The New Zealand Ground-Motion Database: A 2024 Update

A. Rampersad, J. Ridden, R. Lee, B.A. Bradley, C. Zhu, C. Pajaro, C. Schill

University of Canterbury, Christchurch, New Zealand.

ABSTRACT

This article summarizes 2024 updates to the New Zealand Ground-Motion Database (NZ GMDB). The database has been expanded to include 2024 events and addresses challenges with earthquake source metadata, such as earthquake location, magnitude, tectonic classification and finite-fault geometry, among others. Waveform extraction has been improved with longer time windows to better capture all significant arrivals, and improved P-wave picks through an analysis of multiple automatic picking methods. Aftershock flagging of events has been included in this revised version for ease of implementation into seismic hazard assessment. Lastly, we summarize several other improvements that are underway towards the next significant release of the NZ GMDB and inclusion in the NGA-West3 GMDB.

1 INTRODUCTION

The overarching goal of the New Zealand Ground-Motion Database (NZ GMDB) is to provide users with a high-quality dataset of seismic ground motions and their associated metadata. Key applications of such databases include evaluation of empirical ground-motion models (GMMs), development of NZ-specific adjustments of empirical GMMs and validation of physics-based GMMs, among others.

The strong motion database of Van Houtte et al., (2017) laid the framework for a comprehensive collection of ground-motion data specific to NZ and provided invaluable insights but also highlighted numerous areas for enhancement. Hutchinson et al., (2024) subsequently introduced significant improvements leading to a more holistic and comprehensive NZ GMDB. These enhancements included better metadata management, refined data extraction algorithms, and the integration of new methodologies for more accurate seismic event analysis.

A critical aspect of this evolution has also been the collaboration with the NGA (Next Generation Attenuation) projects. The NZ GMDB's active involvement in these projects ensures that it remains at the forefront of seismic data management and analysis procedures by staying aligned with the latest advancements and methodologies promoted by the NGA projects.

This article describes the improvements applied to the previous 2023 NZ GMDB (v3.4, Hutchinson et al., 2024). In particular, this iteration of the NZ GMDB (v4.3) has been meticulously scrutinized and improvements made towards inclusion into the NGA-West3 project. Other improvements relate to the quality of the waveforms extracted and algorithm adjustments to the Intensity Measure (IM) calculations. Overall, these improvements have led to an increased number of quality waveforms and a more robust database.

2 DATABASE OVERVIEW

Figure 1 presents an updated relational schematic of the NZ GMDB with updates since Hutchinson et al. (2024) highlighted in red text. Key changes are discussed in greater detail in the following section. There are no explicit updates to the site table indicated but the work previously carried out by Wotherspoon et al. (2022) has been continuously updated and carried through to the NZ GMDB.

In terms of the event additions, Figure 2 presents a spatial comparison between v3.4 (Hutchinson et al., 2024) and the latest stable version (v4.3). The quantity of high-quality ground motions from earthquakes with moment magnitude (M_w) greater than 3.0 has been updated to include the time period up to 31st December 2024. This has resulted in the present NZ GMDB consisting of 6040 earthquakes recorded at 393 stations with an associated ~70,000 ground motions. The notable increase in ground motions from the former database (~30,000 ground motions) is a result of both the increased time period and improved processing techniques.

