National Liquefaction Model: framework, opportunities and challenges

Maxim D. L. Millen,
Tonkin + Taylor, Nelson, New Zealand & University of Canterbury, Christchurch, New Zealand.

J. Russell
Tonkin + Taylor, Auckland, New Zealand

M. Jacka
Tonkin + Taylor, Christchurch, New Zealand.

ABSTRACT

The substantial liquefaction-induced damage from the Canterbury Earthquake Sequence was significantly underestimated by insurance loss models (Swiss Re, 2014), increasing the cost uncertainty for future events. To address this, the Natural Hazards Commission (NHC) Toka Tū Ake has commissioned Tonkin + Taylor Ltd to develop the National Liquefaction Model (NLM) to support loss modelling in New Zealand. The NLM is a first of its kind model, in that it provides predictions of liquefaction response in terms of common engineering parameters (e.g. LSN) for any scenario peak ground acceleration, earthquake magnitude, and groundwater depth. Several technical challenges shaped the development of the framework including model update requirements, end-use requirements and data scarcity. The model framework, as well as several challenges and adopted solutions are covered in this paper.

1 INTRODUCTION

Tonkin + Taylor Ltd is in the process of developing a National Liquefaction Model (NLM) for the Natural Hazards Commission (NHC) Toka Tū Ake. The model is intended to advance NHC's earthquake loss modelling programme, as well as support informed land use planning and research through a nationally-consistent evaluation of liquefaction hazard.

The model has been deliberately designed to be modular in both the source code and outputs, readily updatable with new data and/or additional functionality, and data-driven, in that it is a fully prescribed process to produce the outputs from the input data (avoiding manual interventions). This paper presents the framework, technical challenges that shaped the framework of the model, and several key opportunities. A full technical report will be made available on release of the model files.

2 KEY STEPS IN MODEL DEVELOPMENT AND APPLICATION

Figure 1 provides the high-level steps for development and application of the NLM to produce the key model outputs. The model development steps and run time steps (i.e. steps to produce outputs using the model) are covered briefly below.

Development steps:

1. Development of a National Flatland Model, which identifies flatland across the country and defines the extent of the NLM.
2. Development of a National Geomorphology Model, which identifies geomorphology across the country. Defined using the terms established by Youd and Perkins (1978) for liquefaction evaluation (i.e. geomorphology and epoch), as well as an additional primary rock mapping classification that was captured in this model (obtained primarily from QMAP v4 from GNS (Heron, 2023)).
3. Liquefaction triggering analysis of over 50,000 ground investigations (cone penetration tests (CPT), boreholes and standard penetration tests (SPT) from the New Zealand Geotechnical Database (NZGD 2023)) for a range of Peak Ground Acceleration (PGA), Magnitude (Mw) and groundwater depths, resulting in over 10 million analyses.
4. Development of a geospatial model, which provides liquefaction vulnerability estimates using relationships between the aggregated national data set of ground investigations and geospatial features.
5. Development of a Liquefaction Vulnerability Model (LVM), which provides estimates of liquefaction demand measures (LDMs) (e.g. of liquefaction severity number (LSN) (Van Ballegooy et al., 2014)) for different PGA, Mw and groundwater levels across the extent of the National Flatland Model. The model combines estimates from both the geospatial model and local ground investigations, and produces output in terms of a lognormal distribution of LSN for each PGA, Mw and groundwater combination.
6. Development of a National Groundwater Model, which provides estimates of the median free groundwater depth, an estimate of uncertainty in that depth, and a seasonal variability trend. Additionally, a model that reflects 1 m of sea level rise was developed.

Run time steps:

1. Selection of the seismicity map of PGA and Mw (either a return period output or an event scenario).
2. Estimation of the likely distribution of an LDM at each location of interest, by combining the PGA map, event Mw (or Mw map for return period calculations), groundwater map, and the LVM.
3. Estimation of land damage or building damage either as a probability or as a realisation of a damage state (i.e. randomly sampled from a distribution) through use of LSN in fragility functions.

