nzsee
NEW ZEALAND SOCIETY FOR EARTHQUAKE ENGINEERING

Evaluation of a geospatial liquefaction model using land damage data from the 2016 Kaikōura earthquake

A. Lin & L. Wotherspoon
University of Auckland, Auckland.

J. Motha
University of Canterbury, Christchurch.

ABSTRACT

Earthquake induced liquefaction and lateral spreading can cause major damage to buildings and infrastructure networks. Common procedures to identify exposed areas such as cone penetration testing require extensive resources and are not suitable for national scale hazard assessments. Statistical models based on geospatial variables offer an alternative approach, allowing for a time and cost efficient estimation of liquefaction and the identification of areas where more detailed assessments would be beneficial. The research focuses on a globally applicable geospatial model developed on publicly accessible variables such as precipitation, distance to closest water body or water table depth. For New Zealand, the model performed well for the 2010 Darfield and 2011 Christchurch event. However, both events were used in the calibration process, suggesting that the performance might be overstated. In order to better understand its prediction potential, the research evaluates the geospatial model using liquefaction observational data from the 2016 Kaikōura earthquake. Results show that the model performs well for the observation cluster in Kaikōura. However, discrepancies can be found for the cluster in Blenheim, which is partly caused by the low resolution of the global model input data. Despite the contradicting outcome, the findings of this research help to identify areas of improvement and illustrate the limitations of the geospatial approach for regional and local application across New Zealand.

1 INTRODUCTION

During the 2010-2011 Canterbury Earthquake Sequence, liquefaction and lateral spreading caused substantial damage to the built environment affecting 60,000 residential properties and major parts of the urban infrastructure system including bridges, stopbanks and water pipes (Cubrinovski et al., 2012). The

events demonstrate the range and extent of liquefaction impacts in New Zealand and emphasise the need for a better understanding in order to make more appropriate design and planning decisions.

Liquefaction hazard maps can help identifying exposed structures. Simplified procedures such as standard penetration testing combined with geostatistical methods and/or geological data are commonly used to generate liquefaction hazard maps on a local or regional level (Holzer et al., 2006; Lenz & Baise, 2007; Maurer et al., 2014). However, they require extensive resources. Most recently, probabilistic models based on geospatial data are considered as an alternative approach (Geyin et al., 2020). Zhu et al. (2015) developed and updated (Zhu et al., 2017) a geospatial liquefaction model by correlating case history data with globally accessible soil property data such as precipitation or water table depth. The model was evaluated across different earthquake scenarios showing satisfying results, especially considering their benefits regarding time and costs compared to in-situ based approaches (Geyin et al., 2020; Maurer, 2017; Rashidian & Baise, 2020). The model also performed well for the 2010 Darfield and 2011 Christchurch earthquakes. However, both events were used in the calibration process, suggesting that the performance might be overstated. In order to better understand its prediction potential, the research compares the observational data from the 2016 M_w7.8 Kaikōura earthquake with the liquefaction estimates calculated by the geospatial model. Based on the findings, general limitations of the approach as well as uncertainties regarding its applicability across New Zealand are discussed.

2 GEOSPATIAL LIQUEFACTION MODEL

Using logistic regression, Zhu et al. (2017) correlated observational data from 27 earthquakes around the globe with geospatial data on the soil properties that are related to liquefaction manifestation. They found that the most promising results were achieved with a combination of peak ground velocity (PGV) in cm/s, shear wave velocity in the upper 30m (Vs_{30}) in m/s, annual precipitation (PRECIP) in mm, distance to the closest water body (DW) in km and water table depth (WTD) in meters below ground level (m.b.g.l.). The liquefaction probability (P) is calculated by the equation

$$P = \frac{1}{1+ e^{-X}} \quad (1)$$

where X equals a function of the explanatory variables

$$X = 8.801 + 0.334 \ln(PGV) - 1.918 \ln(Vs_{30}) + 5.408\ 10^{-4}\ PRECIP - 0.2054\ DW - 0.0333\ WTD \quad (2)$$

In addition to the liquefaction probability, Zhu et al. (2017) introduced the liquefaction spatial extent (LSE) to describe the areal coverage (in percent) of surface manifestation. LSE is estimated by

$$LSE = \frac{49.15}{(1 + 42.4e^{-9.165\,P})^2} \quad (3)$$

where P is the liquefaction probability calculated by Equation 1. While the liquefaction probability allows for the assessment of past and future events on the basis of a binary classification analysis (liquefaction manifests versus liquefaction does not manifest), the liquefaction spatial extent can be used to illustrate the severity of predicted liquefaction manifestation and to develop liquefaction hazard maps for disaster risk management.