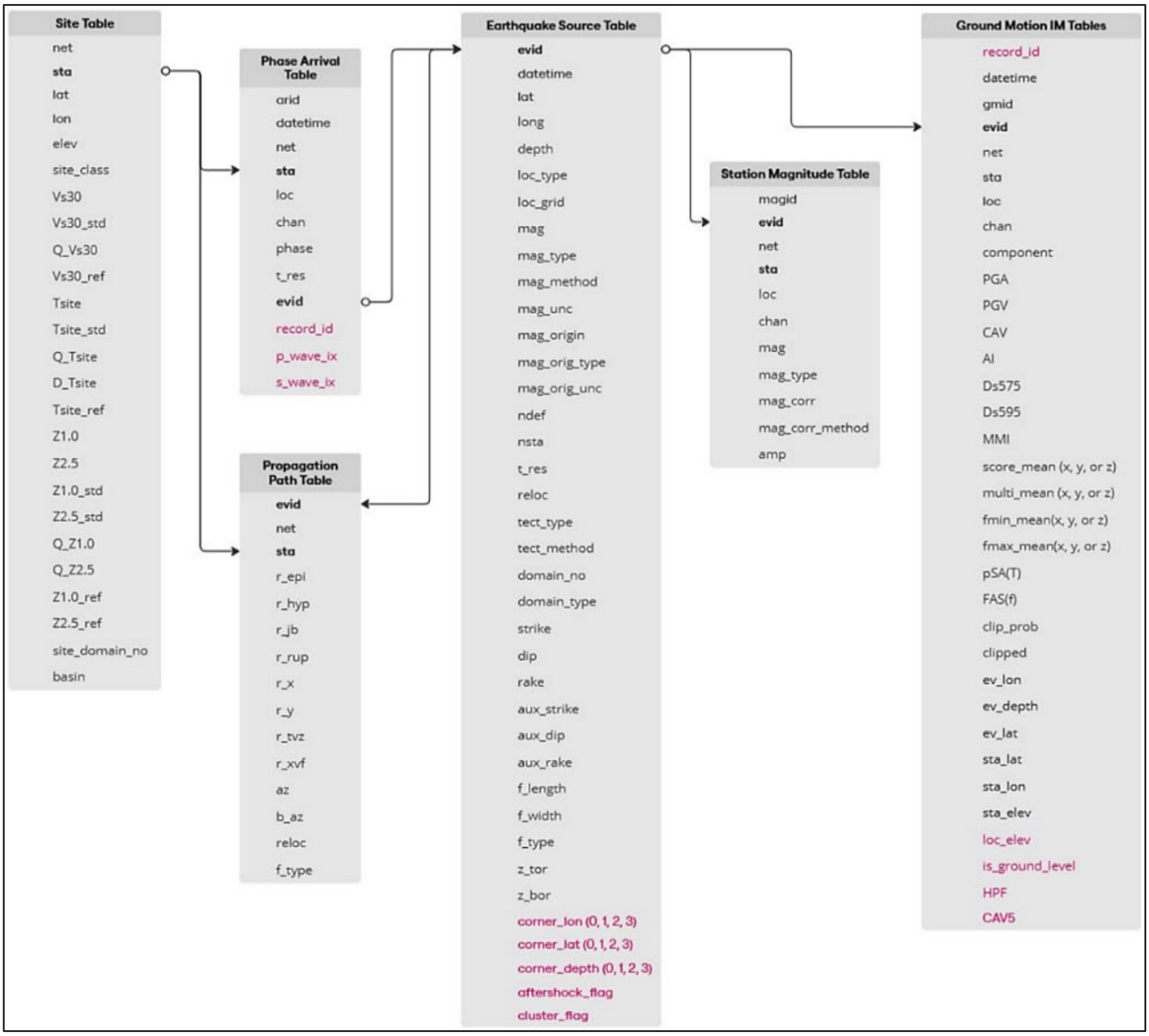


Figure 1 - Updated schematic of the relational setting of the six tables within the NZ GMDB