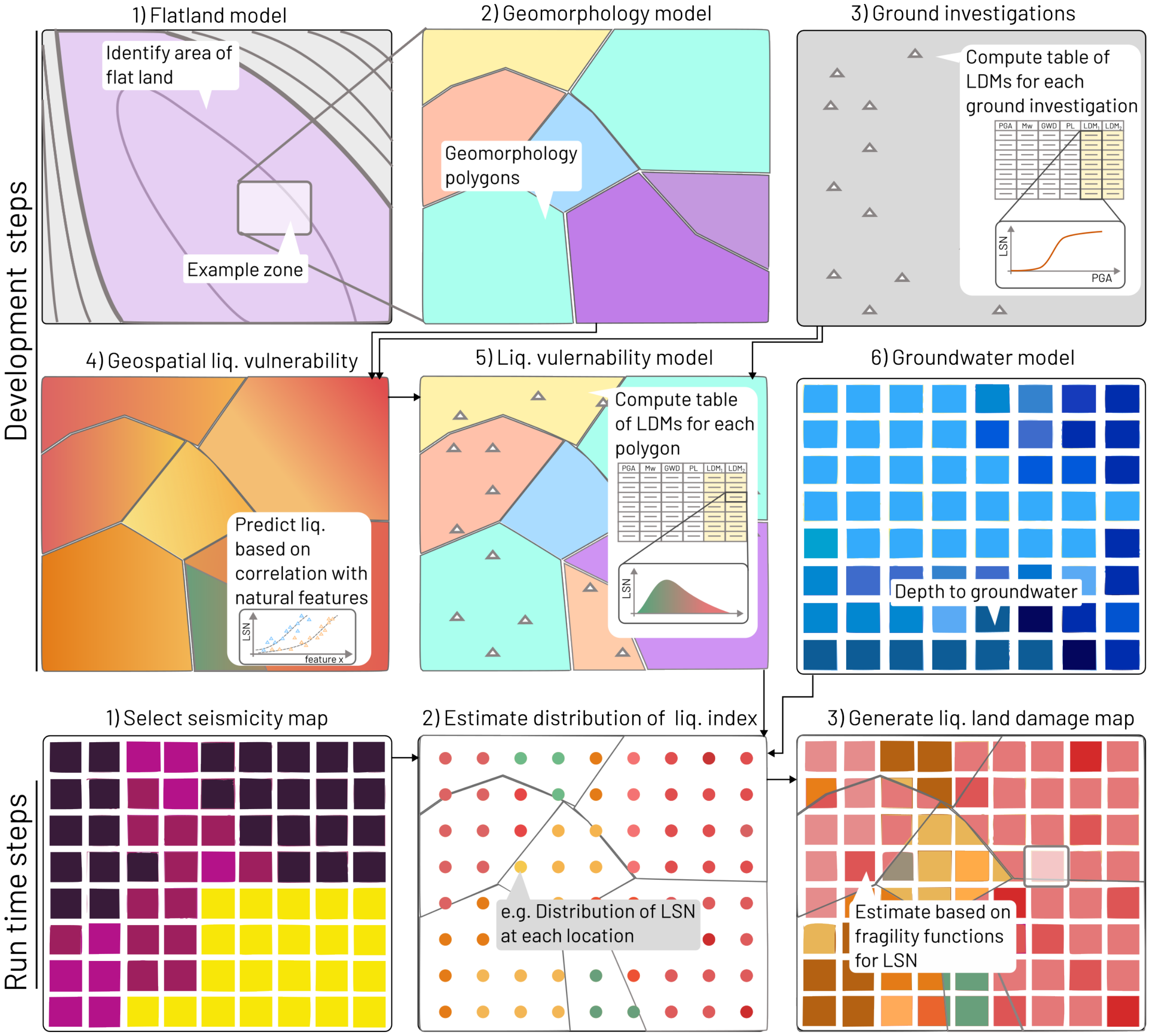

Figure 1 High-level steps to develop and run the National Liquefaction Model

3 KEY CHALLENGES

Some of the most influential technical challenges that shaped the NLM framework and how they were addressed are provided in Table 1.

