w
ND	Normalized distance to coast (Zhu et al., 2015)	–	–
PGA	Peak horizontal ground acceleration (non-USGS)	g	–
$PGA_{M,SM}$	Peak horizontal ground acceleration (from USGS ShakeMaps)	g	–
$PGD (PGA/PLSC)$	HAZUS expected PGDf for a given liquefaction susceptibility zone	m	PGA, liquefaction susceptibility
PGDf	Permanent ground deformation	m	–
$PGDf_H$	Horizontal permanent ground deformation	m	–
$PGDf_V$	Vertical permanent ground deformation	m	–
P_{ml}	HAZUS proportion of map unit susceptible for a given liquefaction susceptibility zone	–	Liquefaction susceptibility
r_d	Shear stress reduction coefficient	–	z
R_f	Horizontal distance to surface projection of fault rupture	km	–
T_d	Duration between first and last occurrence of $PGA \geq 0.05$ g	s	–
V_S	Shear-wave velocity	m s^{-1}	–
V_{S1}	Stress-corrected shear-wave velocity	m s^{-1}	V_S , σ'_v
V_{S1}^*	Limiting upper value of V_{S1} for cyclic liquefaction occurrence	m s^{-1}	Fines content
$V_{S(0-10)}$	Average shear-wave velocity in top 10 m	m s^{-1}	–
$V_{S(10-20)}$	Average shear-wave velocity between 10 and 20 m	m s^{-1}	–
V_{S30}	Average shear-wave velocity in top 30 m	m s^{-1}	–
z	Depth	m	–
σ_v	Total overburden stress	kPa	z , soil density
σ'_v	Effective overburden stress	kPa	σ_v , z

across the entire study area, thus requiring geostatistical techniques to interpolate. Consequently, this method may only be applicable in a small number of study areas.

To extend the applicability of the LPI model, two approaches are proposed to approximate V_S from more readily available data. The first approach uses V_{S30} , the average shear-wave velocity across the top 30 m of soil, as a constant proxy for V_S for all soil layers. Global estimates for V_{S30} at approximately 674 m grid intervals are open-access from the web-based US Geological Survey Global V_{S30} Map Server (USGS, 2013), so this is an appealing option for desktop assessment. One disadvantage of this approach is that the likelihood of liquefaction occurrence in the LPI method is controlled by the presence of soil layers near the surface with low V_S . Furthermore there is a maximum value of V_S at which liquefaction can occur. Hence the use of V_{S30} as a proxy for all layers will result in an overestimation of V_S , CRR and FS*

at layers closer to the surface and, therefore, an underestimation of LPI and liquefaction risk. This is compounded by the weakness of the USGS V_{S30} dataset, since the data are estimated from topographic slope and the correlation between these two variables is weak.

The second approach proposes the manipulation of the same V_{S30} data to simulate a more realistic V_S profile in which velocities decrease towards the surface rather than being constant. Boore (2004) proposes simple linear empirical functions to extrapolate V_{S30} values in situations where shear-wave velocity data are only known up to shallower depths, based on observations from the United States and Japan. It is proposed to invert the Boore (2004) empirical functions in reverse and use them to back-calculate shallower average shear-wave velocities from V_{S30} data from the USGS Global Server (USGS, 2013). However, it should be noted that since the original function was not developed using or-

thogonal regression, this inversion is an additional source of uncertainty. For simplicity it is proposed to only use the empirical functions to calculate V_{S10} (average shear-wave velocity across top 10 m) and V_{S20} (average shear-wave velocity across top 20 m). The calculated value for V_{S10} can then be used as a proxy for V_S at all soil layers between 0 and 10 m depth and both the V_{S10} and V_{S20} values can be used to determine an equivalent proxy for all soil layers between 10 and 20 m. From manipulation of the Boore (2004) empirical functions and the formula for calculating averaged shear-wave velocities, Eqs. (8) and (9) determine the proxies to be used in the two depth ranges.

$$V_{S(0-10)} = 10^{\left(\frac{\log V_{S30} - 0.042062}{1.0292}\right)} \quad (8)$$

$$V_{S(10-20)} = 2 \times 10^{\left(\frac{\log V_{S30} - 0.025439}{1.0095}\right)} - V_{S(0-10)} \quad (9)$$

In this study, both of these approximations are adopted in addition to the use of known V_S profiles, resulting in the assessment of three implementations of the LPI model.

2.2 HAZUS

HAZUS^{®MH} MR4 (from here on referred to as HAZUS) is a loss estimation software package produced by the National Institute of Building Sciences (NIBS) and distributed by the Federal Emergency Management Agency (FEMA) in the United States. The software accounts for the impacts of liquefaction and the Technical Manual (NIBS, 2003) describes the method used to evaluate the probability of liquefaction.

HAZUS divides the assessment area into six zones of liquefaction susceptibility, from none to very high. This can be done either by interpreting surficial geology from a map and cross-referencing with the table published in the manual or by using an existing liquefaction susceptibility map. Surface geology maps are generally not open-access or free to non-academic organizations and some basic geological knowledge is required to be able to cross-reference mapped information with the zones in the HAZUS table. Hence, the first approach may be problematic for insurers who do not have the requisite in-house expertise. Where liquefaction susceptibility maps are available, unless they use the same zonal definitions as HAZUS, it will be necessary to make assumptions on how zones translate between the third-party map and the manual.

For a given liquefaction susceptibility category, the probability of liquefaction occurrence is given by Eq. (10) (NIBS, 2003), where $P[\text{Liq}|PGA = a]$ is the conditional probability of liquefaction occurrence for a given susceptibility zone at a specified level of peak horizontal ground motion, a ; K_m is the moment magnitude correction factor; K_w is the groundwater correction factor; and P_{ml} is the proportion of map unit susceptible to liquefaction, which accounts for the real variation in susceptibility across similar geologic units. The conditional probability is zero for the susceptibility zone “none”.

For the other susceptibility zones, the conditional probabilities are given by linear functions of acceleration (distinct for each zone), which are not repeated here. The moment magnitude and groundwater correction factors are given by Eqs. (11) and (12):

$$P[\text{Liq}] = \frac{P[\text{Liq}|PGA = a]}{K_m K_w} P_{ml}, \quad (10)$$

$$K_m = 0.0027M_w^3 - 0.00267M_w^2 - 0.2055M_w + 2.9188, \quad (11)$$

$$K_w = 0.022d_w + 0.93, \quad (12)$$

where d_w is the depth to groundwater. The map unit factor is a constant for each susceptibility zone, with values of 0.25, 0.20, 0.10, 0.05, 0.02 and 0, going from “very high” to “none”. In addition to the problems identified for determining liquefaction susceptibility, the HAZUS method also requires water table depth to be known or estimated and judgment on selection of appropriate ground motion prediction equation if ShakeMap or equivalent data are not available.

2.3 Zhu et al. (2015)

Zhu et al. (2015) propose empirical functions to estimate liquefaction probability specifically for use in rapid response and loss estimation. They deliberately use predictor variables that are readily accessible, such as V_{S30} , and do not require any specialist knowledge to be applied. The functions have been developed using logistic regression on data from the earthquakes that occurred in Kobe, Japan, on 17 January 1995 and in Christchurch, New Zealand, on 22 February 2011. Forecasts from the resulting functions have been compared to observations from the 12 January 2010 Haiti earthquake. Since these functions have been developed using data from the Christchurch earthquake, there is an element of circularity in assessing their performance against observations from the same event. However, it is worth noting that the datasets used to develop these functions have not come from the same source as the observations used in this case study. Furthermore, the functions have been calibrated to optimize estimation of the areal extent of liquefaction, whereas in this case study it is the ability of the functions to make site-specific forecasts that is being assessed.

For a given set of predictor variables, the probability of liquefaction is given by the function in Eq. (13), where X is a linear function of the predictor variables. Zhu et al. (2015) propose three linear models that are applicable to the Canterbury region and are adopted in this study: a specific local model for Christchurch, a regional model for use in coastal sedimentary basins (including Christchurch) and a global model that is applicable more generally.

$$P[\text{Liq}] = \frac{1}{1 + e^{-X}} \quad (13)$$

For the global model, the linear predictor function, X_G , is given by Eq. (14), where CTI is the compound topo-

graphic index, used as a proxy for saturation, and can be obtained globally from the USGS Earth Explorer web service (USGS, 2014b). V_{S30} is obtained from the USGS Global Server (USGS, 2013) and $PGA_{M,SM}$ is the product of the peak horizontal ground acceleration from ShakeMap estimates (USGS, 2014a) and a magnitude weighting factor, MWF, given by Eq. (15).

$$X_G = 24.1 + \ln PGA_{M,SM} + 0.355CTI - 4.784 \ln V_{S30} \quad (14)$$

$$MWF = \frac{M_W^{2.56}}{10^{2.24}} \quad (15)$$

For the regional model, the linear predictor function, X_R , is given by Eq. (16), where, additionally, ND is the distance to the coast, normalized by the size of the basin, i.e. the ratio between the distance to the coast and the distance between the coast and inland edge of the sedimentary basin (soil–rock boundary). The location of the inland edge can be estimated from a surface roughness calculation based on a digital elevation model (USGS, 2014b) or by using V_{S30} data such that the inland edge is assumed to be the boundary between NEHRP site classes C (soft rock) and D (stiff soil) (i.e. at $V_S = 360 \text{ m s}^{-1}$). For the Christchurch-specific local model, the linear predictor function, X_L , is given by Eq. (17).

$$X_R = 15.83 + 1.443 \ln PGA_{M,SM} + 0.136CTI - 9.759ND - 2.764 \ln V_{S30} \quad (16)$$

$$X_L = 0.316 + 1.225 \ln PGA_{M,SM} + 0.145CTI - 9.708ND \quad (17)$$

For applicability within the insurance sector, this model presents an advantage over LPI and HAZUS since the only parameter that requires engineering judgment is the selection of ground motion prediction equation if ShakeMap or equivalent data are not available.