3 DATA AND METHODOLOGY

Following the approach of Zhu et al. (2017), all input variables are based on publicly available sources: PGV values from the 2016 Kaikōura earthquake are retrieved from the USGS ground motion and shaking intensity tool Shakemap (USGS, 2016); Vs_{30} is estimated using the approach by Wald & Allen (2007) that correlates topographic slope with Vs_{30} measurements; PRECIP is retrieved from the weather and climate database

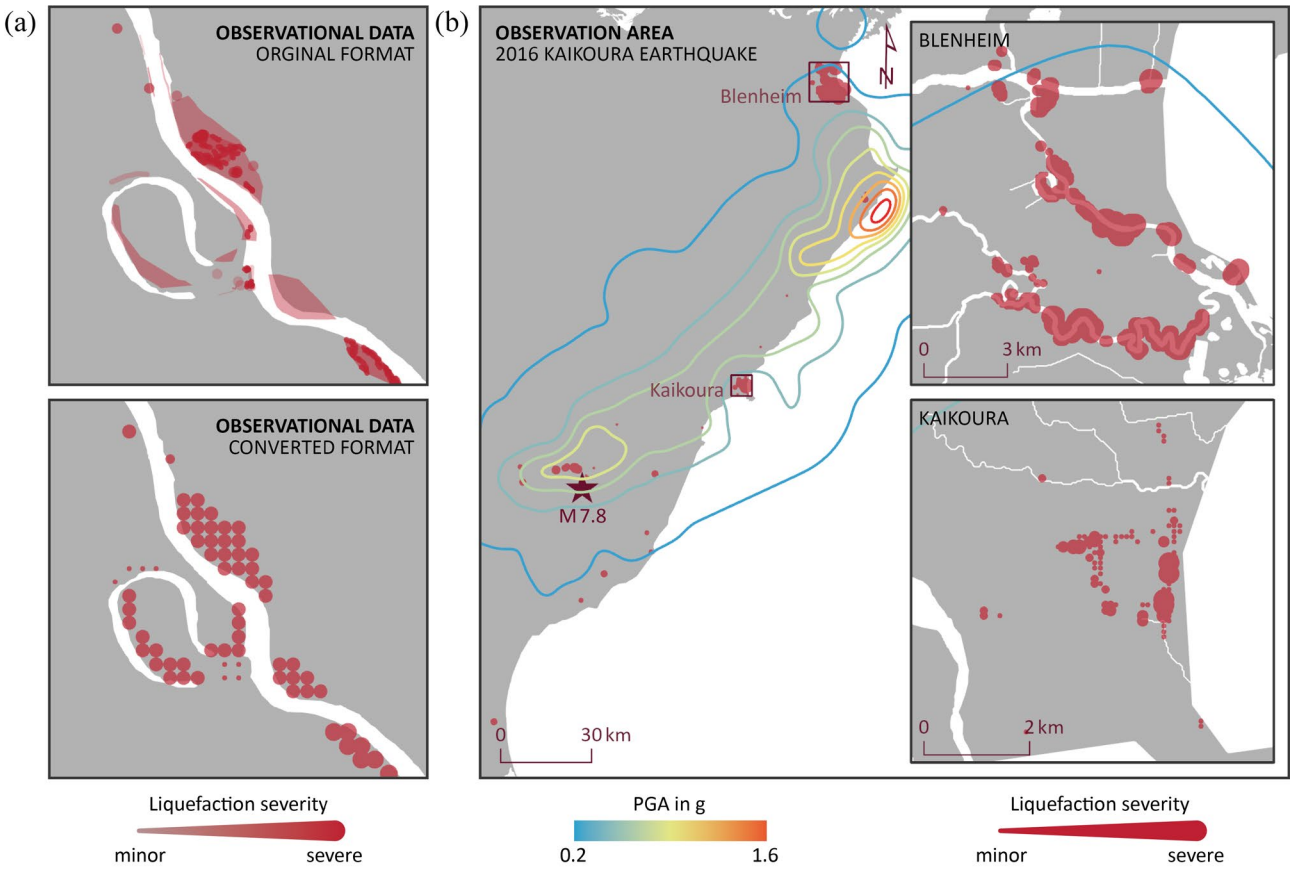

Figure 1: (a) Illustration of the data conversion process and (b) observed liquefaction manifestation including PGA contours across the observation area.

WorldClim (Hijmans et al., 2005); and WTD is based on the groundwater model by Fan et al. (2013) using climate, terrain and sea level measurements. The explanatory variable DW is defined as the minimum of distance to the closest river (HydroSHEDS, 2006) or closest coastline (NASA, 2009); both datasets are generated from remote sensing data. Each explanatory variable is converted to a raster and resampled using bilinear interpolation to achieve a consistent resolution of 100m. While the Zhu et al. (2017) approach uses a resolution of 30 arcsec (approximately 1km), Rashidian & Baise (2020) suggest that 3 arcsec (approximately 100m) is more appropriate for the comparison with observational data, especially in fast changing terrains such as coastal areas.

3.1 Observed liquefaction manifestation

The M_W 7.8 Kaikōura earthquake occurred on 14 November 2016 involving multiple faults in the upper South Island of New Zealand. The rupture originated approx. 15km north east of Culverden and propagated to the north-east along the east coast of Canterbury and Marlborough. The Kaikōura earthquake triggered tens of thousands of landslides and caused wide ranging surface displacements leading to severe damage of buildings and infrastructure networks (Kaiser et al., 2017). Despite high shaking intensity, the extent of liquefaction and lateral spreading was relatively limited (Bastin et al., 2018). The QuakeCoRE (2017) historic liquefaction database contains the observational data collected during post-event reconnaissance. The entries consist of different data formats including points (23%), lines (e.g. roads, 60%) and polygons (e.g. areas, 17%), and describe both the land damage type (e.g. liquefaction without lateral spreading) and the liquefaction severity (e.g. minor).

As the application of the geospatial liquefaction model requires a consistent data format, the lines and polygons are converted to points and point grids, respectively (Fig. 1a). The point spacing along lines and the