Table 1 Key challenges and adopted approaches for the development of the National Liquefaction Model

Challenge	Adopted approach
How can the model support the running of different seismic event scenarios and different groundwater estimates (including more refined estimates developed at regional scale)?	A modular framework was used that decoupled the groundwater and seismicity layers from liquefaction estimates. This was achieved by precomputing LDMs for all combinations of a predefined set of PGA, Mw and groundwater depths. Then LDMs at each location could

	be produced by looking up the LDM that corresponds to the scenario PGA, Mw and groundwater depth of interest.
How can the model be easily updated to allow new data, new liquefaction vulnerability methods, and other data sources in future releases	A modular framework for each of the key steps was adopted, which meant that alternative approaches can be considered at each step (e.g. producing alternate LDM's such as LPI (Iwasaki et al., 1981)). The decoupling occurs between each key step by producing output files which are read in by subsequent steps, as well as a linear dependency of logic in the code base (i.e. no circular references).
How can rational, unbiased estimates of liquefaction-induced ground damage be generated in the absence of local ground investigations? *It was necessary for the model to produce unbiased estimates due to its application for loss modelling, noting that the model can produce different percentile outputs for other purposes (i.e. 15th percentile likelihood of LSN).*	Many existing geospatial models (e.g. Zhu et al. (2017)), only provide estimates on likelihood of liquefaction, not a probability of land damage or an LDM. They are also not directly compatible with LDMs computed from ground investigations via simplified triggering calculations (e.g. Boulanger & Idriss (2016)). For the NLM the geospatial approach by Geyin & Maurer (2020) was adapted, which extrapolates LDMs computed from ground investigations using geospatial features. Additional challenges with this approach are covered in Section 4.1.
How can the model be stored such that it maintains an appropriate level of accuracy both spatially and in terms of the input seismicity and groundwater depth, but could be stored and loaded with conventional geospatial software?	The three primary options for storage of geospatial predictions are raster, vector shape (i.e. polygons), and analytical (i.e. equations). A vector shape approach was adopted. The storage of response parameters in terms of input parameters can either be done analytically or as a lookup table. Both options were used, with the output of the geospatial model using equations, while the LVM output as a lookup table. Additional details and considerations of these decisions are covered in Section 4.2
How can the outputs be communicated in language that can be understood beyond the engineering profession?	The outputs of the model were produced in terms of LSN and converted to land damage probability and building damage probability using fragility functions.

3.1 Development of liquefaction geospatial model

The estimate of liquefaction using a geospatial model allowed predictions across the extent of the National Flatland Model, however, there were several challenges with its development that were resolved with the adopted solution. The initial solution was approach of Geyin & Maurer (2020), where LDMs are computed from all CPTs for the estimated groundwater level, and a range of Mw and PGA levels. For each ground investigation, a series of geospatial features were sampled (e.g. distance to river, annual precipitation). This produces a database of LDM outputs paired with several geospatial features as well as Mw and PGA. A machine learning (ML) model was then trained on the dataset such that the ML model can then predict an LDM from only the geospatial features and Mw and PGA. Geyin & Maurer (2020) note that if a particular ground type is poorly sampled by the CPT database then predictions would be poorly constrained.

This approach was then modified to produce predictions for a range of groundwater values instead of the expected groundwater. Using a range of groundwaters at each CPT meant that LDM predictions could be produced for any groundwater depth. The CPT dataset (NZGD, 2023) was then used to conduct over 10 million simplified triggering analyses, and was combined with several distance to water features, shear wave velocity in the top 30 m (Foster et al., 2019), and the National Geomorphology Model as well as additional features extracted from QMAP (Heron, 2023). This approach yielded several limitations:

1. The high correlation of geospatial features with the National Geomorphology Model meant that unexpected trends were produced from the ML models (e.g. LDMs decreasing with distance from water)
2. The use of only CPT data meant that only data from penetrable ground was in the database, this meant that the ML model extrapolated penetrable ground values into unsampled non-penetrable ground (this problem could be confirmed by looking at borehole records).

To overcome the first issue, predictions were produced using only the categories of the National Geomorphology Model (namely geomorphology class, epoch, and main rock). The predictions were provided as lognormal distributions of LSN and analytical equations were fitted for each category to quantify the trends for changes in groundwater and magnitude adjusted PGA (from Idriss & Boulanger (2008)).

The second issue was addressed by adding predictions using SPT and borehole records to supplement the CPT dataset. This allowed rock sites as well as dense gravel sites to be represented in the training dataset.

A significant advantage of the adopted approach is that is improves with improved geomorphology mapping and as the ground investigation database increases.