3 Model assessment application

This section summarizes the procedure for comparing the model forecasts to observations from the Canterbury earthquake sequence. A brief description is provided of the liquefaction observation dataset and the additional datasets accessed in order to provide the required inputs to the nine models. This is followed by a discussion on the conversion of quantitative model outputs to categorical liquefaction forecasts and an explanation of the diagnostics used to assess model performance.

3.1 Liquefaction observations

The methods described in the previous section are compared for two case studies from the Canterbury earthquake sequence: the M_W 7.1 Darfield earthquake on 4 September 2010 and the M_W 6.2 Christchurch earthquake on 22 February 2011 (GNS Science, 2014), as identified in

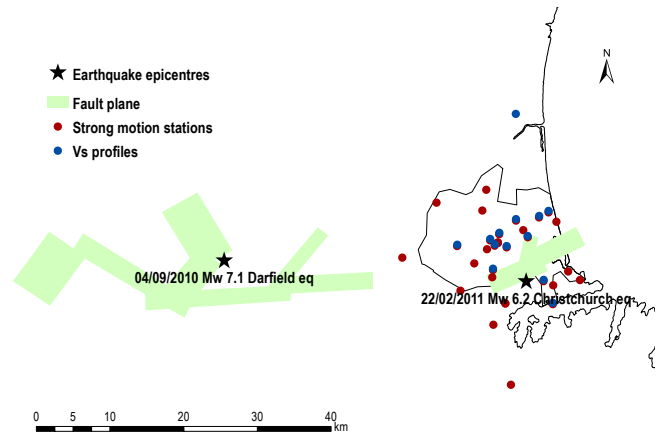


Figure 1. Locations of epicentres and fault planes (Beaven et al., 2012) of the Darfield and Christchurch earthquakes, strong-motion stations from which recordings are used to estimate shaking durations and locations at which shear-wave velocity (V_S) profiles are known (Wood et al., 2011). Note that locations of V_S profiles coincide with strong-motion stations.

Fig. 1. The corresponding peak horizontal ground acceleration contours for each earthquake are shown in Fig. 2.

Surface liquefaction observation data have been obtained from two sources: ground investigation data provided directly from Tonkin & Taylor, geotechnical consultants to the New Zealand Earthquake Commission (EQC) (van Ballegooy et al., 2014), and maps stored within the Canterbury Geotechnical Database (2013a), an online repository of geotechnical data and reports for the region set up by EQC for knowledge sharing after the earthquakes. The data provided by Tonkin & Taylor include records from over 7000 geotechnical investigation sites across Christchurch. After each earthquake, a land damage category is attributed to each site, representing a qualitative assessment of the scale of liquefaction observed. There are six land damage categories, but since this study only investigates liquefaction triggering the categories are converted to a binary classification of liquefaction occurrence. These data are supplemented by the maps from the CGD which show the areal extent of the same land damage categories. To ensure equivalence in the study, all models are applied to the same study area for each earthquake, which is the region for which the input data for all models are available. The study area is divided into a grid of $100 \text{ m} \times 100 \text{ m}$ squares, generating 25 100 observation sites. It is noted, however, that at some locations within Christchurch no liquefaction observations are available so these sites are excluded from the subsequent analysis. As a result, the study area consists of 20 147 sites for the Darfield earthquake and 22 803 sites for the Christchurch earthquake. The observations from the two events are shown in Fig. 3.

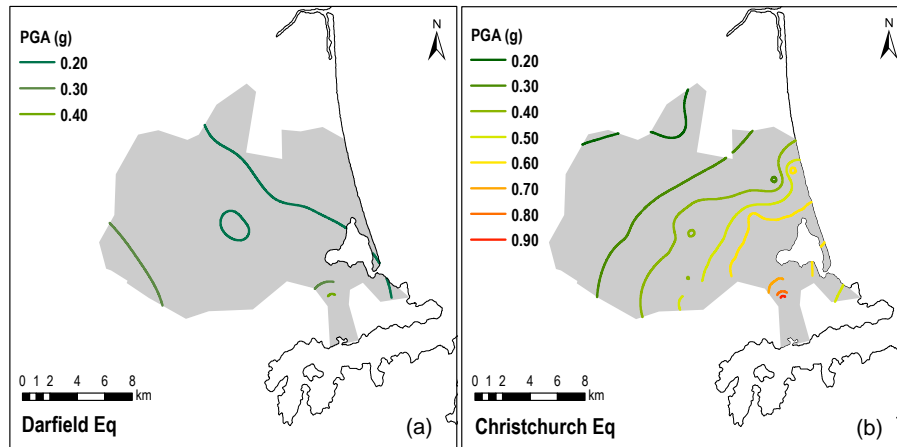


Figure 2. Contours of peak horizontal ground acceleration (PGA) for the Darfield and Christchurch earthquakes (source: Canterbury Geotechnical Database, 2013b).

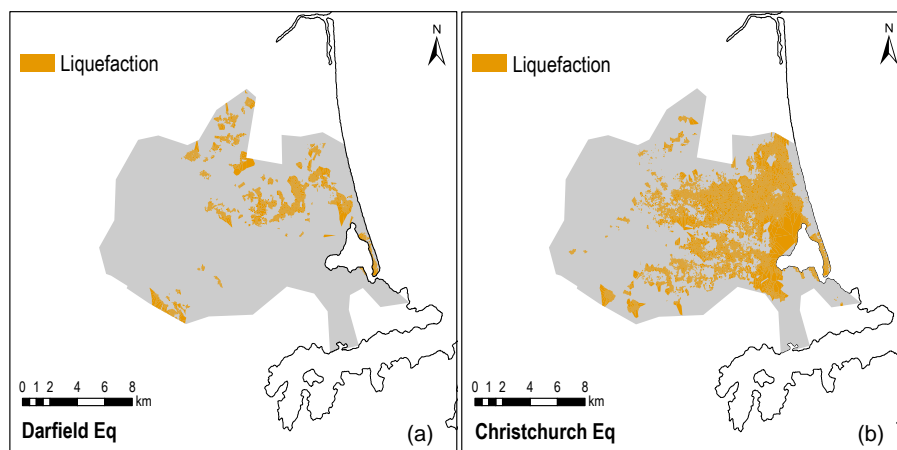


Figure 3. Location of surface liquefaction observations (brown) in Christchurch and surrounding areas due to the Darfield and Christchurch earthquakes, based on data provided by Tonkin and Taylor and published within the Canterbury Geotechnical Database (2013a).

3.2 Forecast model inputs

This study includes three implementations of the LPI model: (1) using known V_S profiles (referred to as LPI1 in this paper); (2) using V_{S30} as a proxy for V_S (LPI2); and (3) using “realistic” V_S profiles simulated from V_{S30} and the Boore (2004) functions (LPI3). The geotechnical investigation data provided by Tonkin & Taylor also include values of LPI calculated at each site from CPT data rather than V_S . Although this approach is not feasible for insurers, for reference its forecasting power is also compared here and this implementation is referred to as LPIref. Historically it has been thought that after liquefaction occurs, soils densify and increase their resistance to future liquefaction. However, Lees et al. (2015) conducted an analysis comparing CPT-based strength profiles and subsequent liquefaction susceptibility at sites in Christchurch both before and after the February 2011 earthquake. They concluded that no significant strengthening

occurred and that the liquefaction risk in Christchurch after the earthquake remained the same as it was beforehand. The study by Orense et al. (2012) came to similar conclusions and therefore, for the purposes of this case study, post-earthquake CPT data are appropriate for assessing liquefaction susceptibility.

A water table depth of 2 m has been assumed across Christchurch, reflecting the averages described by Giovinazzi et al. (2011) – 0 to 2 m in the eastern suburbs and 2 to 3 m in the western suburbs – and soil unit weights of 17 kPa above the water table and 19.5 kPa below the water table are assumed, as suggested by Wotherspoon et al. (2014). V_{S30} data for LPI2 and LPI3 are taken from the USGS web server, with point estimates on an approximately 674 m grid.

