nzsee
NEW ZEALAND SOCIETY FOR EARTHQUAKE ENGINEERING

Development of an expanded earthquake and ground motion catalogue for New Zealand

J.A. Hutchinson, R. Lee & B. Bradley
University of Canterbury, Christchurch

C. van Houtte & A. Kaiser
GNS Science, Lower Hutt.

L. Wotherspoon
The University of Auckland, Auckland.

ABSTRACT

In the last two decades, New Zealand (NZ) has experienced significant earthquakes, including the 2010 M 7.2 Darfield, 2011 M 6.2 Christchurch, and 2016 M 7.8 Kaikōura events. Amongst these large events, tens of thousands of smaller earthquakes have occurred. While previous event and ground-motion databases have analyzed these events, many events below M 4 have gone undetected. The goal of this study is to expand on previous databases, particularly for small magnitude (M<4) and low-amplitude ground motions. This new database enables a greater understanding of regional variations within NZ and contributes to the validity of internationally developed ground-motion models. The database includes event locations and magnitude estimates with uncertainty considerations, and tectonic type assessed in a hierarchical manner. Ground motions are extracted from the GeoNet FDSN server and assessed for quality using a neural network classification approach. A deep neural network approach is also utilized for picking P and S phases for determination of event hypocentres. Relative hypocentres are further improved by double-difference relocation and will contribute toward developing shallow (< 50 km) seismic tomography models. Analysis of the resulting database is compared with previous studies for discussion of implications toward national hazard prediction models.

1 INTRODUCTION

New Zealand (NZ) is a tectonically diverse country with opposing subduction zones beneath the northern and southern islands that are joined by the Alpine fault, a nearly 500 km long right-lateral strike-slip fault.

Dozens of earthquakes are detected on a daily basis. While the majority of earthquakes have small magnitudes (< 4), there have been several large and devastating ruptures since 2010. These include the 2010 M 7.2 Darfield, 2011 M 6.2 Christchurch, and 2016 M 7.8 Kaikōura events. The 22 February 2011 M 6.2 Christchurch earthquake was responsible for at the least an estimated NZ$40 billion in damages to Christchurch city (Noy et al., 2016), and was itself an aftershock of the M 7.2 Darfield earthquake, which occurred on 3 September, 2010. The aftershock sequence of the 2011 Christchurch earthquake led to the discovery a previously unmapped fault system, extending to a depth of approximately 10 km (Bannister et al., 2011). Significant aftershocks proceeded with the 13 June 2011 Mw 6.0 and 23 December 2011 Mw 5.9 Christchurch earthquakes. An examination of these events and other significant earthquakes have informed the drive behind developing seismic hazard models based on ground-motion predictions, fault distributions, and site characteristics (e.g. Bradley, 2019; Dempsey et al., 2020; de la Torre et al., 2020). To further inform and improve on these models, a robust earthquake catalogue is essential.

For nearly two decades, continuous records of earthquakes have been recorded, located, and catalogued in NZ with the GeoNet project. These records can be easily accessed and downloaded via an FDSN webservice with popular tools such as the ObsPy module (Krischer et al., 2015; Wassermann et al., 2010) for the Python programming language. The information for tens of thousands of events can be accessed using the GeoNet Quake Search tool: https://quakesearch.geonet.org.nz. Despite countless hours spent cataloguing these events, many can still go undetected, particularly in regions with sparser geographic coverage by seismic networks. For this reason, we propose the implementation of recent computational techniques for the detection and location of small magnitude earthquakes. The discovery of additional earthquakes (and corresponding seismic zones and/or faults) will in turn lead to improvements of regional and national seismic hazard predictions.

With the freely available GeoNet FDSN seismic waveform catalogue, we develop an approach to automatically detect and locate earthquakes, while computing local magnitudes (M_L). These events will be catalogued alongside the currently available GeoNet dataset and will eventually include regional centroid moment tensor (CMT) solutions, focal mechanism solutions, and moment magnitude (M_W) measurements. Further, we plan to relocate events and update regional seismic tomography with a variety of approaches.

2 DATABASE OVERVIEW

The overarching goal of this project is to develop an earthquake and ground-motion intensity measure (IM) database that is regularly updated and easily accessed and understood. The catalogue will span all instrumentally recorded events until present and allows for expansion via additional tables and columns. We have modeled the overall structure of the database on the NGA-West2 format of Ancheta et al. (2014).