3.2 Challenges with storage and use of the model

Given the need to allow the running of different seismic event scenarios and different groundwater estimates, the storage of the model needed to support estimates at every location for the full range of input parameters (resulting in the order of 1500 outputs per location). To achieve this with a 100 m grid raster would require storage of ~55 billion values, which would result in significant computation challenges. An analytical approach is convenient and was adopted for the geospatial model, however, the analytical approach did not allow the inclusion of local data at the site of interest. Therefore, vector shapes (polygons) were adopted which identified areas that were expected to have a similar liquefaction response and a similar level of uncertainty in that response. The response for each polygon was assigned two lognormal distributions of LSN, one representing only the likely spatial distribution of LSN within the polygon, the second representing the spatial distribution and the uncertainty in that distribution. These polygons are referred to as similar soil polygons (SSPs) and acted as a form of compression of the data, particularly in areas with low resolution (i.e. limited geomorphology mapping and no local ground investigations). The extent of each polygon was governed by minimising information loss (i.e. the larger the area the more averaging that may happen across ground investigations, which inherently loses the spatial trends that may be observed in the ground investigations). To achieve a balance of compression versus information loss, three different tiers of SSPs were developed, which reflect an increasing understanding of the ground conditions:

- Tier 1 polygons are based on the geomorphology polygon boundaries and have limited or no ground investigation data.

- Tier 2 polygons split large Tier 1 polygons based on ground investigations if there is a sufficient difference in liquefaction response of the ground investigations.
- Tier 3 polygons identify closely spaced clusters of ground investigations that have similar liquefaction response.

Figure 2 shows a conceptual example of a polygon where the true response for a cross-section is shown and examined for the following situations:

1. In Case 0 there are no ground investigations, and the distribution of expected response can only be constrained by the geospatial model. The polygon is not split, and it is considered Tier 1.
2. In Case 1 there are a few ground investigations, however, the readings do not show a detectable spatial trend or significant benefit of splitting. The polygon is not split, and it is considered Tier 1.
3. In Case 2 there are a few ground investigations, and the readings show a discernible spatial trend. The polygon is split, and it is considered Tier 2.
4. In Case 3 there is a high density of ground investigations, the readings do show a discernible spatial trend as well as a closely spaced consistent set of readings. The polygon is split to become a Tier 2 polygon, and the closely spaced set is identified as a Tier 3 cluster.

This approach led to a high density of Tier 3 polygons in areas with lots of ground investigations (e.g. Christchurch), while low density areas were covered by Tier 1 and Tier 2 polygons (e.g. farmlands).

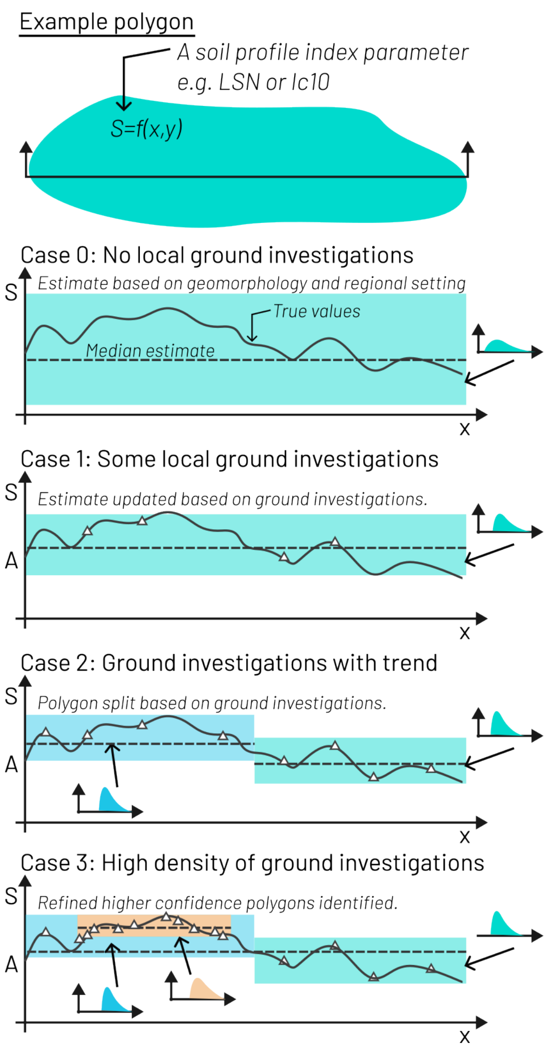

Figure 2 Conceptual example of classifying SSP for a polygon with different amounts of ground investigations

Kriging is often coupled with raster-based geospatial approaches to combine local data with geospatial predictions (applying kriging would have added to the challenge of using raster storage, since the variogram of LSN response is dependent on PGA, Mw and groundwater). However, this approach did not fit with the requirement of no spatial trend within an SSP. Instead, local data was combined with the geospatial prediction using Bayesian updating, where the distribution of likely LSN was modified based on the calculation of LSN from the ground investigations within each polygon.