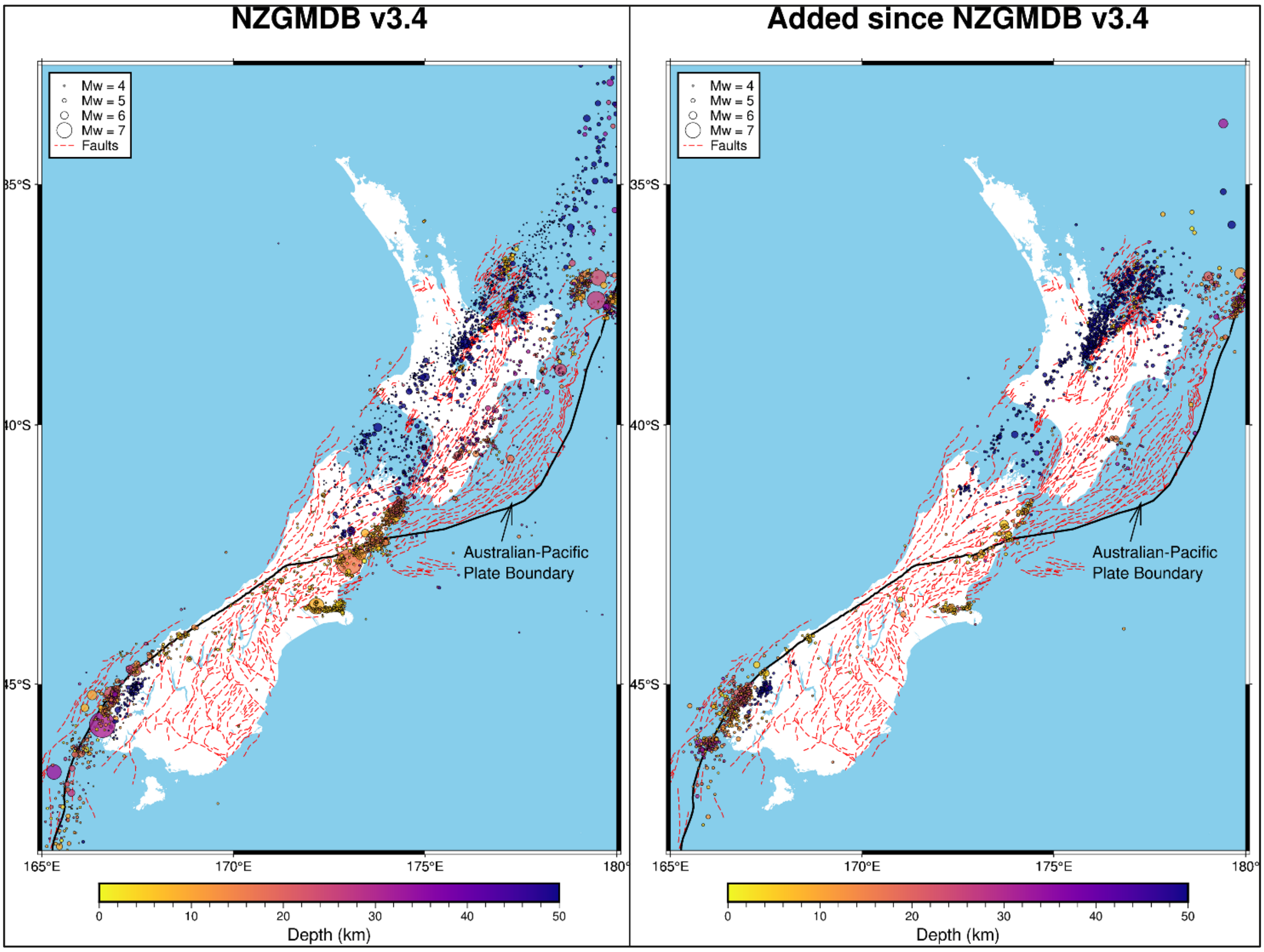


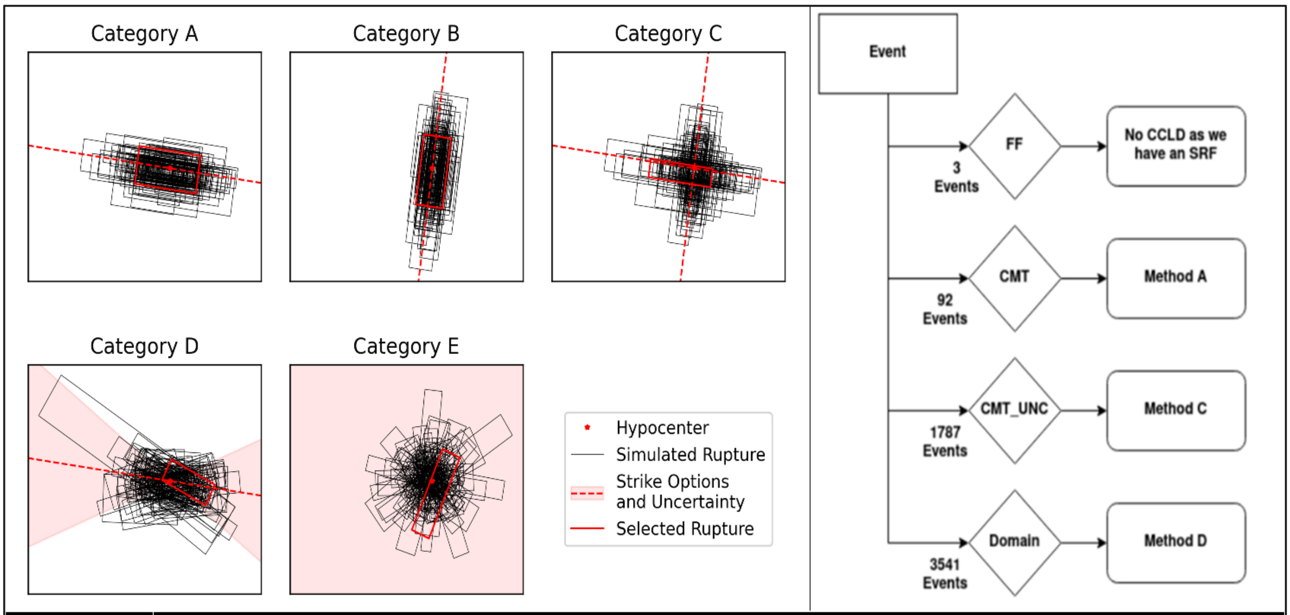
Figure 2 - Earthquake locations (Left) within NZGMDB v3.4 (Right) Added since NZGMDB 3.4

3 EARTHQUAKE CATALOGUE IMPROVEMENTS

3.1 Finite-fault geometry and propagation path table

Finite-fault models are an important aspect of earthquake source descriptions to ensure accurate source-to-site distances. Hutchinson et al. (2024) described bespoke finite-fault geometries for 3 events, with all other events utilising approximate fault geometries based on scaling relations discussed in the former database for estimated finite fault geometries.

The CCLD program (Chiou and Young, 2008) has been implemented into the processing workflow of the current NZGMDB in order to generate more robust estimates of R_{JB} and R_{rup}. It has also been used in the NGA-West2 and NGA-Subduction projects (Ancheta et al., 2014; Contreras et al., 2022). The algorithm, in essence, builds upon the scaling relations defined for various tectonic classes but instead of a single relationship, a stochastic set of possible rupture surfaces are generated for a given earthquake based on the available source metadata (M_w, dip, strike and rake). Preferred nodal planes (if known) can also be utilized to constrain the possible rupture surface via various categories depicted in Figure 3. The final selected rupture plane from which R_{rup} and R_{JB} are computed is based on the geometry which minimizes the squared difference between R_{rup} and the median R_{rup} from each simulation. Figure 3 also illustrates the workflow by which CCLD is implemented in the NZGMDB whereby each category is assigned to an event based on its solution. Finally, the geometry of the rupture for each event is also stored in the NZ GMDB as seen in Figure 1.



Category	Description
A	One or two nodal plane solution are known but the first is preferred
B	Two Nodal plane solutions are known but the second is preferred
C	Two nodal plane solutions are known but neither is preferred where each simulation will randomly select which nodal plane to use
D	Only one nodal plane solution is known or assumed with uncertainty where the given rake is used to assign rupture mechanism.
E	No nodal plane solutions are known/assumed. Rake, strike, and dip are randomly assigned with equal probability during each simulation.