Several tables comprise this database, which are summarized in Figure 1. They are listed as follows: earthquake sources, propagation paths, station site information, and ground-motion IM details. Together, these tables provide the necessary information which can be used for ground-motion prediction as a component of seismic hazard analysis. This section provides an overview of each table and breaks down the information presented in each column.

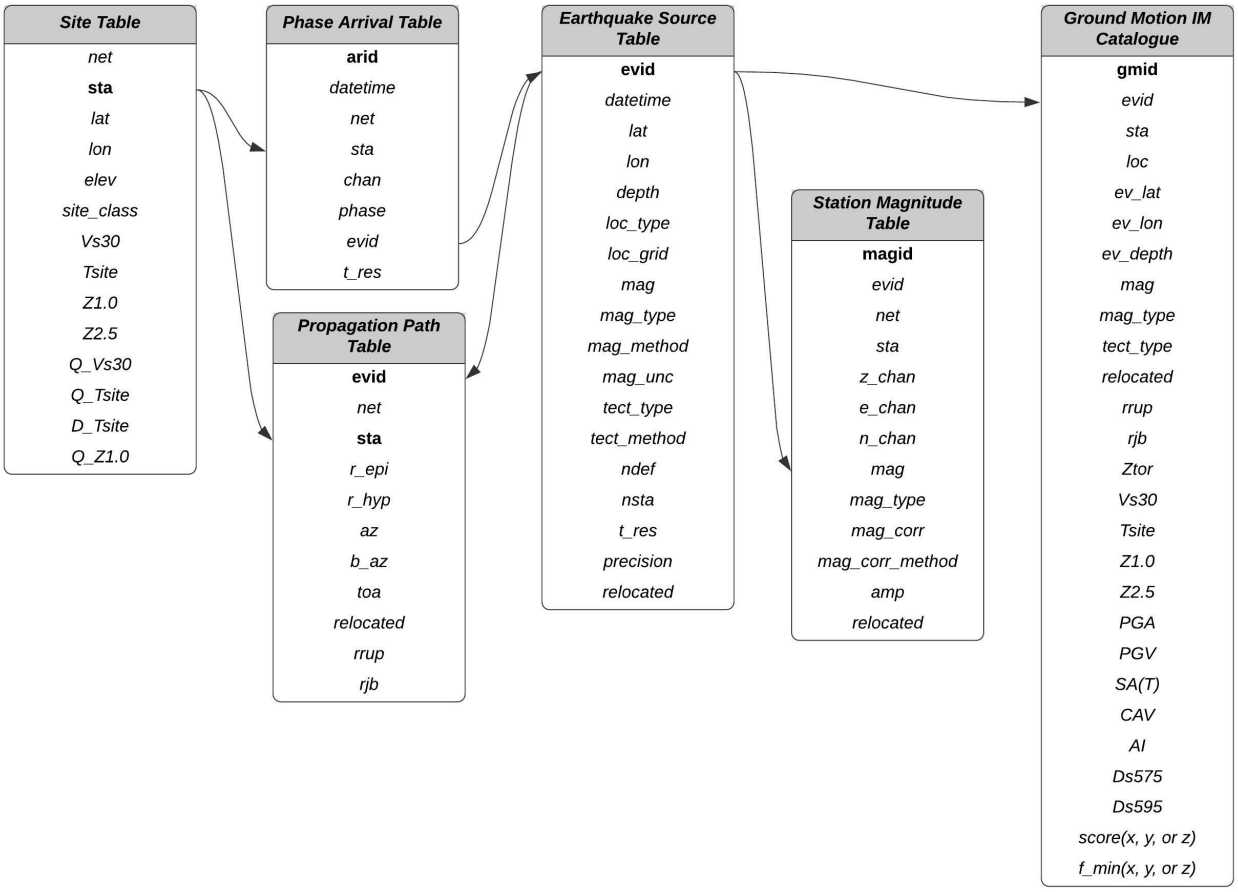

Figure 1: Proposed database relationship chart. Table names are presented at the top of each box. Column names are listed below the table names. Bolded names indicate the key identifiers for that table. Arrows indicate the general input/output relationship between tables. Descriptions for each table and their columns are provided in sections 2.1 – 2.6.