Wood et al. (2011) have published V_S profiles for 13 sites across Christchurch obtained using surface wave testing methods. These sites are identified in Fig. 1. In GIS, the pro-

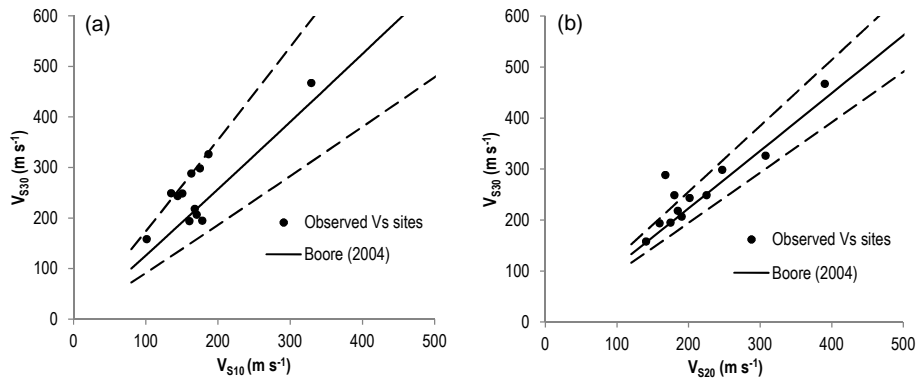


Figure 4. Plots comparing observed V_{S30} with V_{S30} estimated from Boore (2004) equations, with respect to observed V_{S10} (a) and observed V_{S20} (b). The dashed lines represent the 95 % confidence interval around the Boore (2004) relationships. V_{S30} is the average shear-wave velocity in the top 30 m of ground and V_{S10} and V_{S20} are the equivalents at 10 and 20 m depth respectively.

files are converted to point data for each 1 m depth increment from 0 to 20 m, so that each point represents the V_S at that site for a single soil layer and there are a total of 13 points for each soil layer. Ordinary kriging (with log transformation to ensure non-negativity) is applied to the points in each soil layer to create interpolated V_S raster surfaces for each layer. Interpolation over a large area from such a small number of points is likely to result in estimations carrying significant uncertainty. However, from the perspective of commercial loss estimation, this is typical of the type of data that an analyst may be required to work with and so there is value in investigating its efficacy. Whilst Andrus and Stokoe (2000) advise that the maximum V_{S1} can range from 200 to 215 m s^{-1} depending on fines content, subsequent work by Zhou and Chen (2007) indicates that the maximum V_{S1} could range from 200 to 230 m s^{-1} . In the absence of specific fines content data, a median value of 215 m s^{-1} is assumed to be the maximum. In practice, a soil layer may have a value of V_{S1} below this threshold but not be liquefiable because the soil is not predominantly clean sand. Because of the regional scale of this analysis though, site-specific soil profiles (as distinct from V_S profile) are not taken into account in determining whether a soil layer is liquefiable. Goda et al. (2011) suggest the use of “typical” soil profiles to determine the liquefaction susceptibility of a soil layer at a regional scale. Borehole data at sites close to the 13 V_S profile sites are available from the Canterbury Geotechnical Database (2013c). These indicate that in the eastern suburbs of Christchurch, soil typically consists predominantly of clean sand to 20 m depth, with some layers of silty sand. On the western side of Christchurch, however, there is an increasing mix of sand, silt and gravel in soil profiles, particularly at depths down to 10 m. Therefore it is possible, particularly in western suburbs, that the calculated V_{S1} values may indicate liquefiable soil layers when they are in fact not, which would lead to overestimation of LPI and the extent of liquefaction.

For the implementation of model LPI3, it could be argued that rather than using the Boore (2004) relationships to estimate V_S profiles at shallower depths from V_{S30} , the local V_S data published by Wood et al. (2011) could be used to develop a locally calibrated model. This would be preferable from a purely scientific perspective. However, the purpose of this study is to investigate the potential for a simple “global” model for commercial application, and this is defined in part as a model that makes use of methods already in the literature and does not require additional model development. Nevertheless, when using existing models it is useful to assess their applicability to a study area, and the V_S profiles published by Wood et al. (2011) can be used to assess the suitability of the Boore (2004) relationships in Christchurch. Figure 4 shows plots of V_{S30} against V_{S10} and V_{S20} as calculated from the observed profiles and compares these to the Boore (2004) functions. The plots show that the relationships exhibit a small bias towards the underestimation of V_{S30} . When inverted, the application of these relationships to Christchurch may therefore result in the overestimation of V_S at shallower depths and therefore underestimate liquefaction occurrence. However, the majority of observed values are within the 95 % confidence intervals and so the relationships can be deemed to be applicable.

For application of the HAZUS method, liquefaction susceptibility zones have to be identified to determine the values of model input parameters. In this paper liquefaction susceptibility zones are adopted from the liquefaction susceptibility map available from the Canterbury Maps web resource operated by Environment Canterbury Regional Council (ECan, 2014). From the map it is possible to identify four susceptibility zones: “none”, “low”, “moderate” and “high”. However, six susceptibility zones are defined by HAZUS (NIBS, 2003). Since the Canterbury zones cannot be subdivided, it is necessary to map the Canterbury zones onto four of the HAZUS zones. In HAZ1 the zones are mapped simply by matching names; in HAZ2, the “low” and “high” zones in

Table 3. Conversion between Canterbury and HAZUS liquefaction susceptibility zones for three implementations of HAZUS method. Refer to Table 4 for descriptions corresponding to model acronyms.

Canterbury liquefaction susceptibility zones	Equivalent HAZUS liquefaction susceptibility zone and expected settlement amplitude					
	Model HAZ1		Model HAZ2		Model HAZ3	
	Zone	Expected settlement (cm)	Zone	Expected settlement (cm)	Zone	Expected settlement (cm)
None	None	0	None	0	None	0
Low	Low	2.5	Very low	0	Average low and very low	1.25
Moderate	Moderate	5	Moderate	5	Moderate	5
High	High	15	Very high	30	Average high and very high	22.5

Canterbury are mapped to the more extreme “very low” and “very high” zones in HAZUS; and in HAZ3, the relevant input parameters for each zone are taken to be the average of those identified in HAZ1 and HAZ2. The mapping between susceptibility zones in each of the implementations described in Table 3. As with the LPI model, depth to water table is assumed to be 2 m across Christchurch.

Three models proposed by Zhu et al. (2015) are compared in this paper: (1) the global model (referred to as ZHU1), (2) the regional model (ZHU2) and (3) the local model (ZHU3). The PGA “shakefields” from the Canterbury Geotechnical Database (2013b) are used as equivalents to the USGS ShakeMap. CTI (USGS, 2014b), at approximately 1 km resolution and V_{S30} (USGS, 2013) are downloaded from the relevant USGS web resources. In total nine model implementations are being compared, based on three general approaches (see Table 4).

3.3 Site-specific forecasts

When using probabilistic forecasting frameworks, one can interpret the calculated probability as a regional parameter that describes the spatial extent of liquefaction rather than discrete site-specific forecasts, and indeed Zhu et al. (2015) specifically suggest that this is how their model should be interpreted. So, for example, one would expect 30 % of all sites with a liquefaction probability of 0.3 to exhibit liquefaction and 50 % of all sites with a liquefaction probability of 0.5. However, when using liquefaction forecasts as a means to estimate structural damage over a wide area, it is useful to know not just the number of liquefied site but also where these sites are. This is particularly important for infrastructure systems since the complexity of these networks means that damage to two identical components can have significantly different impacts on overall systemic performance depending on the service area of each component and the level of redundancy built in.

There are two ways to generate site-specific forecasts from probabilistic assessments. One approach is to group sites together based on their liquefaction probability and then randomly assign liquefaction occurrence to sites within the group based on that probability, e.g. by sampling a uniformly distributed random variable. This method is good for ensuring that the spatial extent of the site-specific forecasts reflect the probabilities, but since the locations are selected randomly it has limited value for comparison of forecasts to real observations from past earthquakes. It can be more useful for generating site-specific forecasts for simulated earthquake scenarios.

Another method is to set a threshold value for liquefaction occurrence, so all sites with a probability above the threshold are forecast to exhibit liquefaction and all sites with a probability below the threshold are forecast to not exhibit liquefaction. The disadvantage of this approach is that the resulting forecasts may not reflect the original probabilities. For example if the designated threshold probability is 0.5 and all sites have a calculated probability greater than this (even if only marginally), then every site will be forecast to liquefy. Conversely if all sites have a probability below 0.5, then none of the sites will be forecast to liquefy. However, since there is no random element to the determination of liquefaction occurrence, the forecasts are more definitive in spatial terms and hence more useful for the this comparative site-specific study. Although not strictly a probabilistic framework, thresholds can also be used to assign liquefaction occurrence based on LPI by determining a value above which liquefaction is assumed to occur.

For all of the methods however, the issue arises of what value the thresholds should take. No guidance is given for HAZUS, whilst Zhu et al. (2015) propose a threshold of 0.3 to preserve spatial extent, although they also consider thresholds of 0.1 and 0.2. In their original study, Iwasaki et al. (1984) suggest critical values of LPI of 5 and 12 for liquefaction and lateral spreading respectively. However, other lo-

Table 4. Liquefaction forecasting models compared in this paper.

Model	Description
LPI1	Liquefaction potential index (LPI) with known shear-wave velocity, V_S , profiles
LPI2	LPI with average shear-wave velocity in the top 30 m, V_{S30} , as a proxy for V_S
LPI3	LPI with simulated V_S profiles
LPIref	LPI calculated from standard penetration test (SPT) results
HAZ1	HAZUS with “direct” conversion of susceptibility zones
HAZ2	HAZUS with “extreme” susceptibility zones
HAZ3	HAZUS with “average” conversion of susceptibility
ZHU1	Global model by Zhu et al. (2015)
ZHU2	Regional model by Zhu et al. (2015)
ZHU3	Local model by Zhu et al. (2015)

calized studies where the LPI method has been applied have found alternative criteria that provide a better fit for observed data as summarized by Maurer et al. (2014). Since there is uncertainty in the selection of threshold values, this study investigates a range of values for each model. Both the observation and forecast datasets are binary classifications, so standard binary classification measures based on 2×2 contingency tables are used to compare performance.

3.4 Performance diagnostics

Comparison of binary classification forecasts with observations is made by summarizing data into 2×2 contingency tables for each model. The contingency table identifies the true positives (TP), true negatives (TN), false positives (FP, type I error) and false negatives (FN, type II error). A good forecasting model would forecast both positive (occurrence of liquefaction) and negative (non-occurrence of liquefaction) results well. Diagnostic scores for each model can be calculated based on different combinations and functions of the data in the contingency tables. The true positive rate (TPR or sensitivity) is the ratio of true positive forecasts to observed positives. The true negative rate (TNR or specificity) is the ratio of true negative forecasts to observed negatives. The false positive rate (FPR or fallout) is the ratio of false positive forecasts to true negatives. A useful model would have a high TPR and TNR (> 0.5) and low FPR (< 0.5).