Figure 3 - (Left) CCLD framework and rupture simulations (Right) Implementation in the NZGMDB (https://github.com/tristanbuckreis/ccldpy?tab=readme-ov-file)

3.2 Phase arrival table

The phase arrival table contains information about the earthquake phases that are associated with each event at a particular station. In particular, the P-wave onset (Parr) is used in many seismological applications such as earthquake locations and re-locations, source mechanism analysis and ground-motion processing for parameters such as Signal-to-Noise Ratio (SNR). The former NZ GMDB (v3.4) utilized P-wave onsets from the initial pass of the seismic network, GeoNet. Alternatively, if the network did not identify a P-wave onset, a custom APP PhasePicker (Kalkan 2016) was utilized, which idealizes the seismometer as a single-degree-of-freedom (SDOF) oscillator model. However, a manual inspection of 500 waveforms of varying ground motion quality revealed that PhasePicker led to erroneous picks, particularly on noisy datasets. It was, therefore, necessary to revise the choice in the algorithm as a means to secure the quality of the P-wave onsets.

A study was carried out on the 500 waveforms for various APPs as summarized in Table 1. The study involved manually picking P-wave onsets which are then compared to the P-wave onsets determined by the chosen algorithms. These algorithms were selected based on the various domains by which the waveforms were resolved and their popularity within the seismological community. The results of the study are summarized in Figure 4 where the PhaseNet Neural Network achieved the most accurate results with ~95% falling within 1 second of the manual P-wave onset.

Table 1 - Summary of Automatic P-wave Algorithms Assessed

Method	Domain/Type	Reference
TauPyModel - Ray theoretical Travel Times and Paths	Time/Theoretical	The TauP Toolkit: Flexible Seismic Travel-time and Ray-path Utilities (Crotwell H. P., et al 1999)
Short-Term Average/Long-Term Average or STA/LTA (Standard)	Time	Automatic earthquake recognition and timing from single traces (Allen RV., 1978)
Short-Term Average/Long-Term Average or STA/LTA (Recursive)	Time	A comparison of select trigger algorithms for automated global seismic phase and event detection (Withers M., 1998)
Autoregressive – Akaike Information Criterion (AR-AIC)	Time/Statistical	A technique for automatic detection of onset time of P-and S-Phases in strong motion records (Akazawa T., 2004)
Adapted Kurtosis-based P-wave Picker	Frequency/ Statistical	An Automatic Kurtosis-Based P - and S -Phase Picker Designed for Local Seismic Networks (Baillard C., 2014)
PhasePicker	Frequency	An Automatic P-Phase Arrival-Time Picker (Kalkan, 2016)
PhaseNet	Neural Network	PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method (Zhu W. & Beroza G., 2018)