2.1 Site table

The site table contains basic station information, as well as site response data.

net, sta | Network and station name.

lat, lon | Station latitude and longitude coordinate in decimal degrees.

elev | Station elevation, in meters above sea level.

site_class | Site classification by rank (A is best, D is worst)

Vs30 | 30-m averaged shear-wave velocity (m/s).

Tsite | Fundamental site periods (s)

$Z_{1.0}$, $Z_{2.5}$ | Depths (in km) to shear-wave velocities of 1.0 and 2.5 km/s, respectively.

Q_Vs30 | Quality rating of Vs30 measurement (see Kaiser et al. (2017))

Q_Tsite | Quality rating of Tsite measurement (see Kaiser et al. (2017))

D_Tsite | Method used for determining site period (see Kaiser et al. (2017))

Q_Z1.0 | Quality rating of Z1.0 measurement (see Kaiser et al. (2017))

2.2 Phase arrival table

The phase arrival table includes information about earthquake phases that are associated with located earthquakes.

arid | Arrival identification number.

datetime | Time of phase arrival.

net, sta, chan | Seismic network, station, and channel where the arrival was detected.

phase | Phase of the arrival (e.g., *P*- or *S*-phases).

evid | Event identification number.

t_res | Time residual of the observed arrival from the predicted arrival (seconds)

2.3 Earthquake source table

The earthquake source table includes location and magnitude information for recorded events. These data will be collated from both the available GeoNet catalogue, global catalogues, as well as our internally developed catalogue.

evid | Event identification number.

datetime | Origin time of the event.

lat, lon, depth | Origin latitude and longitude of the event in decimal degrees, with depth in km below sea level.

loc_type, loc_grid | Location method or catalogue used for the event and when applicable, the location grid used for location.

mag, mag_type, mag_method, mag_unc | The magnitude and magnitude type of the event (e.g. direct M_W vs M_L corrected to M_W) and method used to calculate it, as well as the uncertainty.

tect_type, tect_method | The tectonic type of the event (crustal, interface, or slab) and the method used to determine it

ndef, nsta | The number of picks and stations used in locating the earthquake

t_res, precision | When applicable, the origin time residual in seconds and the decimal precision of the hypocenter in kilometers.

relocated | Whether the associated event has been relocated (yes or no).

2.4 Station magnitude table

The station magnitude table contains local and corrected local magnitudes for source-receiver pairs.

magid | Magnitude identification number.

evid | Event identification number associated with the source-receiver pair.

net, sta, z_chan, e_chan, n_chan | Network, station, and 3-component channels used to determine magnitude.

mag | Uncorrected magnitude.

mag_type | Magnitude type of the uncorrected magnitude.

mag_corr | Corrected local magnitude.

mag_corr_method | Method used to correct local magnitude. Currently using equations from Rhoades et al. (2020).

amp | Peak amplitude of the signal in mm.

relocated | Whether the associated event has been relocated (yes or no).

2.5 Propagation path table

The propagation path table contains source – receiver information, such as distance and azimuth.

evid | Event identification number.

net, sta | Network and station name.

r_epi, r_hyp | Epicentral and hypocentral distances between station and event (km).

az, b_az | Azimuth from the event to the station and back-azimuth from the station to the event.

toa | Takeoff angle from the event to the station. Currently all toas are derived from those reported in the GeoNet database.

relocated | Whether the associated event has been relocated (yes or no).

rrup, rjb | Shortest distance to the rupture plane and Joyner-Boore distance, the shortest distance to the surface projection of a rupture plane (km). These are roughly equivalent to r_hyp and r_epi, respectively.

2.6 Ground motion intensity measure tables

The ground motion table provides specific information about IMs for each event across stations with recorded data. IMs include peak-ground-acceleration (PGA), peak-ground-velocity (PGV), and the pseudo-spectral-accelerations at vibration periods ranging from T = 0-10 s. When possible, ground motions are classified with a quality score from 0-1 for each record (as long as it is part of a 3-component seismogram). This is achieved via the ground-motion classifier neural network (GMC), developed by Dupuis et al. (2020). This catalogue is split into several tables, which depend on the component (000, 090, ver, rotD50, and rotD100). The columns listed here are for the flat-file version of the database tables.

gmid | Ground motion identification number.

evid | Event identification number.

net, sta, loc | Network, location (number associated to different instruments at a site/station), and station name.

ev_lat, ev_lon, ev_depth, mag, mag_type, tect_type | Event latitude, longitude, depth (km), magnitude, magnitude type, and tectonic type.

relocated | Whether the associated event has been relocated (yes or no).