The results presented in a contingency table and associated diagnostic scores assume a single initial threshold value. However, further statistical analysis is undertaken to optimize the thresholds in accordance with the observed data. For a single model, at a specified threshold, the receiver operating characteristic (ROC) is a graphical plot of TPR against FPR. The line representing $TPR = FPR$ is equivalent to random guessing (known as the chance or no-discrimination line). A good model has a ROC above and to the left of the chance line, with perfect classification occurring at (0,1). The diagnostic scores for each model are re-calculated with different thresholds and the resulting ROC values are plotted as a curve for the model. Since better models have points towards

the top left of the plot, the area under the ROC curve (AUC) is a generalized measure of model quality that assumes no specific threshold. Since the diagonal of the plot is equivalent to random guessing, $AUC = 0.5$ suggests a model has no value, while $AUC = 1$ is a perfect model. For a single point on the ROC curve, Youden’s J statistic is the height between the point and the chance line. The point along the curve which maximizes the J statistic represents the TPR and FPR values obtained from the optimum threshold for that model.

In addition to comparing the performance of simplified models to each other, it is useful to measure the absolute quality of each model. Simply counting the proportion of correct forecasts does not adequately measure model performance since it does not take into account the proportion of positive and negative observations; e.g. a negatively biased model will result in a high proportion of correct forecasts if the majority of observations are negative. The Matthews correlation coefficient (MCC) is more useful for cases where there is a large difference in the number of positive and negative observations (Matthews, 1975). It is proportional to the chi-squared statistic for a 2×2 contingency table and its interpretation is similar to Pearson’s correlation coefficient, so it can be treated as a measure of the goodness of fit of a binary classification model (Powers, 2011). From contingency table data, MCC is given by Eq. (18).

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (18)$$

4 Results

This section summarizes the results of the model applications using contingency table analysis. The results are first presented for analysis using a set of initial assumed thresholds for positive forecasts and subsequently for analysis in which thresholds are optimized for performance. The sensitivity of the forecasts to variation of V_{S30} and PGA inputs is also assessed.

Table 5. Summary of contingency table data and diagnostic scores for all models using initial threshold estimates, including “LPI” models subject to sensitivity test without Juang et al. (2005) correction factors being applied to the factor of safety. Refer to Table 4 for descriptions corresponding to model acronyms.

Model	True positives (TP)	True negatives (TN)	False positives (FP)	False negatives (FN)	True positive rate (TPR)	True negative rate (TNR)	False positive rate (FPR)
LPI1	6345	25 685	9442	1478	0.811	0.731	0.269
LPI2	147	35 063	64	7676	0.019	0.998	0.002
LPI3	4287	30 578	4549	3536	0.548	0.870	0.130
LPIref	5964	20 826	14 301	1859	0.762	0.593	0.407
HAZ1	0	35 127	0	7823	0.000	1.000	0.000
HAZ2	0	35 127	0	7823	0.000	1.000	0.000
HAZ3	0	35 127	0	7823	0.000	1.000	0.000
ZHU1	1880	33 483	1644	5943	0.240	0.953	0.047
ZHU2	3135	31 931	3196	4688	0.401	0.909	0.091
ZHU3	2754	31 017	4110	5069	0.352	0.883	0.117
LPI2b	610	34 902	225	7213	0.078	0.994	0.006
LPI3b	6068	20 509	14 618	1755	0.806	0.584	0.416

4.1 Contingency table analysis – initial thresholds

An initial set of results using 5 as a threshold value for the LPI models, 0.3 as a threshold for the ZHU models and 0.5 as a threshold value for the HAZUS models is shown in Table 5, alongside the corresponding diagnostic scores.

The LPI1, LPI3 and LPIref models are the only models that meet the criteria of having TPR and TNR > 0.5 and FPR < 0.5, with the LPI1 model performing better despite being based on V_S rather than ground investigation data. Table 3 shows that all HAZUS models are very good at forecasting non-occurrence of liquefaction. However, this is only due to the fact that they are forecasting no liquefaction all the time, and so their ability to forecast the occurrence of liquefaction is extremely poor. The high TNR but relatively low TPR of the three ZHU models indicate that they all show a bias towards forecasts of non-occurrence of liquefaction. The difference between TPR and TNR is indicative of the level of bias in the model and in this regard, ZHU2, the regional model shows less bias than in ZHU1, the global model, as would be expected. The bias in the ZHU2 and ZHU3 models is approximately similar although ZHU2 performs slightly better.

The LPI2 model, using V_{S30} as a proxy, also shows a very strong bias towards forecasting non-occurrence, which is expected since V_{S30} generally provides an overestimate of V_S for soil layers at shallow depth. At sites where the soil profile of the top 30 m is characterized by some liquefiable layers at shallow depth with underlying rock or very stiff soil (e.g. in western and central areas close to the inland edge of the sedimentary basin), V_{S30} will be high. Hence, this leads to false classification of shallow layers as non-liquefiable. The LPI3 model with simulated V_S profiles exhibits good performance in forecasting non-occurrence of liquefaction and correctly forecasts just over half of the positive liquefaction observa-

tions, indicating bias towards negative forecasts. Although the V_S profiles generated through this approach are more realistic than using a constant V_{S30} value, the V_S at each layer is related to V_{S30} . Therefore, at sites characterized by a high V_{S30} value with low V_S values at shallow depths, even using Eqs. (8) and (9) may not estimate sufficiently low values of V_{S1} to classify the shallow layers as liquefiable. Another factor in the LPI models is the use of the bias-correction factor proposed by Juang et al. (2005). Whilst this correction factor is appropriate when actual V_S profiles are used, as in LPI1, it may not be appropriate for LPI2 and LPI3 where non-conservative proxies for V_S are used and the resulting misclassification of liquefiable soil layers balances the conservativeness of the Andrus and Stokoe (2000) CRR model. The sensitivity of the models to the correction factor is investigated by reproducing the contingency tables for LPI2 and LPI3 with the same threshold values but ignoring the correction factor for FS. These models are referred to as LPI2b and LPI3b and the new contingency table analysis is presented in Table 5.

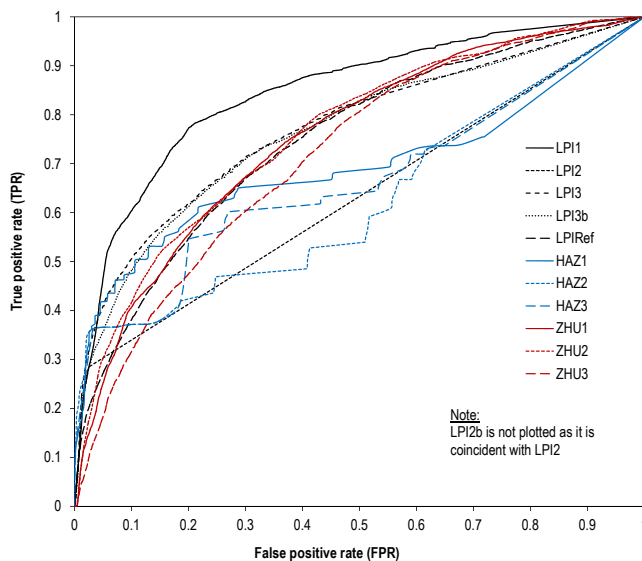
These results show that not using the bias correction makes little difference to the performance of LPI2, as LPI2b still exhibits an extremely strong bias towards forecasting non-occurrence of liquefaction. For LPI3, however, the difference is more significant. Without the correction factor, the TPR and TNR values for LPI3b reverse, with only just over half negative liquefaction occurrences being correctly forecast. LPI3b therefore exhibits a bias towards positive liquefaction forecasts and so it confers no advantage over LPI3.

4.2 Contingency table analysis – optimized thresholds

The results in Tables 3 and 4 demonstrate the performance of each model with a single initial threshold value. ROC analysis is used to optimize the thresholds and curves for the 11

Table 6. Model quality diagnostics and optimum threshold values for each model from ROC curves. Refer to Table 4 for descriptions corresponding to model acronyms.

Model	Area under curve (AUC)	<i>J</i> statistic	Threshold	True positive rate (TPR)	True negative rate (TNR)
LPI1	0.845	0.573	7	0.774	0.799
LPI2	0.630	0.122	1	0.131	0.991
LPI2b	0.630	0.206	1	0.224	0.982
LPI3	0.772	0.420	4	0.581	0.839
LPI3b	0.766	0.414	10	0.617	0.797
LPIref	0.748	0.366	6	0.689	0.678
HAZ1	0.679	0.238	0.1	0.073	0.999
HAZ2	0.608	0.316	0.1	0.134	0.997
HAZ3	0.661	0.315	0.1	0.133	0.998
ZHU1	0.753	0.355	0.1	0.556	0.799
ZHU2	0.760	0.371	0.1	0.767	0.604
ZHU3	0.718	0.306	0.1	0.712	0.594

**Figure 5.** Receiver operating characteristic (ROC) curves for forecasting models. Refer to Table 4 for descriptions corresponding to model acronyms.

simplified models and reference model are generated using the ROC package in R (Sing et al., 2005), as shown in Fig. 5. For this study, the threshold for the LPI models is assumed to be a whole number, while for the HAZ and ZHU models, the threshold is assumed to be a multiple of 0.05 subject to a minimum value of 0.1, which is the minimum applied by the Zhu et al. (2015). The AUC values, maximum *J* statistics, optimum thresholds and corresponding TPR and TNR values for all models are shown in Table 6.