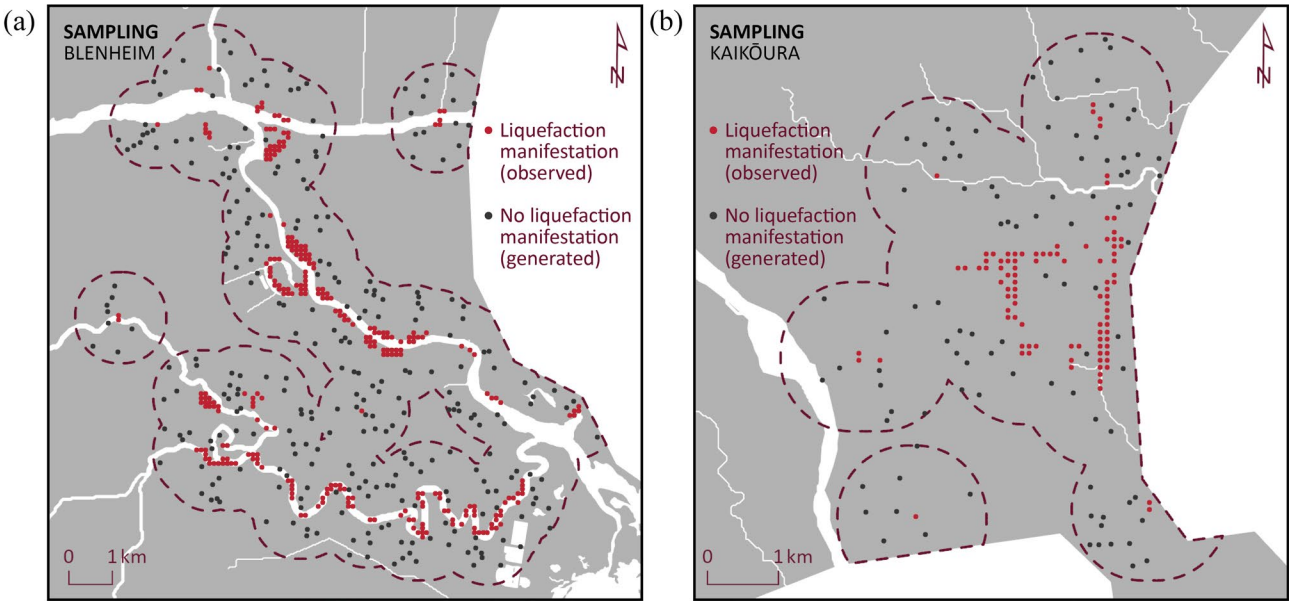

Figure 2: Illustration of the sampling method leading to the points with liquefaction manifestation (observed) and no liquefaction (generated) for the (a) Blenheim and (b) Kaikōura clusters.

grid size across the polygons are set to 100m in order to be in alignment with the resolution of the input variables. Overlaying datasets that could result in multiple points in one raster cell and potentially lead to information redundancy are treated as a single layer. In cases where overlaying datasets contained different information on the observed liquefaction severity, the higher severity is chosen. All points are merged to one sample. The outcome of the conversion process is shown in Figure 1b. Although the observation area covers the coastal region from the north of Marlborough to the centre of Canterbury as well as the area around the epicentre, the assessment focuses on the areas with the highest point density (clusters), which can be found in Blenheim and Kaikōura. 83% of the liquefaction manifestation was observed in these clusters. Despite a lower ground shaking intensity, the observational data in Blenheim shows a higher manifestation severity, with most of the points located along the Wairau River (north) and the Ōpaoa River (south). The distribution pattern in the Kaikōura cluster does not directly relate to the surrounding rivers; however, many observation points are located close to the eastern coastline.

The geospatial liquefaction model by Zhu et al. (2017) is applied to the entire observation area. Model performance is assessed by comparing the LSE map with the observed liquefaction manifestation across the Blenheim and Kaikōura clusters.

3.2 Generated cases (no liquefaction manifestation)

In addition to the LSE comparison, prediction performance is assessed using the receiver operating classification (ROC) area under the curve (AUC), a common indicator to evaluate binary classification problems (Fawcett, 2006). Since the observational data only consists of positive cases (liquefaction manifested), the negative sample (liquefaction did not manifest) requires the generation of additional points. Zhu et al. (2017) used random sampling, which involved the creation of buffers with a radius of 15km around the positive cases. In this area, random points were generated to complete the locations where no liquefaction manifested. A sensitivity analysis proved that the chosen radius did not affect the outcome of the ROC evaluation. Due to the differences in spatial distribution of the observational data and the resolution of the explanatory variables, the buffer dimensions proposed by Zhu et al. (2017) seem unsuitable for the 2016 Kaikōura data. Having a relative high density within the observation points, it is expected that the positive cases will result in similar probability values (positive spatial autocorrelation). A large buffer area would lead to a more spread distribution of the negative cases, hence, a higher variation across the probability

