16th World Conference on Earthquake, 16WCEE 2017
Santiago Chile, January 9th to 13th 2017
Paper N° 2987
Registration Code: S-L1463098482

CHALLENGES AND OPPORTUNITIES IN DEVELOPING A NATIONAL SEISMIC HAZARD AND RISK MODEL WITH OPENQUAKE FOR NEW ZEALAND

N. Horspool[1], E.R. Abbott[2], S. Canessa[3], C. Van Houtte[4], M. Gerstenberger[5], R. Huso[6], A. King[7]

[1] *Earthquake Risk Scientist, GNS Science, New Zealand, n.horspool@gns.cri.nz*
[2] *Seismic Hazard Analyst, GNS Science, New Zealand, e.abbott@gns.cri.nz*
[3] *Software Engineer, GNS Science, New Zealand, s.canessa@gns.cri.nz*
[4] *Seismologist, GNS Science, New Zealand, c.vanhoutte@gns.cri.nz*
[5] *Seismologist, GNS Science, New Zealand, m.gerstenberger@gns.cri.nz*
[6] *Software Engineer, GNS Science, New Zealand, r.huso@gns.cri.nz*
[7] *Risk Engineer, GNS Science, New Zealand, a.king@gns.cri.nz*

Abstract

National seismic hazard and risk models are often developed with 'in-house' software that, while meeting the needs for that particular model, may be limited in functionality. Over time the models can become out-dated and cumbersome to maintain and upgrade. Further, these codes are often not in a suitable state for distribution and collaboration with others. This is common across many countries around the world including New Zealand.

Through the Global Earthquake Model (GEM) initiative, the seismology and engineering community has access to a free and open source seismic hazard and risk engine, OpenQuake, that can be used as the calculation engine for national models. OpenQuake provides flexibility in calculation types and its open source philosophy allows new functionality to be added by users if required.

Recently, the main components of the New Zealand Seismic Hazard and Risk Model have been converted from legacy software formats to the Natural Risk MarkUp Language (NRML) format and incorporated into an end-to-end hazard to risk OpenQuake model for New Zealand. This process has also contributed back to the GEM community by the incorporation of new Ground Motion Prediction Equations (GMPEs) specific to New Zealand into the OpenQuake repository.

This paper will present some of the challenges of implementing the existing seismic hazard model in OpenQuake, provide insight into the process of comparing results from multiple codes, as well as the opportunities of using a state-of-the-art computation engine such as OpenQuake.

Keywords: seismic hazard, risk assessment, OpenQuake, Global Earthquake Model, New Zealand

1. Introduction

The OpenQuake (OQ) engine developed by the Global Earthquake Model (GEM) is a free and open source seismic hazard and loss calculation engine written in the Python programming language [1, 2]. GEM has not only developed a transparent and uniform computation engine for calculating seismic hazard and risk, but has also provided a number of repositories of key input models, such as ground motion prediction equations (GMPEs), scaling laws, and vulnerability models that can be contributed to by the GEM community. Further, GEM has also developed a suite of Global Components, such as a uniform and homogenized global instrumental seismicity catalogue, a global faulted-earth model, a global strain rate model, and global exposure models. These datasets allow users to construct a seismic hazard or risk calculation using OpenQuake for any location around the globe. However, OpenQuake can also be used as the calculation engine where the user supplies their own input models if these are available.

The New Zealand National Seismic Hazard Model (NSHM) uses a bespoke software code written in Fortran 77 that was developed over 20 years ago. The code has been modified by a few researchers familiar with the code, but the core functionality has remained unchanged since it was developed. The code initially started out as what could be classed 'research software' and was then utilized for the 2000 update of the NSHM [4]. Following this it became the software of choice for the New Zealand NSHM and was also used in the 2010 update [5]. The fact that it has been used for over 20 years attests to the huge value of this piece of software, but even the primary author of the code notes that the code was not envisioned to be in use for this long, and would have been superseded much earlier [5]. However, as with most code that evolved from research software there are issues with limited testing coverage, poor version control, hard-coded variables, limited documentation, and many code redundancies. These features mean the software carries significant technical debt, which makes it difficult for new users to modify or for new functionality to be added.

For these reasons the source models and GMPEs used in the NSHM were recently implemented in the OpenQuake engine. This creates the opportunity to run the NSHM in a state-of-the-art seismic hazard and risk calculation engine that provides improved functionality. This will enable new research in seismic hazard assessment to be implemented in the NSHM with less effort than using the legacy code.

This paper outlines the details of the New Zealand NSHM and how it was implemented in OpenQuake, preserving as much as possible the same implementation as that in the legacy NSHM software. Results from the legacy NSHM and the new OQ-NSHM are then compared to provide confidence that OQ-NSHM is giving similar results to the original NSHM, which is important for users of the OQ-NSHM. Finally, the challenges and opportunities presented by implementing the NSHM in OpenQuake are summarized which may be of use for others undertaking a similar task.