With optimized thresholds all the LPI models, except LPI2, and all the ZHU models meet the TPR and TNR criteria (> 0.5). All HAZ models and both versions of LPI2 have AUC values closer to the “no value” criterion, suggesting

that the problems with these models lie not just with threshold selection but more fundamentally with their composition and/or relevance to the case study (noting that the HAZ models have been developed for analysis in the United States). The reason these are to the left of the chance line is that they are forecasting non-occurrence of liquefaction at nearly every site and hence they are guaranteed a low FPR value. LPI1 is the best-performing model according to both of the ROC diagnostics and although the optimum threshold value of 7 is higher than proposed by Iwasaki et al. (1984), it is within the range for marginal liquefaction – 4 to 8 – proposed by Maurer et al. (2014) and so may be considered plausible. The two versions of the LPI3 model perform similarly and have reasonable diagnostic scores but LPI3, with the correction factor, produces a more plausible optimum threshold value of 4. It is noted, however, that although the optimum threshold for LPI3b is 10, the TPR and TNR criteria are met with a threshold of 4 but with a lower model performance and greater positive forecast bias (*J* statistic = 0.344, TPR = 0.806, TNR = 0.538).

The ZHU1 and ZHU2 models perform reasonably with AUC values and *J* statistics slightly lower than the LPI3 models, but the optimum thresholds are at the minimum of the range that has been investigated, confirming the degree to which these models underestimate liquefaction occurrence. The ZHU2 model also meets the TPR and TNR criteria with a threshold value of 0.2, albeit with a greater forecast bias (*J* statistic = 0.370, TPR = 0.555, TNR = 0.815). The ZHU3 model, despite being specific to Christchurch, does not perform as well as ZHU1 or ZHU2. There are potential reasons for this anomaly, such as that the ZHU models were calibrated to preserve the extent of liquefaction rather than to make site-specific forecasts or because the data used to develop the models have not come from the same source as the observation data used for comparison. Therefore these results do not contradict or invalidate the original findings of Zhu et

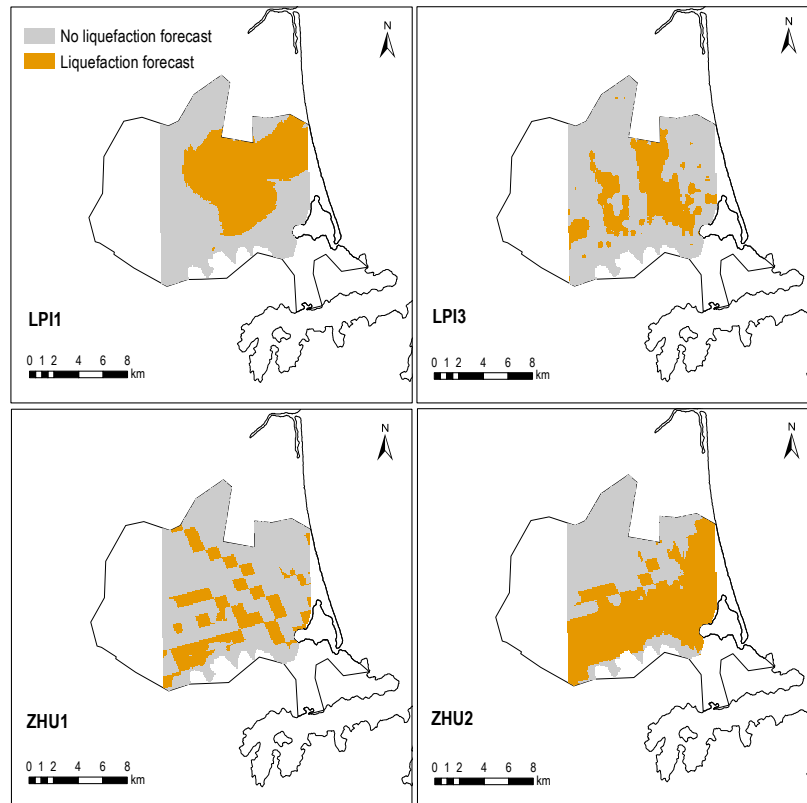


Figure 6. Maps of liquefaction forecasts from selected models for the Darfield earthquake. Unshaded areas are where no forecast was made due to unavailability of input data. Refer to Table 4 for descriptions corresponding to model acronyms.

al. (2015). The absolute quality of models is evaluated by calculating MCC. In the preceding analysis, the best-performing model is LPI1 and this has a value of $MCC = 0.48$. The correlation is only moderate but nevertheless indicates that the model is better than random guessing. As part of a rapid assessment or desktop study for insurance purposes, this may be sufficient. LPI3 and LPI3b have $MCC = 0.380$ and 0.357 respectively, whilst LPIref has $MCC = 0.29$.

4.3 Mapping of model forecasts

The maps in Figs. 6 and 7 show how forecasts of liquefaction occurrence, relating to the Darfield and Christchurch earthquakes respectively, are distributed across the city for four of the best-performing models identified in Table 6: LPI1, LPI3, ZHU1 and ZHU2. Figure 3 shows that a greater extent of liquefaction was observed in the Christchurch earthquake than in the Darfield earthquake and this is reflected by all four models represented in Figs. 6 and 7. However, for both earthquakes, each of the models forecasts a greater extent of liquefaction than was observed. In the Darfield earthquake, most of the liquefaction was observed in the north and east of the city. Whilst to some degree this spatial distribution is matched by model LPI1, the remaining models do not represent the observed distribution well. In particular the models

ZHU1 and ZHU2 estimate a greater proportion of liquefaction in the south of the city. In the Christchurch earthquake, liquefaction was mostly observed in the eastern suburbs of the city. All the models forecast the majority of liquefaction to occur in these areas, although model ZHU2 forecasts more liquefaction occurring in western suburbs than actually occurred, while model ZHU1 forecasts no liquefaction occurring to the west of the city at all. The spatial distributions of the forecasts from the LPI models exhibit only limited accuracy, yet they are better than the forecasts from the two ZHU models. This can be explained partly by the fact the LPI method is designed for site-specific estimation, whereas the ZHU models have been calibrated to optimize the extent rather than the location of liquefaction.

4.4 Sensitivity test – V_{S30}

The sensitivity of the forecasts to variation in V_{S30} is assessed for models LPI3 and ZHU1. LPI3 is the best-performing model that requires V_{S30} and ZHU2 is the best-performing ZHU model. The forecasting procedure and contingency table analysis for the two models are repeated for two scenarios, one where V_{S30} is decreased by 10% at all sites and one where V_{S30} is increased by 10% at all sites.

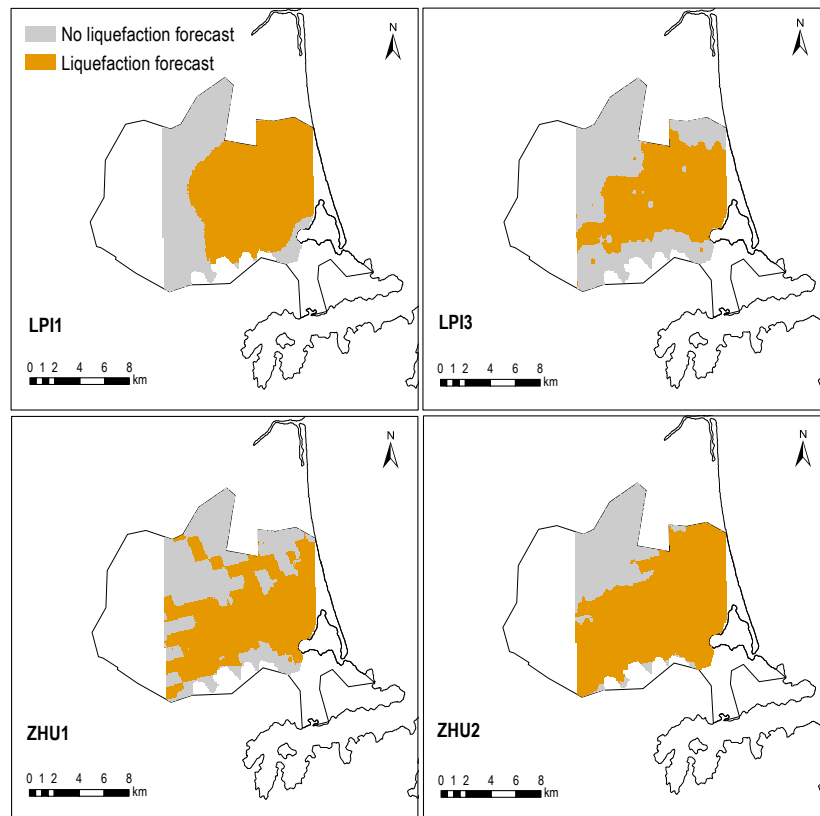


Figure 7. Maps of liquefaction forecasts from selected models for the Christchurch earthquake. Unshaded areas are where no forecast was made due to unavailability of input data. Refer to Table 4 for descriptions corresponding to model acronyms.

In the scenario where V_{S30} is decreased, the TPR for model LPI3 increases to 0.819 with a threshold of 4 (the optimized threshold from Table 6), while the TNR decreases to 0.536, effectively reversing the bias demonstrated by the original model. The J statistic reduces significantly to 0.356 indicating lower performance than the original model. With the new V_{S30} values, the optimized threshold increases to 9, with J statistic = 0.426, which is higher than the original model, TPR = 0.654 and TNR = 0.773. When V_{S30} is increased, TPR = 0.308 with a threshold of 4, which is lower than criterion for good performance (TPR > 0.5), TNR = 0.974. This demonstrates a strengthening of the negative bias in the original model and poor performance since the J statistic reduces to 0.282. The optimum threshold changes to 1, yet even with this threshold, while the J statistic improves to 0.388, TPR = 0.489, which is still below the performance criterion. These results show that LPI3 forecasts are sensitive to variation in V_{S30} . Therefore, although currently the optimum LPI3 threshold for Christchurch has been identified as 4, if in the future more accurate V_{S30} becomes available, then the analysis presented in this paper should be repeated to recalibrate model LPI3 with a new optimum threshold.