R_{rup}, R_{jb} | Shortest distance to the rupture plane and Joyner-Boore distance, the shortest distance to the surface projection of a rupture plane (in km). These are roughly equivalent to r_hyp and r_epi, respectively.

Z_{TOR}| Depth to the top of the rupture (in km).

Vs_{30} | 30-m averaged shear-wave velocity (m/s).

Tsite | Fundamental site periods (s).

$Z_{1.0}$, $Z_{2.5}$ | Depths (in km) to shear-wave velocities of 1.0 and 2.5 km/s, respectively.

PGA, PGV | Peak ground acceleration (g) and velocity (cm/s).

SA(T) | Pseudo-acceleration response spectra (spectral acceleration, g) for various periods, T.

CAV, AI | Cumulative absolute velocity (g*s) and Arias intensity (m/s).

Ds575, Ds595 | Significant duration (s) for different energy thresholds (5-75% and 5-95%).

score_X, score_Y, score_Z | Classification score (generally from 0-1) on the quality of the GM for X, Y, and Z channels.

f_min_X, f_min_Y, f_min_Z | Minimum viable frequency (Hz) of the GM for the X, Y, and Z components.

3 METHODOLOGY FOR AN IMPROVED EARTHQUAKE CATALOGUE

An essential component of this project is to expand the number of available events, and improve on locations and moment tensors for the earthquake catalogue. By lowering the magnitude of completeness of the earthquake catalogue, we can obtain a larger number of events and recorded ground motions, which is essential to understand systematic ground motion features that can be used to improve ground motion prediction. This section outlines our approach to expanding and improving on the earthquake catalogue.

3.1 Automated location procedure

Automatic detection of earthquakes is a necessary approach when dealing with an enormous dataset. The standard approach has been to use a simple short-term/long-term average (STA/LTA) search to find peak signals that stand out from background noise on seismograms. Unfortunately, seismic records can be quite noisy, which often masks the earthquake signal. Because signals are often missed (due to a low STA/LTA ratio) or are falsely identified due to spikes in noise, an analyst must painfully review seismograms for the best results. This practice is time-consuming and strongly dependent on the analyst.

With recent advances in neural network and machine learning, we can apply an approach to identifying the *P* and *S*-phases on less-than-ideal data. We implement EQTransformer (Mousavi et al., 2020), a deep-learning model that simultaneously detects and picks both *P* and *S* phases. This method applies weighting to attention layers of specific portions of a detected waveform in order to attain accurate arrival times in combination with the trained model.

Following earthquake phase detection and picking, the arrivals are associated as events. This association is based on detection times across multiple stations within a moving window of 15 seconds, requiring at least 4 stations to be considered an event.

The associated arrivals are used to locate the earthquake in a 3-D velocity structure. First, we interpolate between data points from the New Zealand-wide V2.2 seismic velocity model (Eberhart-Phillips et al., 2020), so that we have a grid spacing of 10 km in horizontal, and 4 km in the vertical direction. The velocity grid is used to calculate separate *P* and *S* travel-times for each station in the New Zealand national network (and including the IRIS station SNZO). Travel-times are solved for with PyKonal, the Eikonal-based Python package developed by White et al. (2020). Finally, the associated arrival times and 3-D travel-time tables are used to solve for the most likely location using a modified application of the maximum intersection (MAXI) method (Font et al., 2004), which provides time residual statistics as a means of evaluating the precision of the solution.

The complete daily event and arrival catalogues are used to compute M_L, which is written out to a separate database. In addition, we use the M_L to M_W relationship established by Ristau et al. (2016), and improved by Rhoades et al. (2020), to report a corrected M_L, which accurately translates to the M_W scale. Event information is also compiled in QuakeML XML files for convenient, portable delivery.