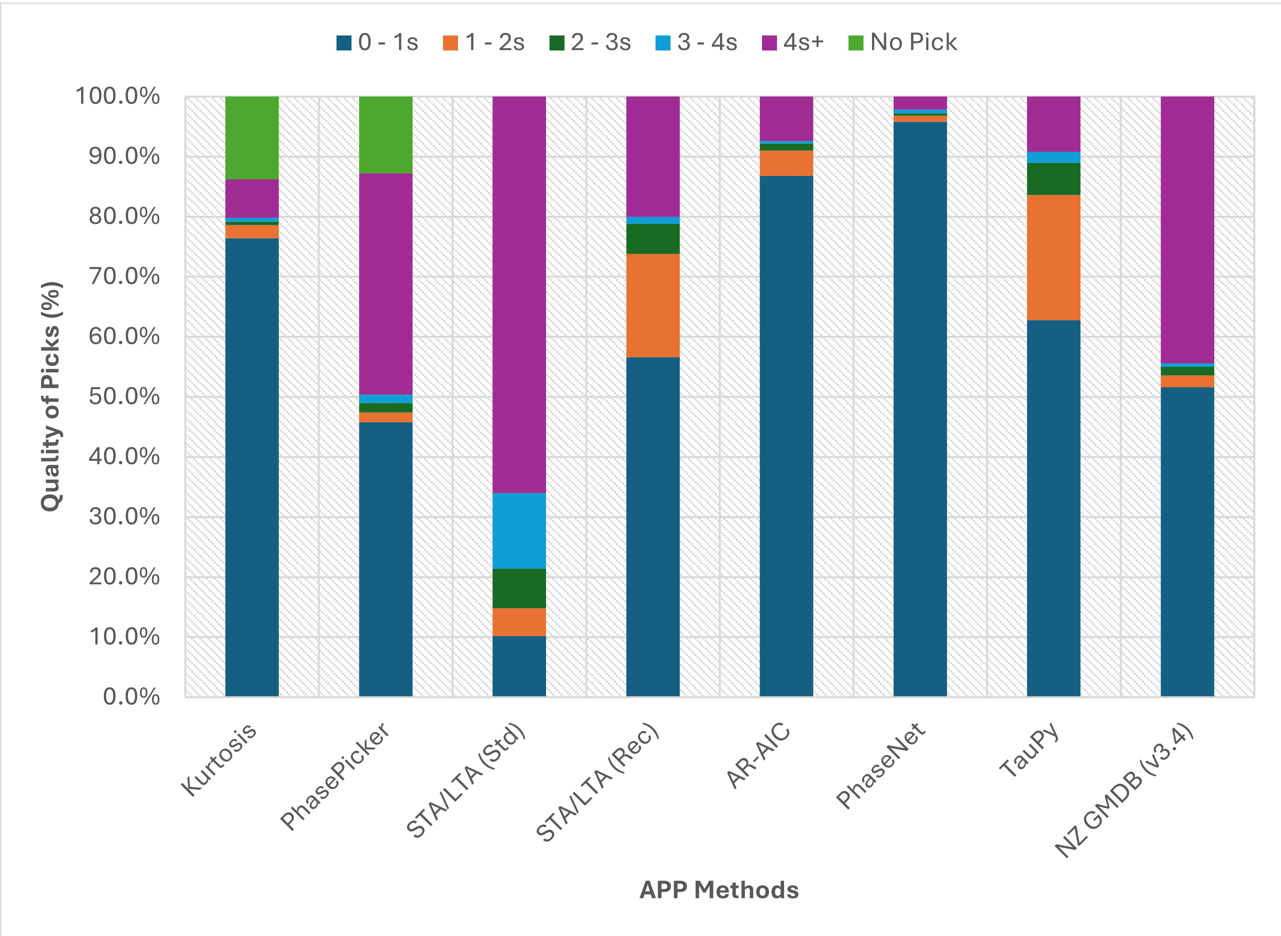


Figure 4 - Performance summary of the APP algorithms, showing how well automated P-wave arrival picks match the manual P-wave arrivals with the NZ GMDB v3.4 for reference

3.3 Ground motion extraction

Ground-motion waveforms are extracted from the GeoNet FDSN web service (https://www.geonet.org.nz/data/access/FDSN) using the time window defined by Equation 1. However, as depicted in Figure 5 there were several instances where waveforms were cut abruptly and/or extracted late leaving portions of the waveform missing. A manual inspection of 200 waveforms revealed that: (a) the 1.2 multiplier in Equation 1 applied to the 5%-95% significant earthquake duration (D_{s595}) was insufficient and (b) the late extraction was compounded by the poor P-wave onset determined by PhasePicker. The latter issue was addressed in the previous subsection and the multiplier has been increased to achieve a balance between window length and storage. The final multiplier in Equation 1 was adjusted from 1.2 to 2.0.