Model ZHU2 experiences much smaller changes as a result of changes to V_{S30} . When V_{S30} is decreased, and with

a threshold of 0.1 (the optimized threshold from Table 6), TPR = 0.820, TNR = 0.532 and J statistic = 0.352. When V_{S30} is increased, TPR = 0.700, TNR = 0.662 and J statistic = 0.362. For both scenarios all performance criteria are met and there only small reductions in J statistic. When the models are optimized, the thresholds change to 0.25 for the decrease (J statistic = 0.370) and to 0.15 for the increase (J statistic = 0.368). These results suggest that ZHU2 forecasts are relatively stable in response to variations in V_{S30} , but if more accurate V_{S30} becomes available in the future, then some performance improvement can be achieved through recalibration of the optimum threshold.

4.5 Sensitivity test – PGA

The sensitivity of the forecasts to uncertainty in PGA measurements is also assessed for models LPI3 and ZHU1. In the two sensitivity test scenarios, the forecasting procedure and contingency table analysis are repeated for two scenarios: one where PGA is decreased by 10% at all sites and one where PGA is increased by 10% at all sites.

In the scenario where PGA is decreased, the TPR for model LPI3 decreases to 0.503 with a threshold of 4, while the TNR increases to 0.905 and there is only a small re-

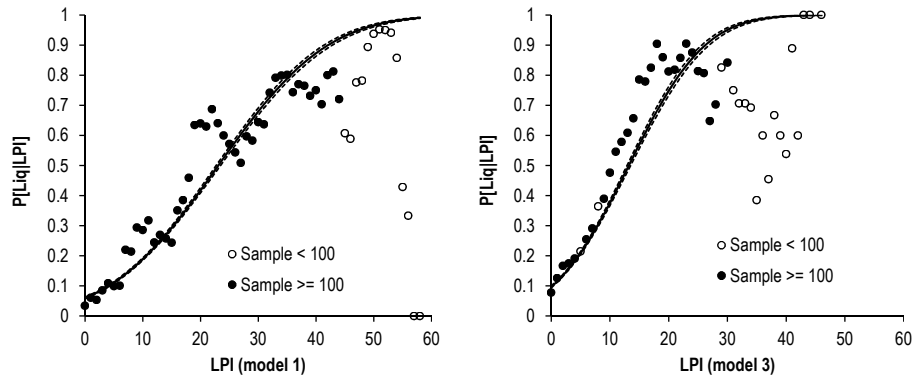


Figure 8. Plots of liquefaction probability against liquefaction potential index (LPI) derived from site-specific observations by generalized linear model with probit link function for two best-performing LPI models. Plots also display the observed liquefaction rates at each LPI value and classified by sample size.

duction in J statistic to 0.408. The optimized threshold decreases to 2, with J statistic = 0.424, which is higher than the original model, TPR = 0.594 and TNR = 0.830. When PGA is increased, TPR = 0.652 with a threshold of 4 and TNR = 0.765, with corresponding J statistic = 0.417. The optimum threshold changes to 6, with J statistic = 0.419, TPR = 0.576 and TNR = 0.843. In general, changes in PGA do affect the scores but, in all cases, the changes are relatively small, particularly with respect to the J statistic, and the performance criteria are still met.

Model ZHU2 also experiences small changes as a result of changes to PGA. When PGA is decreased, and with a threshold of 0.1, TPR = 0.725, TNR = 0.637 and J statistic = 0.362. When PGA is increased, TPR = 0.798, TNR = 0.574 and J statistic = 0.372, which is a small increase over the original model. For both scenarios all performance criteria are met and there only small changes to J statistic. When the models are optimized, the threshold changes to 0.2 for the decrease scenario (J statistic = 0.369), but for the increase scenario the optimum threshold is still 0.1. These results suggest that both LPI3 and ZHU2 forecasts are relatively stable in response to variations in PGA and so while small uncertainties in PGA measurements will change the rates of true positive and true negative forecasts, overall performance in terms of J statistic remains similar.

5 Probability of liquefaction

When the threshold-based approach to liquefaction occurrence is applied to the LPI models, it provides a deterministic forecast. This may be considered sufficient for the simplified regional-scale analyses conducted for catastrophe modelling and loss estimation. However, a modeller may also want to establish a probabilistic view of liquefaction risk by relating values of LPI to probability of liquefaction occurrence. Since the occurrence of liquefaction at a site is a binary classification variable, it can be modelled by a Bernoulli distri-

bution with probability of liquefaction, p , which depends on the value of LPI. With data from past earthquakes, functions relating p to LPI can be derived using a generalized linear model with probit link function. The probability of liquefaction occurring given a particular value of LPI, λ , is given by Eq. (19), where Φ is cumulative normal probability distribution function and Y^* is the probit link function given by Eq. (20).

$$p_{\text{LPI}=\lambda} = P[\text{Liq}|\text{LPI} = \lambda] = \Phi(Y^*) \quad (19)$$

$$Y^* = \beta_0 + \beta_1 \lambda \quad (20)$$

The link function is a linear model with LPI as a predictor variable and is derived from the individual site observations. Figure 8 displays the relationships between liquefaction probability and LPI fit by this method for the two best-performing LPI models, LPI1 and LPI3, including 95% confidence intervals. The relationships are accompanied by plots of the observed liquefaction rates, aggregated at each value of LPI. The plot for model LPI3 shows greater scatter of observed rates around the fit line than the plot for model LPI1, although in both cases the confidence interval is very narrow, which is a reflection of the large sample size. The confidence interval for LPI1 (± 0.0014) is slightly narrower than the confidence interval for LPI3 (± 0.0021), indicating that LPI1 is the better model for estimating liquefaction probability, just as it is better at forecasting liquefaction occurrence by LPI threshold. For both models, the observed rates that are furthest away from the best-fit line are predominantly those that are based on smaller sample sizes (arbitrarily defined here as 100). These have less influence on the regression – since the use of individual site observations implicitly gives more weight to observations in the region of LPI values for which sample sizes are larger. Furthermore, the observed rates are themselves more unreliable for smaller sample sizes. For example, for model LPI1, observations based on more than 100 samples have an average margin of error of 0.05, whereas

Table 7. Coefficients of link function and summary of contingency table analysis for the two best-performing LPI models. Refer to Table 4 for descriptions corresponding to model acronyms.

Model	β_1	B_0	True positive rate (TPR)	True negative rate (TNR)	J statistic	Area under curve (AUC)
LPI1	0.067	-1.555	0.683	0.869	0.551	0.843
LPI3	0.098	-1.299	0.784	0.856	0.641	0.766

the average margin of error for smaller samples is 0.19. For model LPI3, observations based on more than 100 samples have an average margin of error of 0.05 and this increases to 0.22 when considering the observations based on smaller samples.

The Hosmer–Lemeshow test (Hosmer and Lemeshow, 1980) is a commonly used procedure for assessing the goodness of fit of a generalized linear model when the outcome is a binary classification. However, Paul et al. (2013) show that the test is biased with respect to large sample sizes, with even small departures from the proposed model being classified as significant and consequently recommend that the test is not used for sample sizes above 25 000. Pseudo- R^2 metrics are also commonly used to test model performance (Smith and McKenna, 2013), but these compare the proposed model to a null intercept-only model rather than comparing the model forecasts to observations. Although the purpose of the analysis in this section is to relate LPI to liquefaction probabilistically, contingency table analysis with a threshold probability to determine liquefaction occurrence remains an appropriate technique to test the fit of the model (Steyerberg et al., 2010). Assuming a threshold probability of 0.5, Table 7 presents summary statistics from the contingency table analysis of each model and also the coefficients of the corresponding probit link function.

Both models have values of TPR and TNR above 0.5 and the values are of a similar order to those obtained in Table 5 for the same models. The exception is the TPR for model LPI3, which is significantly higher when the threshold probability is used and eliminates much of the bias towards negative forecasts. The difference in values of AUC between Tables 6 and 7 are negligible but the J statistic for LPI3 with a threshold probability of 0.5 is considerably higher than the J statistic for the optimal threshold found for LPI3 in Table 6. This suggests that LPI3 is best implemented as a probabilistic model for liquefaction occurrence. Overall these statistics indicate that both of the probabilistic LPI models proposed are good fits to the observed data.

6 Permanent ground deformation

The preceding sections have analysed methods for forecasting liquefaction triggering, but for assessing the fragility of structures and infrastructure it is more informative to be able to estimate the scale of liquefaction, in terms of the PGDf. In

fact, fragility functions for liquefaction-induced damage are commonly expressed in these terms (Pitilakis et al., 2014). A summary of the available approaches for quantifying PGDf is provided by Bird et al. (2006), who also compare approaches for lateral movement, settlement and combined movement (volumetric strain). The majority of these approaches require detailed geotechnical data as inputs (e.g. median particle size, fines content). The likelihood that insurers possess or are able to acquire such data is low, which means that these approaches are not suitable for regional-scale rapid assessment. The lack of simplified models is not surprising given the small number of models that exist for liquefaction triggering assessment and that by definition measuring the scale of liquefaction is more complex. From the available models in the literature, there are three that can be applied without the need for detailed geotechnical data: the EPOLLS regional model for lateral movement (Rauch and Martin, 2000) and the HAZUS models for lateral movement and vertical settlement (NIBS, 2003). To demonstrate the challenge faced by insurers looking to improve their liquefaction modelling capability, these models are compared to PGDf observations from the Darfield and Christchurch earthquakes. It should be noted that the HAZUS model has been developed specifically for the United States and the empirical data used to develop its constituent parts come mainly from California and Japan. The EPOLLS model is based on empirical data from the United States, Japan, Costa Rica and the Philippines.