3.2 Relocation and local seismic tomography improvements

As mentioned in the Introduction, we intend to test various relocation and seismic tomography methods, including SIMUL2014 (Thurber and Eberhart-Phillips, 1999) and hypoTD (Guo and Zhang, 2017). SIMUL2014 derives local tomography from earthquake travel times differences. Earthquake raypaths from events outside of the study area can be implemented into the solution, and the program requires little configuration in order to function correctly. hypoTD is a triple-difference hypocentre relocation program that derives solutions from catalogues containing absolute travel time, travel-time difference, and cross-correlation difference data. Event-pair and station-pair double difference data can be combined into double-pair double difference data to ascertain absolute locations with more accuracy than either individual method, improving upon the frequently cited HypoDD (Waldhauser and Ellsworth, 2000) and TomoDD (Zhang and Thurber, 2003) methods.

3.3 Regional moment tensor

As a final step in developing an expanded earthquake catalogue, we intend to add to, and hopefully improve on, the regional CMT data. A probabilistic approach to computing source inversions is provided by the Grond package (Heimann et al., 2018), which is part of the Pyrocko suite of Python packages (Heimann et al., 2017). This method is highly configurable, allowing the user to control several inputs such as waveform data, GNSS data, and Green's function stores (which are calculated externally). Forward modelling is used to invert for a solution over thousands of iterations, while jointly bootstrapping to generate meaningful statistics. The data is reported in a HTML-based document and can be used to quickly assess whether a CMT solution has been accurately determined for an event.

For this project, initial tests on $M_W > 5+$ events have been successful. At lower magnitudes (and at greater distances from the source), however, seismograms are often contaminated by background noise. This problem could potentially be alleviated by a noise removal approach, such as with the DeepDenoiser neural network (Zhu et al., 2018)(Zhu et al., 2019). Unfortunately, in our exploration to date, the software currently struggles in removing noisy signals at the low frequencies required for CMT inversions (approximately 0.01 – 0.05 Hz) over signal lengths exceeding a 30 second time-window.

4 DISCUSSION

Expanded earthquake and ground-motion catalogues for use in ground-motion modelling as part of the National Seismic Hazard Model update are currently in the early stages of development. The entire GeoNet earthquake database from 2003 – August 2020 has been converted to the format shown in Figure 1 and described in Section 2.3. From this catalogue, we have computed ground motion qualities for all earthquakes with M >= 4. Example low- and high-quality ground motions are shown in Figure 2. This catalogue will be further expanded to include ground motion IMs and will be implemented with the automated catalogue.