$$t_{win} = (S_{arr} + 1.2 \times D_{s595}) - (P_{arr} - 15), \quad (1)$$

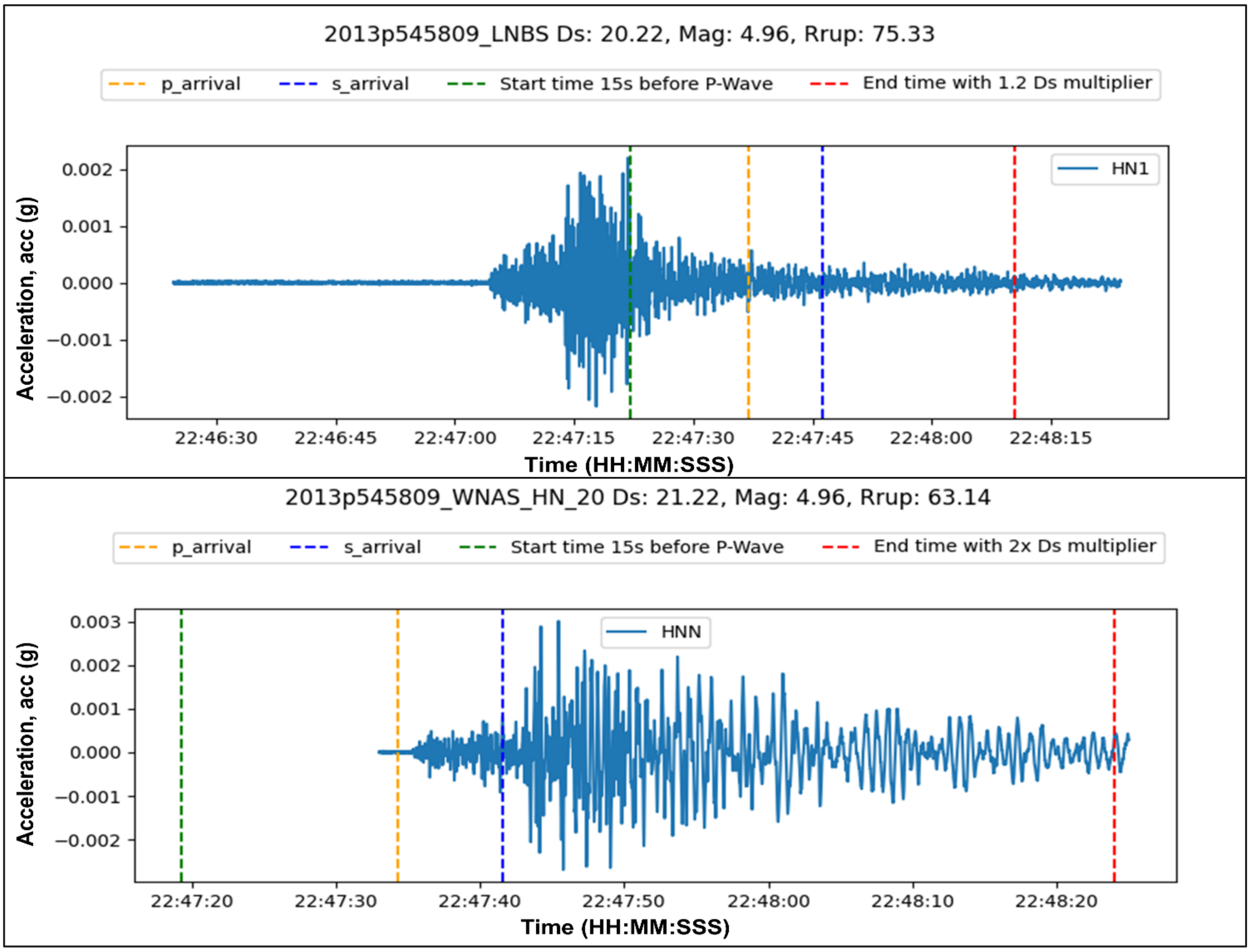


Figure 5 - (Top) Waveform extraction with 1.2 multiplier and PhasePicker APP (Bottom) Waveform extraction with 2.0 multiplier and PhaseNet APP

3.4 Ground-motion classifications

The quality of waveforms is assessed by a ground-motion classification (GMC) neural network (Dupuis et al., 2023), a deep learning-based model trained on NZ-specific data. A quality score of 0 – 1 is determined for each waveform. The criterion for inclusion requires that the minimum score of the three waveform components (X, Y and Z) must be greater than 0.5.

This criterion ensures that all three components of the records are of desirable quality. However, an inspection of 500 ground motions with variable GMC scores revealed that in several instances high quality horizontal records (>0.5) were being rejected due to poor quality vertical records. Therefore, in NZGMDB v4.3 the horizontal IM flatfiles (e.g., RotD50 and geomean) will require a quality of score greater than 0.5 on the X and Y waveforms. Similarly, vertical IM flatfiles will require a quality of score greater than 0.5 on the Z waveform.

3.5 Ground-motion intensity measures (IMs)

The range of IMs present in Hutchinson et al. (2024) are largely retained in this update but recent work with NGA-West3 recommends additional IMs. A summary of the additions are presented in Table 2 where the notable inclusions are the RotD0 component, cumulative absolute velocity above 5 cm/s^2 (CAV_5) and the effective amplitude spectrum (EAS). The utilization of RotD0, RotD50 and RotD100 as a three-component assessment of infrastructure and building performance during earthquakes is becoming increasingly popular due to their simplicity and robust estimations. Similarly, EAS represents an orientation-independent FAS metric which essentially preserves the mean power of the corresponding horizontal components similar to RotD50. CAV_5, on the other hand, was an IM introduced by Kramer and Mitchell (2006) which proved to be more closely aligned with excess pore pressure generation in potentially liquifiable soils. It has since gained popularity amongst practitioners and researchers for use in liquefaction hazard evaluations for rock motion inputs.

Other improvements to the IMs relate to optimized algorithms for the computation of the smoothed cumulative absolute velocity (for CAV_5) and Arias intensity (I_A). Occasionally, the resampling of the waveforms to compute these metrics would lead to gaps or plateaus in the time series resulting in inaccuracies in the calculated IMs. This is a common challenge in the processing of waveforms which often requires a rigorously optimized algorithm for resampling techniques. In this regard, the algorithm was revised and quality checked against manual calculations which led to overall more accurate estimates on the IMs.

Table 2 - Comparison of IMs between NZGMDB v3.4 and 4.3

Component	NZGMDB v3.4			NZGMDB v4.3		
	pSA	**FAS**	**Non-Spectral**	**pSA**	**FAS**	**Non-Spectral**
NS (000)	x	x	x	x	x	x
EW (090)	x	x	x	x	x	x
Z (Vertical)	x	x	x	x	x	x
geomean	x	x	x	x	x	x
RotD0	-	-	-	x	-	-
RotD50	x	x	x	x	-	-
RotD100	x	x	x	x	-	-
EAS	-	-	-	-	x	-
Non-spectral IMs: PGA, PGV, I_A, CAV, Ds575, Ds595, ***CAV_5 (New)***						

3.6 Aftershock flags

NGA projects employ the Woodell and Abrahamson (2014) methodology to discern between mainshocks and aftershocks denoted as C0 and C1 events, respectively. This methodology was preferred due to its physical accuracy in locating earthquakes as a function of the mainshock's rupture area. Traditional time-distance-window metrics developed by Gardner and Knopoff (1974) assume that the distance between C0 and C1 events are purely circular thus ignoring true rupture geometry. Three distance metrics were therefore proposed by Woodell and Abrahamson (2014); Woodell (2018) as summarized in Figure 6 which physically links C0 and C1 events via their ruptures.