4 OPPORTUNITIES

The framework has intentionally been designed such that it is one model that can produce multiple outputs to support a variety of use cases. The nationally consistent data-driven approach for liquefaction modelling presents unique opportunities for a variety of applications including loss modelling, post-disaster recovery planning, land use planning, research, and refining/testing policy settings. The precomputation of probability of land damage across a spectrum of PGA, Mw and groundwater depth values, provides blind prediction outputs of future events, thus allowing rigorous review of the models predictive capabilities. The model development is full prescribed such that more data (e.g. ground investigations, geomorphology maps, groundwater depth measurements) directly feeds through to more refined outputs. Furthermore, the framework is fully modular to allow alternative processes, alternative data sets, or alternative sub models (e.g. more refined groundwater models) to be inserted within the model development steps or run time steps (refer Figure 1).

For geoprofessionals, the model development has addressed several industry challenges including novel processes to quantify liquefaction spatial variability, analysis of the correlation between the age of the deposit and LDMs, as well as comparisons between LDMs computed from cone penetration tests versus standard penetration tests and boreholes.

5 CONCLUSIONS

The development of the National Liquefaction Model had several geotechnical engineering, software engineering, and data science challenges that required balanced consideration of making something useable while capturing a spatial resolution of response appropriate for loss modelling. These challenges shaped the model framework to produce a decoupled framework consisting of several sub models. The model development combined a novel geospatial estimate with the results of over 10 million precomputed liquefaction triggering analyses to allow predictions to be generated for different seismic scenarios and groundwater estimates across the flatland of New Zealand. The model has been developed to support loss modelling, however, will contribute to informed land-use planning and wider research into liquefaction and hazard management. The full technical report that will follow the release of the model will provided additional details.

6 ACKNOWLEDGEMENTS

The authors would like to acknowledge the significant contributions to both the project direction and technical solutions from the peer review team, steering group, end user group and the Natural Hazards Commission (NHC) Toka Tū Ake, as well as from Sjoerd van Ballegooy, Tonkin + Talyor Geotechnical and Data and Analytics staff, and Geotechnics staff.

7 REFERENCES

Boulanger RW, Idriss IM (2016). “CPT-based liquefaction triggering procedure”. *Journal of Geotechnical and Geoenvironmental Engineering*, 142(2): 04015065. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001388

Foster KM, Bradley BA, McGann CR and Wotherspoon LM (2019). “A VS30 map for New Zealand based on geologic and terrain proxy variables and field measurements”. *Earthquake Spectra*, 35(4): 1865–1897. https://doi.org/10.1193/121118EQS281M

Geyin M and Maurer BW (2020). “Fragility functions for liquefaction-induced ground failure”. *Journal of Geotechnical and Geoenvironmental Engineering*, 146(12): 04020142. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002416

Heron (2023). *Geological map of New Zealand 1:250,000*. 4th ed [Dataset]. GNS Science. https://doi.org/10.21420/5XTJ-5718 Idriss IM and Boulanger RW (2008). *Soil liquefaction during earthquakes*. Earthquake Engineering Research Institute (EERI).

Iwasaki T, Tokida K and Tatsuoka F (1981). “Soil liquefaction potential evaluation with use of the simplified procedure”.

NZGD (2023). *New Zealand Geotechnical Database (NZGD)* [Dataset]. https://www.nzgd.org.nz

Swiss Re (2014). *Small quakes, big impact: Lessons learned from Christchurch*. Swiss Re. https://www.swissre.com/dam/jcr:57469cc6-0a23-4100-8fe4-00663af35ad0/lessons_from_christchurch_web.pdf

Van Ballegooy S, Malan P, Lacrosse V, Jacka ME, Cubrinovski M, Bray JD, O'Rourke TD, Crawford SA and Cowan H (2014). "Assessment of liquefaction-induced land damage for residential Christchurch". *Earthquake Spectra*, 30(1): 31–55. https://doi.org/10.1193/031813EQS070M

Zhu J, Baise LG and Thompson EM (2017). "An updated geospatial liquefaction model for global application". *Bulletin of the Seismological Society of America*, 107(3): 1365–1385. https://doi.org/10.1785/0120160198