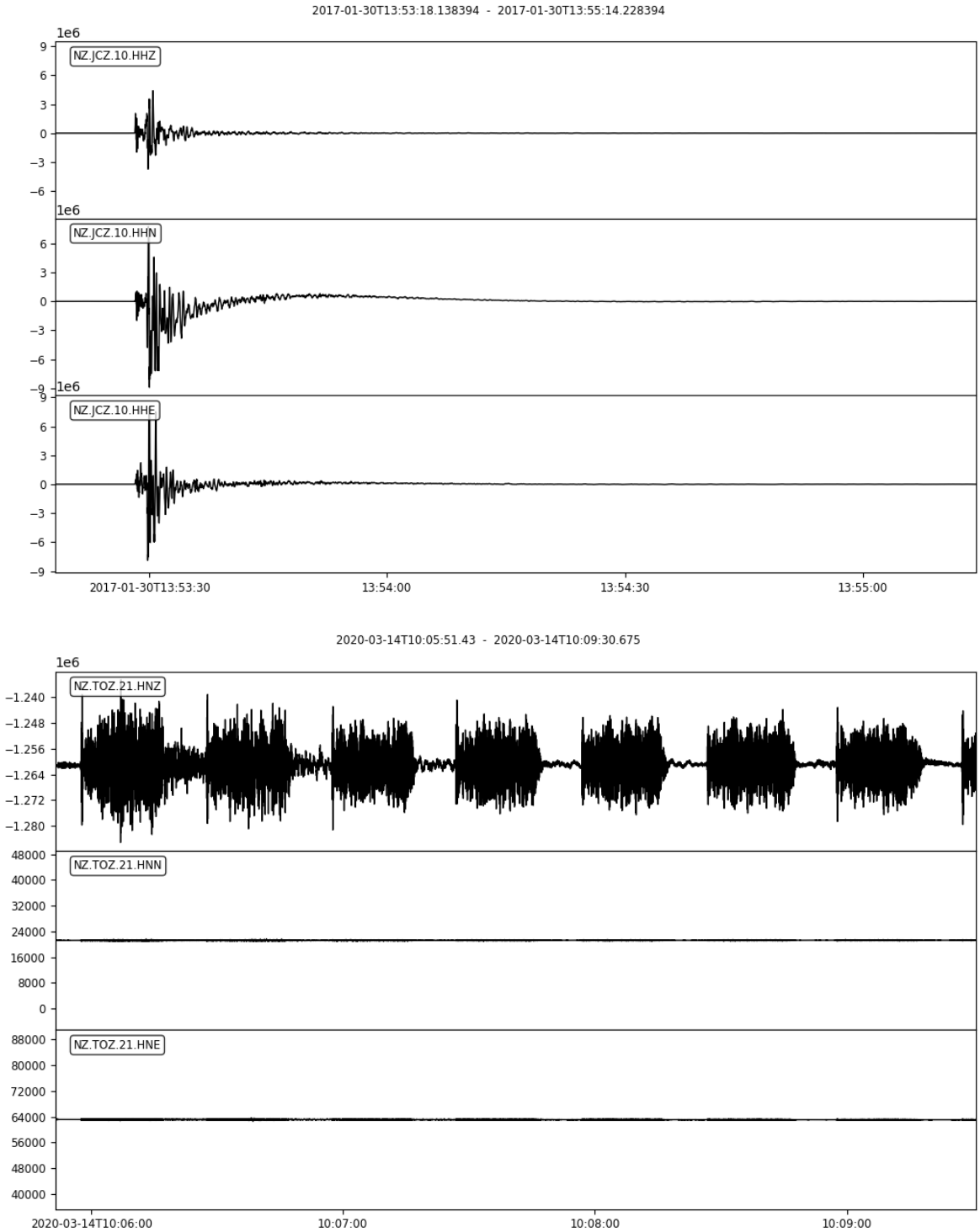

Figure 2: An example high quality ground motion detected by the GMC (top) vs a poor quality ground motion (bottom).

For a comparison between the GeoNet earthquake catalogue and the automated location procedure (Section 3.1), we have selected data from the first week of January 2011, the year of the Mw 6.2 Christchurch earthquake. As an example, on 2 January 2011, 87 events are provided in the GeoNet catalogue. Our automated approach identified and located 110 events (Figure 3). Of these events, 55 have closely matched origin times within 10 seconds. Many stations were not included in this test, particularly from the South Island. Further matches will be possible with refinement of the association and location parameters and the

inclusion of all stations with phase detections. Over half of the GeoNet solutions for this date have operator assigned depth values, while our results provide depth constraint based on location within a 3-D velocity structure. As an added advantage, our location method also includes computation of the corrected M_L (Rhoades et al., 2020) as part of the automated process.