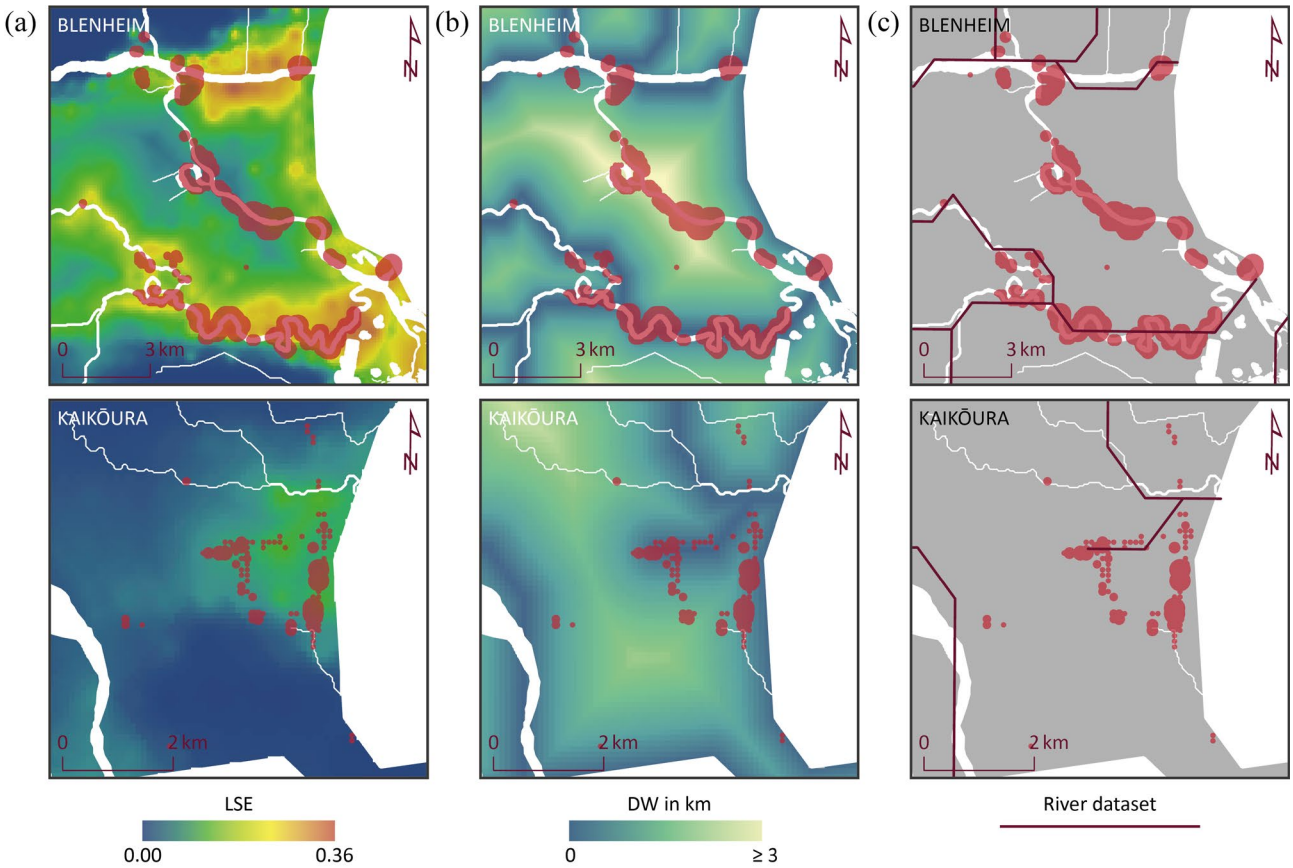

Figure 3: (a) Liquefaction spatial extent, (b) distance to closest water and (c) river dataset across the Blenheim and Kaikōura clusters. The red areas indicate the observed liquefaction severity.

results, which likely affects the ROC analysis. On the other hand, a small buffer area would not sufficiently cover the study area limiting the validity of the overall assessment. As a compromise, a buffer radius of 1km is chosen. Negative cases are generated by creating random points within this buffer (Fig. 2). For class balance, the number of negative cases is equal to the number of positive cases. This leads to a total sample size of 570 points for the Blenheim cluster and 190 points for the Kaikōura cluster. Similar to the (converted) observational data, an intersample distance of 100m or above is chosen.

For both the positive and negative samples, liquefaction probability is estimated in each point by extracting the values from the explanatory variable rasters and applying Equations 1 to 3. Plotting the true positive rate (TPR) (percentage of positives predicted correctly) against the false positive rate (FPR) (percentage of negatives predicted incorrectly) for different probability thresholds results in the ROC curve. The area under the curve (AUC) indicates how well the model distinguishes between positive and negative cases. Usually, the ROC AUC ranges from 0.5 (random prediction) to 1.0 (perfect prediction) (Fawcett, 2006).