To determine the suitability of the Woodell and Abrahamson (2014) methodology for the NZ GMDB, other common declustering techniques were assessed, as summarized in Table 3. Rupture areas for each event in the absence of explicit finite-fault geometries were computed with CCLD as discussed previously. The results revealed that the traditional time-distance-window methods aggressively characterized several likely C0 events as C1 whereas the probabilistic methods submissively characterized likely C1 events as C0. On the other hand, the Woodell and Abrahamson (2014) method provided a reasonable balance between physical accuracy and probabilistic likelihood based on the time-window adapted from Gardner & Kopff (1974). However, a specified cut-off distance by which searches are limited for nearby ruptures are required where a cut-off distance of 0 km represents events within the time-window that are located directly within a mainshock rupture. For a cut-off distance of 0 km, 80% of events are classified as C0 whereas a cut-off distance of 5 km equates 61%. This is a substantially more feasible prediction as opposed to the traditional time-distance-windows (e.g., Gardner and Knopoff 1974) which classifies 45% of events as C0.

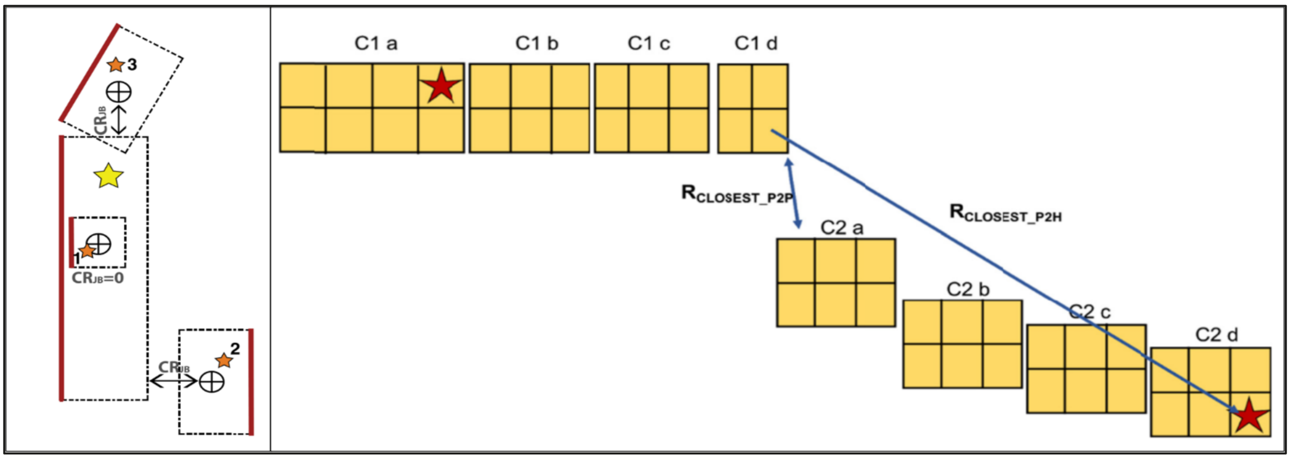


Figure 6 - (Left) crjb distance as defined by Woodell and Abrahamson (2014) (Right) rclosestp2h and rclosestp2p additional distances as defined by Woodell (2018)

Table 3 - Summary of declustering algorithms assessed

Method	Model/Type	Reference
Gardner & Knopoff 1974	Time-Distance-Window	Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian - Gardner & Knopoff 1974
Gruenthal 1985		The updated earthquake catalogue for the German Democratic Republic and adjacent areas - statistical data characteristics and conclusions for hazard assessment - G. Gruenthal 1985
Uhrhammer 1986		Characteristics of northern and central California seismicity, Earthquake Notes, 57 - Uhrhammer 1986
Musson 1999 (Afteran)		Probabilistic seismic hazard maps for the North Balkan region - Musson 1999
Woodell & Abrahamson 2014		Classification of Main Shocks and Aftershocks in the NGA-West2 Database - Abrahamson & Woodell 2014
Reasenberg 1985	Time-Distance-Window/Probabilistic	Second-order moment of central California seismicity - Reasenberg 1985
Nearest Neighbour Distance (NND)	Probabilistic	Earthquake Declustering Using the Nearest-Neighbor Approach in Space-Time-Magnitude Domain - Zaliapin & Ben-Zion 2020
Stochastic Reconstruction	ETAS/Stochastic	Stochastic declustering of space-time earthquake occurrences - Zhuang et al., 2002

4 FURTHER IMPROVEMENTS

4.1 Ground-motion classification bypass

The GMC Neural Network currently serves as the base filter for automatically rejecting or keeping records within the NZ GMDB. However, as noted in the previous sections, a visual inspection of records often provides a more reliable assessment of ground motion quality (but is time-consuming). Currently, several ground-motion records from large M_w events are rejected which is undesirable due to their importance in seismic studies. We plan to implement a bypass system for events of interest (M_w> 5) if GMC criteria are not met during the automated classification process. These events will then be manually screened where a new score can be assigned to determine the most suitable outcome before rejection or inclusion into the NZ GMDB.

4.2 Inclusion of broadband seismometer waveforms

There are many broadband seismometers installed and monitored by GeoNet which are not currently utilized in the NZ GMDB due to reliability issues. Inclusion of these records can substantially improve the density of waveforms across NZ, particularly in locations where the strong motion instruments (HN and BN Channels) contain gaps but the broadband seismometers (e.g., HH channels) do not. Efforts to determine the reliability of broadband instruments towards the inclusion of their recorded ground motions are underway.