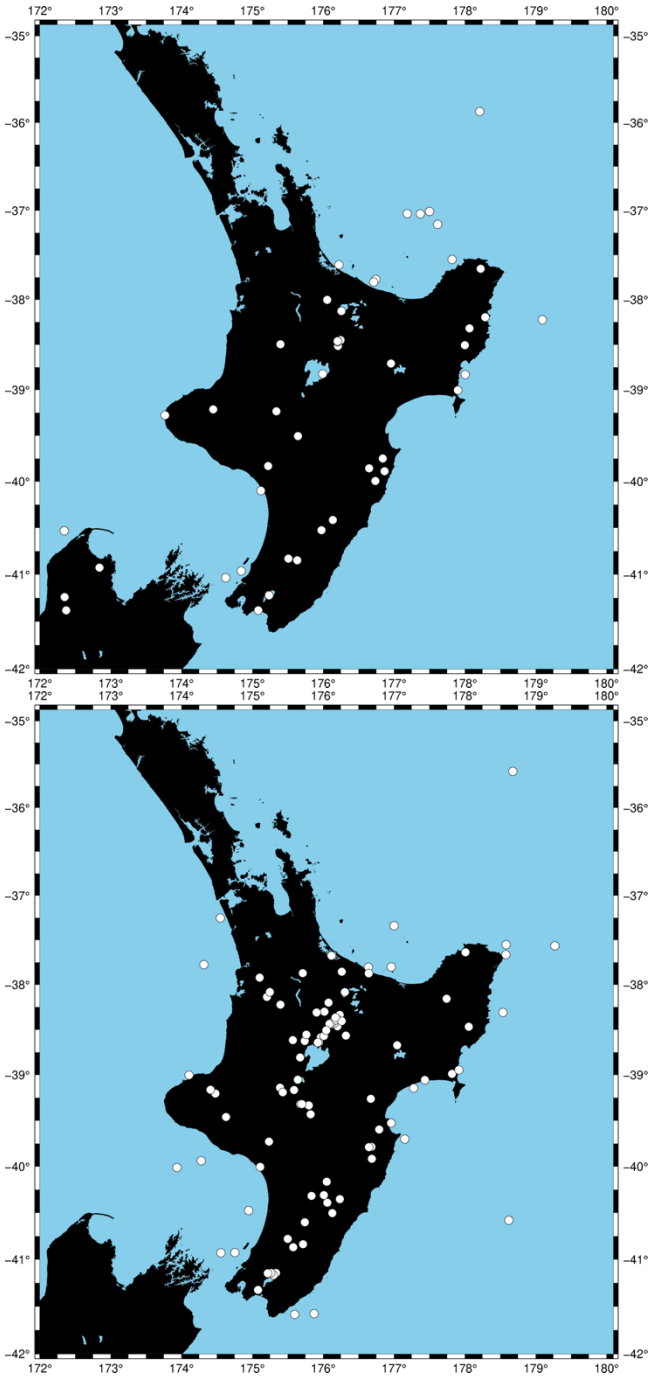

Figure 3: Comparison of GeoNet earthquake epicentres (left) to the new automated catalogue (right) for 2 January 2011. Note that many of the stations were not used for the automated solutions, but will be included as the project matures.

5 CONCLUSION

The expanded earthquake and ground motion catalogues are in the earliest stages of development. As procedure is firmly implemented and thoroughly tested, we plan to analyze the back-catalogue of data from

2003 until 2020. Once completed, we plan to implement regular updates on a weekly or monthly basis to the established catalogue. The complete earthquake catalogue will expand on our understanding of regional hazards, with ground motion intensity measures providing a means of assessing site response to various styles and distances from rupture.

Further objectives include extending the regional CMT catalogue to lower magnitudes and to develop detailed regional seismic tomography models. A greater number of CMT solutions will inform fault rupture models. Higher resolution seismic tomography will lead to more accurate locations and inform hazard models.

6 REFERENCES

Ancheta, T. D., Darragh, R. B., Stewart, J. P., Seyhan, E., Silva, W. J., Chiou, B. S. J., Wooddell, K. E., Graves, R. W., Kottke, A. R., Boore, D. M., et al. (2014). NGA-West2 database. *Earthquake Spectra*, *30*(3), 989–1005. https://doi.org/10.1193/070913EQS197M

Bannister, S., Fry, B., Reyners, M., Ristau, J., and Zhang, H. (2011). Fine-scale relocation of aftershocks of the 22 February Mw 6.2 Christchurch earthquake using double-difference tomography. *Seismological Research Letters*, *82*(6), 839–845. https://doi.org/10.1785/gssrl.82.6.839

Bradley, B. A. (2019). On-going challenges in physics-based ground motion prediction and insights from the 2010–2011 Canterbury and 2016 Kaikoura, New Zealand earthquakes. *Soil Dynamics and Earthquake Engineering*, *124*, 354–364. https://doi.org/10.1016/j.soildyn.2018.04.042

Dempsey, D., Eccles, J. D., Huang, J., Jeong, S., Nicolin, E., Stolte, A., Wotherspoon, L., and Bradley, B. A. (2020). Ground motion simulation of hypothetical earthquakes in the upper North Island of New Zealand. *New Zealand Journal of Geology and Geophysics*. https://doi.org/10.1080/00288306.2020.1842469

Dupuis, M., Schill, C., Lee, R., and Bradley, B. (2020). A neutal network for ground motion quality classification from New Zealand earthquakes of variable magnitudes and tectonic types. Retrieved from https://ir.canterbury.ac.nz/handle/10092/101432

Eberhart-Phillips, D., Bannister, S., Reyners, M., and Henrys, S. (2020). New Zealand Wide model 2.2 seismic velocity and Qs and Qp models for New Zealand. https://doi.org/10.5281/ZENODO.3779523