4 RESULTS

Figure 3a presents the LSE maps for the Blenheim and Kaikōura areas. In Blenheim, higher values can be found along the coast and the main streams. The observed liquefaction manifestation in the south (Ōpaoa River) seems to be well captured by the model, while the low LSE around the remaining observation points (e.g. Wairau River) suggests rather poor prediction performance. The LSE map of the Kaikōura area shows lower values, which reflects the lower severity of observed liquefaction manifestation. With a few exceptions in the outer areas, the observation points appear to be well represented by the LSE map.

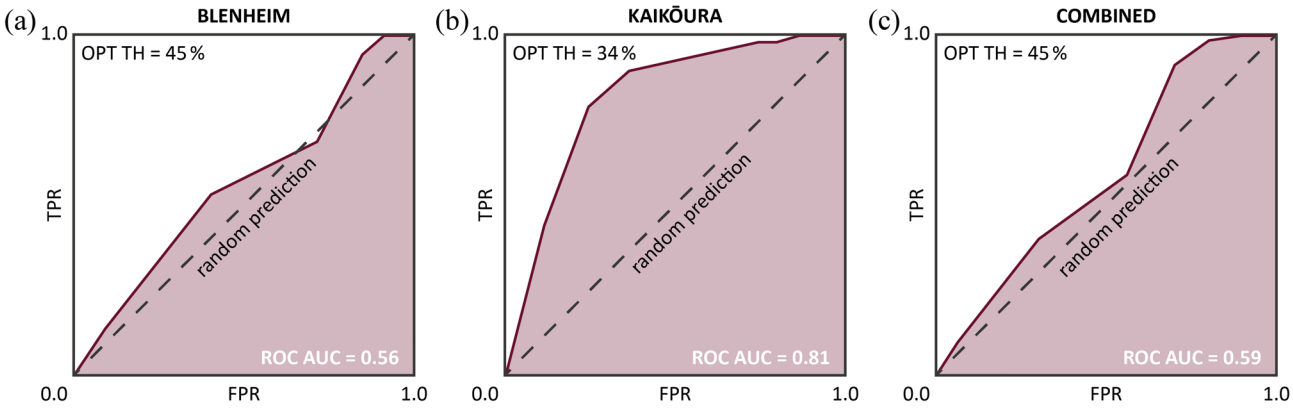

Figure 4: Results of the receiver operating classification (ROC) analysis for the (a) Blenheim cluster, (b) Kaikōura cluster and (c) both clusters combined.

One reason for the low LSE values in the central area of the Blenheim cluster is the explanatory variable DW. As presented in Figure 3b, DW is relatively high around the Wairau River leading to the false perception of a larger distance to the nearest water body. The discrepancy is due to the low resolution of the input variable: The dataset for the location of the rivers is based on HydroSHEDS (2006), a platform aiming to provide consistent hydrographic information on a global scale. However, its use of radar based information can lead to errors (Lehner et al., 2006, 2008). The purple lines in Figure 3c present the HydroSHEDS (2006) river dataset showing an oversimplification of the stream flow. Furthermore, some rivers such as the Wairau River are not covered by the dataset at all. For these areas, it is not surprising that the geospatial model underestimates liquefaction probability and spatial extent. The discrepancies in the LSE map demonstrate the limitations of the global input variables and their use for regional or local liquefaction assessments.

The findings of the LSE assessment are also reflected by the ROC results, which are presented in Figure 4. The Blenheim cluster leads to a very low AUC value (0.56) close to the random threshold, indicating that the geospatial model is not able to separate the positive cases from the negative cases. In contrast, the reasonably high AUC of the Kaikōura cluster (0.81) suggests that the model performs well this area. Comparing the findings of the 2016 Kaikōura event with the results of the ROC analysis conducted by Zhu et al. (2017) for the 2010 Darfield (AUC = 0.69) and 2011 Christchurch (AUC = 0.77) earthquake, shows inconsistency within the ROC performance values, which do not allow for a reliable conclusion on whether the geospatial model is able to predict liquefaction manifestation.