6.1 Vertical settlement

A time series of lidar surface data for Christchurch has been produced from aerial surveys over the city, initially prior to the earthquake sequence in 2003, and subsequently repeated after the Darfield and Christchurch earthquakes. The surveys are obtained from the Canterbury Geotechnical Database (2012a). The lidar surveys recorded the surface elevation as a raster at 5 m cell resolution. The difference between the post-Darfield earthquake survey and the 2003 survey represents the vertical movement due to the Darfield earthquake. Similarly the difference between the post-Christchurch earthquake and the post-Darfield earthquake surveys represents the movement due to the Christchurch earthquake. In addition to liquefaction, elevation changes recorded by lidar can also be caused by tectonic movements. Therefore, to evaluate the vertical movement due to liquefaction effects only (PGDf_v), the differences between lidar

Table 8. Summary statistics of vertical ground deformation (PGD_{fV}) estimates for Darfield and Christchurch earthquakes from HAZUS models. Refer to Table 4 for descriptions corresponding to model acronyms.

Score	Observed	Vertical permanent ground deformation, PGD _{fV} (m)		
		HAZ1	HAZ2	HAZ3
Pearson R^2	n/a	0.064	0.051	0.058
Mean	0.118	0.003	0.008	0.005
Minimum	0.000	0.000	0.000	0.000
Lower quartile	0.051	0.000	0.000	0.000
Median	0.100	0.001	0.000	0.000
Upper quartile	0.162	0.004	0.004	0.004
Maximum	1.464	0.022	0.066	0.043
Residual mean	n/a	-0.114	-0.110	-0.112
Root-mean-square error	n/a	0.146	0.142	0.144

n/a = not applicable.

surveys have been corrected to remove the effect of the tectonic movement. Tectonic movement maps have been acquired from the Canterbury Geotechnical Database (2013d). The only simplified method for calculating vertical settlement is from HAZUS (NIBS, 2003), in which the settlement is the product of the probability of liquefaction, as in Eq. (10), and the expected settlement amplitude, which varies according to liquefaction susceptibility zone, as described in Table 3.

The HAZUS model is applied with each of the three implementations used for forecasting liquefaction probability in the liquefaction triggering analysis. Summary statistics of the PGD_{fV} estimates from each implementation are presented in Table 8. This shows that the HAZUS model significantly underestimates the scale of liquefaction, regardless of how liquefaction susceptibility zones are mapped between the Canterbury and HAZUS classifications. The residuals have a negative mean in each implementation indicating an underestimation bias. Furthermore, the maximum value estimated by HAZ1 and HAZ3 is smaller than the observed lower quartile. The coefficient of determination is also extremely low in each case, implying that there is little or no value in the estimates. It is important to note that there is a measurement error in the lidar data itself of up to 150 mm, as well as a uniform probability prediction interval around the HAZUS estimates. However, even when using the upper bound of the HAZUS estimates (2 times the mean), only around 50 % of estimates fall within the observation error range. These results suggest that the HAZUS model for estimating vertical settlement is not suitable for application in Christchurch.

6.2 Lateral spread

The lidar surveys for Christchurch also record the locations of reference points within a horizontal plane and the differences between these data have been used to generate maps

identifying the lateral displacements caused by each earthquake on a grid of points at 56 m intervals. Similarly to the elevation data, the lateral displacements have to be corrected for tectonic movements, although in this case the corrected maps have been obtained directly from the Canterbury Geotechnical Database (2012b).

The HAZUS model (NIBS, 2003) for estimating ground deformation due to lateral spread is given by Eq. (19), where K_{Δ} is a displacement correction factor, which is a cubic function of earthquake magnitude, and the term on the right-hand side is the expected ground deformation for a given liquefaction susceptibility zone, which is a function of the normalized peak ground acceleration (observed PGA divided by liquefaction triggering threshold PGA for that zone). The formulae for calculating these terms are not repeated here but can be found in the HAZUS manual (NIBS, 2003).

$$\text{PGDf}_H = K_{\Delta} \times E [\text{PGD} | (\text{PGA}/\text{PL}_{SC}) = a] \quad (21)$$

The EPOLLS suite of models for lateral spread (Rauch and Martin, 2000) includes proposed relationships for estimating ground deformation at a regional scale (least complex), at site-specific scale without detailed geotechnical data and at site-specific scale with detailed geotechnical data (most complex). In the regional EPOLLS model, PGD_{fH} is given by Eq. (20), where R_f is the shortest horizontal distance to the surface projection of the fault rupture, and T_d is the duration of ground motion between the first and last occurrence of accelerations ≥ 0.05 g at each site.

$$\text{PGDf}_H = (0.613M_W - 0.0139R_f - 2.42\text{PGA} - 0.01147T_d - 2.21)^2 + 0.149 \quad (22)$$

Durations have been calculated from ground motion records (at 0.02 s intervals) obtained from 19 strong-motion accelerometer stations in Christchurch, identified in Fig. 1. The records from each station for both earthquakes are available from the GeoNet website (GNS Science, 2014). T_d is calculated at each station and then the value at intermediate sites is interpolated by ordinary kriging. Summary statistics of the estimates from the regional EPOLLS and HAZUS models are presented in Table 9. The statistics show that none of the models estimate PGD_{fH} well. The EPOLLS model overestimates the scale of liquefaction, while the HAZUS models each show an underestimation bias. The mean residuals and root-mean-square error (RMSE) are higher for the EPOLLS model, suggesting that the HAZUS models perform slightly better, but this is of little significance since the coefficients of determination of the HAZUS models are all extremely low. A mitigating factor is that the lidar data have a very large error – up to 0.5 m – in the horizontal plane. Taking this into account, over 90% of HAZUS estimates are within the observation error range, although this needs to be interpreted in the context of the mean observed PGD_{fH} being 0.269 m.

Table 9. Summary statistics of horizontal permanent ground deformation (PGDf_H) estimates for Darfield and Christchurch earthquakes from EPOLLS and HAZUS models. Refer to Table 4 for descriptions corresponding to model acronyms.

Score	Horizontal permanent ground deformation, PGDf _H (m)				
	Observed	EPOLLS	HAZ1	HAZ2	HAZ3
Pearson R^2	n/a	0.000	0.022	0.032	0.027
Mean	0.269	0.682	0.141	0.172	0.150
Minimum	0.001	0.149	0.000	0.000	0.000
Lower quartile	0.124	0.418	0.000	0.000	0.000
Median	0.206	0.748	0.084	0.050	0.067
Upper quartile	0.312	0.964	0.184	0.191	0.182
Maximum	3.856	1.989	1.872	3.205	2.443
Residual mean	n/a	0.413	−0.128	−0.096	−0.118
Root-mean-square error	n/a	0.582	0.345	0.438	0.376

Since the HAZUS model underestimates PGDf_H, and PGDf_H cannot be negative, the fact that so many estimates are within this error range is more a reflection of the size of the error relative to the values being observed. Consequently the statistics in Table 9 are more informative and these show that the simplified models all perform poorly.

7 Conclusions

This study compares a range of simplified desktop liquefaction assessment methods that may be suitable for insurance sector where data availability and resources are key constraints. It finds that the liquefaction potential index, when calculated using shear-wave velocity profiles (LPI1), is the best-performing model in terms of its ability to correctly forecast liquefaction occurrence both positively and negatively, although it must be noted that its predictive power is not high. Shear-wave velocity profiles are not always available to practitioners and it is notable therefore that the analysis shows that the next best-performing model is the liquefaction potential index calculated with shear-wave velocity profiles simulated from USGS V_{S30} data (LPI3). Since it is based on USGS data, which are publicly accessible online, this method is particularly attractive to those undertaking rapid and/or regional-scale desktop assessments.

The HAZUS method for estimating liquefaction probabilities performs poorly irrespective of triggering threshold. This is significant since HAZUS methods (not only in respect to liquefaction) are often used as a default model outside of the US when no specific local (or regional) model is available. Models proposed by Zhu et al. (2015) perform reasonably and, since they are also based on publicly accessible data, represent another viable option for desktop assessment. The only issue with these models is that they perform optimally with a low threshold probability of 0.1, which may lead to overestimation of liquefaction when applied to other locations.

As an extension of the liquefaction triggering analysis, this study also uses the observations to relate LPI to liquefaction probability for the two best-performing models. In the case of LPI3, the model performance (as measured by Youden's J statistic) actually improves significantly when employed with a threshold based on corresponding probability rather than based directly on LPI. The final stage of liquefaction assessment is to measure the scale of liquefaction as PGDf. This study only briefly considers this aspect but shows that existing simplified models perform extremely poorly. Existing models show very low correlation with observations and strong estimation bias – underestimation in the case of HAZUS and overestimation in the case of regional EPOLLS. Based on this analysis the estimations from these simplified models are highly uncertain and it is questionable whether they genuinely add any value to loss estimation analysis outside of the regions for which they have been developed.