4.3 Earthquake relocation and seismic tomography improvements

Current earthquake locations utilize 1D location grids and a global 3D velocity model respectively which can lead to uncertainties. Reyners et al., (2011) relocated earthquakes utilizing a NZ-specific 3D velocity model (Eberhart-Phillips et al., 2010) between January 2001 – March 2010. Subsequent advancements in NZ-specific 3D velocity models (Eberhart-Phillips, 2022) present an opportunity for more robust estimates of earthquake locations. These advancements will be utilized and build upon past earthquake relocations via well-established relocation algorithms such as HypoDD (Waldhauser and Ellsworth, 2000) and SIMUL14 (Thurber and Eberhart-Phillips, 2021).

4.4 Finite-fault geometry

Large M_w events often require additional attention to finite-fault geometry due to inherent rupture complexities e.g., multi-fault rupture. Currently, only 3 events have inverted finite rupture models; the Darfield, Christchurch and Kaikoura earthquakes. Through parallel work, we have developed finite fault geometries for 11 additional events, including the Fiordland and Dusky Sound earthquakes (Pajaro et al., 2025), which will be added in a subsequent database version.

5 CONCLUSION

In this article, various shortcomings of the NZGMDB v3.4 were addressed which led to an increased quantity of ground motions and an overall improvement in the quality of the database. Additionally, in the continued efforts to align with the NGA-West3 project, various additions were made such as the implementation of CCLD for source-to-site distances and the inclusion of aftershock flags. Further improvements have also been highlighted which are scheduled to be implemented prior to the next stable release and publication of the NZGMDB.

6 REFERENCES

Abrahamson, N. A., & Silva, W. J. (1997). Empirical Response Spectral Attenuation Relations for Shallow Crustal Earthquakes. *Seismological Research Letters*, *68*(1), 94–127. https://doi.org/10.1785/gssrl.68.1.94

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., Kishida, T., & Donahue, J. L. (2014). NGA-West2 Database. *Earthquake Spectra*, *30*(3), 989–1005. https://doi.org/10.1193/070913EQS197M

Bannister, S., Fry, B., Reyners, M., Ristau, J., & 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

Contreras, V., Stewart, J. P., Kishida, T., Darragh, R. B., Chiou, B. S., Mazzoni, S., Youngs, R. R., Kuehn, N. M., Ahdi, S. K., Wooddell, K., Boroschek, R., Rojas, F., & Órdenes, J. (2022). NGA-Sub source and path database. *Earthquake Spectra*, *38*(2), 799–840. https://doi.org/10.1177/87552930211065054

Eberhart-Phillips, D., Reyners, M., Bannister, S., Chadwick, M., & Ellis, S. (2010). Establishing a Versatile 3-D Seismic Velocity Model for New Zealand. *Seismological Research Letters*, *81*(6), 992–1000. https://doi.org/10.1785/gssrl.81.6.992

Gardner, J. K., & Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? Bull. Seismol. Soc. Am., 64, 1363-1367.

Hutchinson, J. A., Zhu, C., Bradley, B. A., Lee, R. L., Wotherspoon, L. M., Dupuis, M., Schill, C., Motha, J., Manea, E. F., & Kaiser, A. E. (2024). The 2023 New Zealand Ground-Motion Database. *Bulletin of the Seismological Society of America*, *114*(1), 291–310. https://doi.org/10.1785/0120230184

Hutchinson, J., Bradley, B. A., Lee, R., Wotherspoon, L., Dupuis, M., Schill, C., Motha, J., Kaiser, A. E., & Manea, E. F. (2022). *2021 New Zealand Ground-Motion Database*. https://doi.org/10.21420/Z20E-5507

Kalkan, E. (2016). An Automatic *P* -Phase Arrival-Time Picker. *Bulletin of the Seismological Society of America*, *106*(3), 971–986. https://doi.org/10.1785/0120150111

Robert R. Youngs, B. S.-J. C. (2008). *NGA Model for Average Horizontal Component of Peak Ground Motion and Response Spectra, PEER Report 2008-09 | Pacific Earthquake Engineering Research Center*. Pacific Earthquake Engineering Research Center.

Thurber, C., & Eberhart-Phillips, D. (2021). *SIMUL codes for local earthquake tomography* [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.5547889

Waldhauser, F. (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/10.1785/0120000006

Wooddell, K. (2018). New earthquake classification for the NGA-subduction project. *Proceedings, 11th National Conference on Earthquake Engineering, Los Angeles, CA, Paper 1275*.

Wooddell, K. E., & Abrahamson, N. A. (2014). Classification of Main Shocks and Aftershocks in the NGA-West2 Database. *Earthquake Spectra*, *30*(3), 1257–1267. https://doi.org/10.1193/071913EQS208M

Wotherspoon, L. M., Kaiser, A. E., Manea, E. F., & Stolte, A. C. (2022). *National Seismic Hazard Model: Site Characterisation Database summary report*. https://doi.org/10.21420/363X-CK83

Zhu, W., & Beroza, G. C. (2018). PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method. *Geophysical Journal International*. https://doi.org/10.1093/gji/ggy423