Another conflict arises when comparing the optimal thresholds (OPT TH) of the two areas. The optimal threshold is the probability value that best distinguishes between positive and negative cases. It can be used as a reference for the assessment of future earthquake scenarios. The optimal threshold of the Blenheim cluster (45%) is much higher compared to the Kaikōura cluster (34%), preventing the definition of a suitable reference value. The obstacles that result from inconsistent optimal thresholds become clear when evaluating the ROC performance for both clusters combined, which leads to an AUC (0.59) and optimal threshold (45%) similar to the Blenheim cluster. While the tendency towards the Blenheim outcome is not surprising as its sample size is three times larger than the Kaikōura cluster, the combined analysis demonstrates that the high ROC AUC of a single sample does not necessarily reflect the model's overall performance.

5 DISCUSSION AND CONCLUSION

Based on the 2016 Kaikōura earthquake, liquefaction probability and spatial extent were calculated using the geospatial model developed by Zhu et al. (2017). The evaluation of both the LSE maps and the ROC AUC values indicates that the model performs well across the Kaikōura area but poorly for the Blenheim region.

The contradicting results as well as the inconsistent optimal thresholds do not allow for a clear conclusion regarding the model's potential to predict liquefaction manifestation. The discrepancy in the performance outcome is partly caused by the low resolution of the river dataset used to estimate DW. As highlighted by the observational data in the Blenheim cluster, the spatial accuracy of the river data is particularly relevant as liquefaction manifestation often occurs close to rivers. While the explanatory variables proposed by Zhu et al. (2017) are appropriate for large scale hazard assessments as they provide global coverage, they might not be suitable for regional or local application. Further research needs to evaluate the quality of the input variables for New Zealand terrain and potentially consider the use of national datasets, e.g. the New Zealand River Environment Classification (MfE, 2010) as an alternative source to estimate DW.

Another limitation that affects the validity of the findings arises from the sampling method to generate the negative cases. As mentioned previously, a relative small buffer radius was chosen to reduce the effects of the different point densities on the outcome of the ROC analysis. However, as a result of the limited assessment area, the evaluation does not represent the wider area of the Blenheim and Kaikōura clusters; for instance, the buffer of the Blenheim cluster does not cover the misleading LSE values in the north, consequently, the ROC analysis does not account for the overprediction in this area. This suggests that further research should look into a more robust approach to measure the model's prediction potential. Although the comparison of the LSE maps with the observational data allows for a qualitative assessment of the model performance, a quantitative approach is aspired in order to provide a probability value as a reference for the evaluation of future earthquake scenarios.

Although the evaluation does not provide a clear picture on the model's potential to estimate liquefaction manifestation, the findings of this research illustrated the general challenges in measuring the model performance (e.g. sampling approach) and identified areas for improvement (e.g. quality of data input). Further research with observational data from other earthquake scenarios and the consideration of more appropriate data sources for the explanatory variables could lead to a more robust evaluation of the geospatial liquefaction model and its applicability for regional and local hazard assessments across New Zealand.

REFERENCES

Bastin, S.H., Stringer, M., Green, R., Wotherspoon, L., van Ballegooy, S., Cox, B.R. & Osuchowski, A. 2018. Geomorphological Controls on the Distribution of Liquefaction in Blenheim, New Zealand, during the 2016 Mw7.8 Kaikōura Earthquake. Geotechnical Earthquake Engineering and Soil Dynamics V, 264-272. https://doi.org/10.1061/9780784481455.026.

Cubrinovski, M., Robinson, K., Taylor, M., Hughes, M. & Orense, R. 2012. Lateral spreading and its impacts in urban areas in the 2010–2011 Christchurch earthquakes. *New Zealand Journal of Geology and Geophysics*, 55(3), 255-269. https://doi.org/10.1080/00288306.2012.699895.