Font, Y., Kao, H., Lallemand, S., Liu, C. S., and Chiao, L. Y. (2004). Hypocentre determination offshore of eastern Taiwan using the maximum intersection method. *Geophysical Journal International*, *158*(2), 655–675. https://doi.org/10.1111/j.1365-246X.2004.02317.x

Guo, H., and Zhang, H. (2017). Development of double-pair double difference earthquake location algorithm for improving earthquake locations. *Geophysical Journal International*, *208*(1), 333–348. https://doi.org/10.1093/gji/ggw397

Heimann, S., Kriegerowski, M., Isken, M., and Cesca, S. (2017). Pyrocko-An open-source seismology toolbox and library. Retrieved from https://gfzpublic.gfz-potsdam.de/pubman/faces/ViewItemFullPage.jsp?itemId=item_2280895_2

Heimann, S., Isken, M., Kühn, D., Sudhaus, H., and Steinberg, A. (2018). Grond: A probabilistic earthquake source inversion framework. Retrieved from https://gfzpublic.gfz-potsdam.de/pubman/faces/ViewItemFullPage.jsp?itemId=item_3615907_6

Kaiser, A., Van Houtte, C., Perrin, N., Wotherspoon, L., and Mcverry, G. (2017). Site characterisation of GeoN et stations for the New Zealand strong motion database. *Bulletin of the New Zealand Society for Earthquake Engineering*, *50*(1), 39–49. https://doi.org/10.5459/bnzsee.50.1.39-49

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. (2015). ObsPy: a bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, *8*(1), 014003. https://doi.org/10.1088/1749-4699/8/1/014003

de la Torre, C. A., Bradley, B. A., and Lee, R. L. (2020). Modeling nonlinear site effects in physics-based ground motion simulations of the 2010–2011 Canterbury earthquake sequence. *Earthquake Spectra*, *36*(2), 856–879. https://doi.org/10.1177/8755293019891729

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., and Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-17591-w

Noy, I., Parker, M., and Wood, A. (2016). The Canterbury rebuild five years on from the Christchurch earthquake. *Reserve Bank of New Zealand Bulletin*, *79*(3), 3. Retrieved from https://www.questia.com/library/journal/1G1-452290686/the-canterbury-rebuild-five-years-on-from-the-christchurch

Rhoades, D. A., Christophersen, A., Bourguignon, S., Ristau, J., and Salichon, J. (2020). A Depth-Dependent Local Magnitude Scale for New Zealand Earthquakes Consistent with Moment Magnitude. *Bulletin of the Seismological Society of America*. https://doi.org/10.1785/0120200252

Ristau, J., Harte, D., and Salichon, J. (2016). A revised local magnitude (ML) scale for New Zealand earthquakes. *Bulletin of the Seismological Society of America*, *106*(2), 398–407. https://doi.org/10.1785/0120150293

Thurber, C., and Eberhart-Phillips, D. (1999). Local earthquake tomography with flexible gridding. *Computers & Geosciences*, *25*(7), 809–818. https://doi.org/https://doi.org/10.1016/S0098-3004(99)00007-2

Waldhauser, F., and Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America*, *90*(6), 1353–1368. https://doi.org/https://doi.org/10.1785/0120000006

Wassermann, J., Barsch, R., Beyreuther, M., Behr, Y., Megies, T., and Krischer, L. (2010). ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, *81*(3), 530–533. https://doi.org/10.1785/gssrl.81.3.530

White, M. C. A., Fang, H., Nakata, N., and Ben-Zion, Y. (2020). PyKonal: A python package for solving the eikonal equation in spherical and cartesian coordinates using the fast marching method. *Seismological Research Letters*, *91*(4), 2378–2389. https://doi.org/10.1785/0220190318

Zhang, H., and Thurber, C. H. (2003). Double-difference tomography: The method and its application to the Hayward Fault, California. *Bulletin of the Seismological Society of America*, *93*(5), 1875–1889. https://doi.org/10.1785/0120020190

Zhu, W., Mousavi, S. M., and Beroza, G. C. (2019). Seismic Signal Denoising and Decomposition Using Deep Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(11), 9476–9488. https://doi.org/10.1109/TGRS.2019.2926772