Data availability. A number of datasets have been used in this study. One dataset containing liquefaction observations was provided directly by Tonkin and Taylor and is not publicly available. Other datasets that have been used are accessible through the Canterbury Geotechnical Database (CGD) or the US Geological Survey (USGS). These datasets include quantitative observations of vertical (CGD, 2012a) and horizontal (CGD, 2012b) permanent ground deformations, qualitative liquefaction and lateral spreading observations (CGD, 2013a), observed peak ground accelerations (CGD, 2013b), borehole site data (CGD, 2013c), tectonic movement measurement (CGD, 2013d), shear-wave velocity estimates (USGS, 2013) and Earth Explorer data (USGS, 2014b) for input into the Zhu et al. (2015) liquefaction assessment models.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors are grateful to Sjoerd van Ballegooy of Tonkin and Taylor for providing data on observed

land damage categories and liquefaction metrics. Funding for this research project has been provided by the UK Engineering and Physical Sciences Research Council and the Willis Research Network, through the Urban Sustainability and Resilience Doctoral Training School at University College London.

Edited by: B. D. Malamud

Reviewed by: J. Douglas and one anonymous referee

References

- Andrus, R. D. and Stokoe, K. H.: Liquefaction resistance of soils from shear-wave velocity, *J. Geotech. Geoenviron.*, 126, 1015–1025, 2000.
- Beaven, J., Motagh, M., Fielding, E. J., Donnelly, N., and Collett, D.: Fault slip models of the 2010–2011 Canterbury, New Zealand, earthquakes from geodetic data and observations of postseismic ground deformation, *New Zeal. J. Geol. Geop.*, 55, 207–221, <https://doi.org/10.1080/00288306.2012.697472>, 2012.
- Bird, J. F. and Bommer, J. J.: Earthquake losses due to ground failure, *Eng. Geol.*, 75, 147–179, 2004.
- Bird, J. F., Bommer, J. J., Crowley, H., and Pinho, R.: Modelling liquefaction-induced building damage in earthquake loss estimation, *Soil Dyn. Earthq. Eng.*, 26, 15–30, 2006.
- Boore, D. M.: Estimating V_{S30} (or NEHRP site classes) from shallow velocity models (depths < 30 m), *B. Seismol. Soc. Am.*, 94, 591–597, 2004.
- Canterbury Geotechnical Database: Vertical Ground Surface Movements, Map Layer CGD0600, available at: <https://canterburygeotechnicaldatabase.projectorbit.com/> (last access: 9 October 2014), 2012a.
- Canterbury Geotechnical Database: Horizontal Ground Surface Movements, Map Layer CGD0700, available at: <https://canterburygeotechnicaldatabase.projectorbit.com/> (last access: 9 October 2014), 2012b.
- Canterbury Geotechnical Database: Liquefaction and Lateral Spreading Observations, Map Layer CGD0300, available at: <https://canterburygeotechnicaldatabase.projectorbit.com/> (last access: 20 September 2014), 2013a.
- Canterbury Geotechnical Database: Conditional PGA for Liquefaction Assessment, Map Layer CGD5110, available at: <https://canterburygeotechnicaldatabase.projectorbit.com/> (last access: 21 February 2014), 2013b.
- Canterbury Geotechnical Database: Geotechnical Investigation Data, Map Layer CGD0010, available at: <https://canterburygeotechnicaldatabase.projectorbit.com/>, last access: 20 September 2014, 2013c.
- Canterbury Geotechnical Database: LiDAR and Digital Elevation Models, Map Layer CGD0500, available at: <https://canterburygeotechnicaldatabase.projectorbit.com/> (last access: 9 October 2014), 2013d.
- Drayton, M. J. and Verdon, C. L.: Consequences of the Canterbury earthquake sequence for insurance loss modelling, 2013 NZSEE Conference, Wellington, New Zealand, 26–28 April 2013.
- ECan: OpenData Portal Canterbury Maps, available at: <http://opendata.canterburymaps.govt.nz/>, last access: 20 September 2014.
- Giovinazzi, S., Wilson, T., Davis, C., Bristow, D., Gallagher, M., Schofield, A., Villemure, M., Eidinger, J., and Tang, A.: Lifelines performance and management following the 22nd February 2011 Christchurch earthquake, New Zealand: highlights of resilience, *Bull. NZ. Soc. Earthq. Eng.*, 44, 402–417, 2011.
- GNS Science: Strong-Motion Data, available at: <http://info.geonet.org.nz/display/appdata/Strong-Motion+Data>, last access: 15 November 2014.
- Goda, K., Atkinson, G. M., Hunter, A. J., Crow, H., and Motazedian, D.: Probabilistic liquefaction hazard analysis for four Canadian cities, *B. Seismol. Soc. Am.*, 101, 190–201, 2011.
- Hosmer, D. W. and Lemeshow, S.: A goodness-of-fit test for multiple logistic regression model, *Commun. Stat.*, A10, 1043–1069, 1980.
- Iwasaki, T., Arakawa, T., and Tokida, K.-I.: Simplified procedures for assessing soil liquefaction during earthquakes, *Soil Dyn. Earthq. Eng.*, 3, 49–58, 1984.
- Juang, C. H., Yang, S. H., and Yuan, H.: Model uncertainty of shear wave velocity-based method for liquefaction potential evaluation, *J. Geotech. Geoenviron.*, 131, 1274–1282, 2005.
- Lees, J. J., Ballagh, R. H., Orense, R. P., and van Ballegooy, S.: CPT-based analysis of liquefaction and re-liquefaction following the Canterbury earthquake sequence, *Soil Dyn. Earthq. Eng.*, 79, 304–314, 2015.
- Matthews, B. W.: Comparison of the predicted and observed secondary structure of T4 phage Iysozyme, *Biochimica et Biophysica Acta – Protein Structure*, 405, 442–451, 1975.
- Maurer, B. W., Green, R. A., Cubrinovski, M., and Bradley, B. A.: Evaluation of the liquefaction potential index for assessing liquefaction potential in Christchurch, New Zealand, *J. Geotech. Geoenviron.*, 140, 04014032, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0001117](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001117), 2014.
- NIBS (National Institute of Building Sciences): HAZUS[®]MH Technical Manual, NIBS, Washington, D.C., 2003.
- Orense, R. P., Pender, M. J., and Wotherspoon, L. M.: Analysis of soil liquefaction during the recent Canterbury (New Zealand) earthquakes, *Geotechnical Engineering Journal of the SEAGS and AGSSEA*, 43, 8–17, 2012.
- Paul, P., Pennell, M. L., and Lemeshow, S.: Standardizing the power of the Hosmer-Lemeshow goodness of fit tests in larger datasets, *Stat. Med.*, 32, 67–80, 2013.
- Pitilakis, K., Crowley, H., and Kaynia, A. (Eds.): SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk, Vol. 27 Geotechnical, Geological and Earthquake Engineering, Springer, the Netherlands, 2014.
- Powers, D. M. W.: Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation, *J. Mach. Learn. Tech.*, 2, 37–63, 2011.
- Quigley, M., Bastin, S., and Bradley, B. A.: Recurrent liquefaction in Christchurch, New Zealand, during the Canterbury earthquake sequence, *Geology*, 41, 419–422, 2013.
- Rauch, A. F. and Martin, J. R.: EPOLLS model for predicting average displacements on lateral spreads, *J. Geotech. Geoenviron.*, 126, 360–371, 2000.
- Seed, H. B. and Idriss, I. M.: Simplified procedure for evaluating soil liquefaction potential, *J. Geotech. Eng. ASCE*, 97, 1249–1273, 1971.

- Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T.: ROCR: visualizing classifier performance in R, *Bioinformatics*, 21, 3940–3941, 2005.
- Smith, T. J. and McKenna, C. M.: A comparison of logistic regression pseudo- R^2 indices, *Multiple Linear Regression Viewpoints*, 39, 17–26, 2013.
- Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. J., and Kattan, M. W.: Assessing the performance of prediction models: a framework for some traditional and novel measures, *Epidemiology*, 21, 128–138, 2010.
- USGS: Global V_{S30} map server, Earthquake Hazards Program, available at: <http://earthquake.usgs.gov/hazards/apps/vs30/>, last access: 10 November 2013.
- USGS: ShakeMaps, Earthquake Hazards Program, available at: <http://earthquake.usgs.gov/earthquakes/shakemap/>, last access: 13 March 2014a.
- USGS: Earth Explorer, available at: <http://earthexplorer.usgs.gov/>, last access: 13 October 2014b.
- 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.: Assessment of liquefaction-induced land damage for residential Christchurch, *Earthq. Spectra*, 30, 31–55, 2014.
- Wood, C., Cox, B. R., Wotherspoon, L., and Green, R. A.: Dynamic site characterization of Christchurch strong motion stations, *B. NZ Soc. Earthq. Eng.*, 44, 195–204, 2011.
- Wotherspoon, L., Orense, R. P., Green, R., Bradley, B. A., Cox, B., and Wood, C.: Analysis of liquefaction characteristics at Christchurch strong motion stations, in: *Soil Liquefaction During Recent Large-Scale Earthquakes*, edited by: Orense, R. P., Towhata, I., and Chouw, N., Taylor and Francis, London, UK, 33–43, 2014.
- Zhou, Y.-G. and Chen, Y.-M.: Laboratory investigation on assessing liquefaction resistance of sandy soils by shear wave velocity, *J. Geotech. Geoenviron.*, 133, 959–972, 2007.
- Zhu, J., Daley, D., Baise, L. G., Thompson, E. M., Wald, D. J., and Knudsen, K. L.: A geospatial liquefaction model for rapid response and loss estimation, *Earthq. Spectra*, 31, 1813–1837, 2015.