Fan, Y., Li, H. & Miguez-Macho, G. 2013. Global patterns of groundwater table depth. *Science*, 339(6122), 940-943. https://doi.org/10.1126/science.1229881.

Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010.

Geyin, M., Baird, A.J. & Maurer, B.W. 2020. Field assessment of liquefaction prediction models based on geotechnical versus geospatial data, with lessons for each. *Earthquake Spectra*, 36(3), 1386-1411. https://doi.org/10.1177%2F8755293019899951.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965-1978. https://doi.org/10.1002/joc.1276.

Holzer, T.L., Bennett, M.J., Noce, T.E., Padovani, A.C. & Tinsley Iii, J.C. 2006. Liquefaction hazard mapping with LPI in the Greater Oakland, California, area. *Earthquake Spectra*, 22(3), 693-708. https://doi.org/10.1193/1.2218591.

Hydrological data and maps based on Shuttle Elevation Derivatives at multiple scales (HydroSHEDS). 2006. River

network (stream lines) at 30s resolution [GIS Shapefile]. https://www.hydrosheds.org/page/hydrorivers (2019, August 27).

Kaiser, A., Balfour, N., Fry, B., Holden, C., Litchfield, N., Gerstenberger, M., . . . Gledhill, K. 2017. The 2016 Kaikōura, New Zealand, Earthquake: Preliminary Seismological Report. *Seismological Research Letters*, 88(3), 727-739. https://doi.org/10.1785/0220170018.

Lehner, B., Verdin, K. & Jarvis, A. 2006. HydroSHEDS technical documentation. World Wildlife Fund. https://www.hydrosheds.org/images/inpages/HydroSHEDS_TechDoc_v1_2.pdf.

Lehner, B., Verdin, K. & Jarvis, A. 2008. New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10), 93-94. https://doi.org/10.1029/2008eo100001.

Lenz, J.A. & 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. https://doi.org/10.1016/j.soildyn.2006.11.005.

Maurer, B.W. 2017. Field-testing liquefaction models based on geospatial vs. geotechnical data. 6th International Young Geotechnical Engineers' Conference (iYGEC6) (2017, September 16-17), Seoul, Korea.

Maurer, B.W., Green, R.A., Cubrinovski, M. & 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)04014031-04014011. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001117.

Ministry for the Environment (MfE). 2010. River Environment Classification New Zealand (2010) [GIS Shapefile]. https://data.mfe.govt.nz/layer/51845-river-environment-classification-new-zealand-2010/ (2019, March 4).

National Aeronautics and Space Administration (NASA). 2009. Distance to the Nearest Coast [csv file]. https://oceancolor.gsfc.nasa.gov/docs/distfromcoast/

QuakeCoRE. 2017. Historic Earthquake Events [GIS Shapefile]. https://projectorbit.maps.arcgis.com/apps/webappviewer/index.html?id=140265d6f8754f28851c92dee5491c9a (2020, 27 October).

Rashidian, V. & Baise, L.G. 2020. Regional efficacy of a global geospatial liquefaction model. *Engineering Geology*, 272(2020), 105644. https://doi.org/10.1016/j.enggeo.2020.105644.

U.S. Geological Survey (USGS). 2016. M 7.8 - 54km NNE of Amberley, New Zealand, 2016-11-13 11:02:56 (UTC) [kmz file]. https://earthquake.usgs.gov/earthquakes/eventpage/us1000778i/shakemap/pgv (2019, 21 December).

Wald, D.J. & Allen, T.I. 2007. Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, 97(5), 1379-1395. https://doi.org/10.1785/0120060267.

Zhu, J., Baise, L.G. & Thompson, E.M. 2017. An updated geospatial liquefaction model for global application. *Bulletin of the Seismological Society of America*, 107(3), 1365-1385. https://doi.org/10.1785/0120160198.

Zhu, J., Daley, D., Baise, L.G., Thompson, E.M., Wald, D.J. & Knudsen, K.L. 2015. A geospatial liquefaction model for rapid response and loss Estimation. *Earthquake Spectra*, 31(3), 1813-1837. https://doi.org/10.1193/121912eqs353m.