2. New Zealand National Seismic Hazard Model

The first NSHM for New Zealand was developed in 1985 [3], and was subsequently updated in 2000 [4] and 2010 [5]. The 1985 model [3] consisted of zonal area sources with no explicitly defined fault sources. The ground motion prediction equation (GMPE) was developed from Japanese data and was compared with a limited New Zealand dataset of strong motion recordings [6]. The 2000 iteration [4] was a major step change as it included over 305 fault sources in the model and a NZ specific GMPE [7]. The most recent iteration of the NSHM [3] now includes 535 fault sources while using the same ground motion prediction equation as the 2000 model [8].

The New Zealand NSHM has three main components. The background seismicity source model, the active fault source model, and the GMPEs. The details of these and how they are implemented in OpenQuake are outlined in the following sections.

2.1 Background Sources

The background seismicity source model accounts for small to moderate magnitude events that occur off the known mapped faults (Section 2.2). The background seismicity model in the NSHM is modeled as point sources on a 0.1° x 0.1° grid with depth layers at 10, 30, 50, 70 and 90 km that has been created using a spatial smoothing technique of historical earthquake catalogue data using a Gaussian Kernel filter to estimate the a-value for each grid point. The b-value is calculated for large homogenous seismo-tectonic zones and the estimated b-value is then applied uniformly across all grid points in the zone [5]. The NSHM background seismicity source model file consists of a list of grid points with the b-value, faulting type, maximum magnitude (M_{max}), and the coordinates of the grid point. The activity rates are expressed as the number of earthquakes above a given magnitude for three different time periods of completeness. The header information contains the time period (number of years prior to 2010) of completeness for M4.0, M5.0 and M6.5. These source parameters are then parsed into NRML format for use in OpenQuake. The following describes how the parameters are mapped to the appropriate OpenQuake inputs.

First, the fault sense type of the background sources need to be converted from classes such as 'reverse', 'normal', and 'reverse-strike slip' that are used in the NSHM to the equivalent rake, which is used by OpenQuake. This mapping is shown in Table 1.

Table 1 – Fault sense to rake conversions

NSHM Fault Sense Type	OQ-NSHM Fault Rake
RV, IF	90
NN, NV, NS, DS	-90
SS, SN, SR	0
RS, RO	45

Second, the activity rates of each grid need to be calculated and represented as a Magnitude Frequency Distribution (MFD) in OpenQuake. The NSHM uses a truncated Gutenberg Richter distribution [5] for background seismicity. OpenQuake requires this to be defined by an a-value, b-value, minimum magnitude (M_{min}) and a maximum magnitude (M_{max}). The M_{min} is set to 5.25 for the NSHM, and M_{max} and b-value is taken from the NSHM input file. M_{max} for background sources is M6.5 in the Taupo Volcanic Zone and M7.2 elsewhere. To calculate the a-value at M0 we need to evaluate the activity rate defined as the number of events greater than magnitude, *i*, (N_i) with an associated time periods of completeness ($T_{i,}$) which are parameters supplied in the NSHM input file. This is done using the following;

$$a_{M0} = log_{10}\left(\frac{N_{M4.0} + N_{M5.0} + N_{M6.5}}{T_{M4.0} + T_{M5.0} + T_{M6.5}}\right)$$

where,

$$T_i = cTime \times 10^{(-M_i b)}$$

and *cTime* is the time period of completeness of magnitude M_i in years If the a-value at M0 is less than 0.0008 a floor level of 0.0008 is assigned. The number of M>4 earthquakes is shown Fig. 1.

2.2 Fault Sources

The 2010 NSHM contains 535 active fault sources [5]. Each fault is required to be characterized in OpenQuake by a geometry, such as surface fault trace coordinates that must adhere to the right-hand rule (fault dipping to the right when looking from the first to last point on the trace), fault dip, fault rake, and upper and lower seismogenic depth. It also requires a magnitude-frequency distribution (MFD). The NSHM fault source model file, contains the surface coordinates of the upper seismogenic zone (not surface fault trace), which may or may not adhere to the right-hand rule, the depth of the upper and lower seismogenic zone, fault dip, dip direction, faulting type, the characteristic magnitude of the fault, and average recurrence interval in years of this magnitude. This MFD can be considered a delta function with no uncertainty in magnitude. To convert the NSHM to OpenQuake, first the geometry is defined. This requires the coordinates of the upper seismogenic zone of the fault to be reprojected to a surface fault trace (i.e. at zero depth). For faults with zero seismogenic depth (i.e. suface rupturing) these coordinates are the same. Next, the coordinates need to be reordered to the right-hand rule. The fault sense is also converted to rake as outlined in Tab. 1. The recurrence model is a single annual rate, and is converted to annual rate from the recurrence interval in years. All faults are modeled as 'simple fault sources' in OpenQuake. The fault sources are shown in Fig. 1.

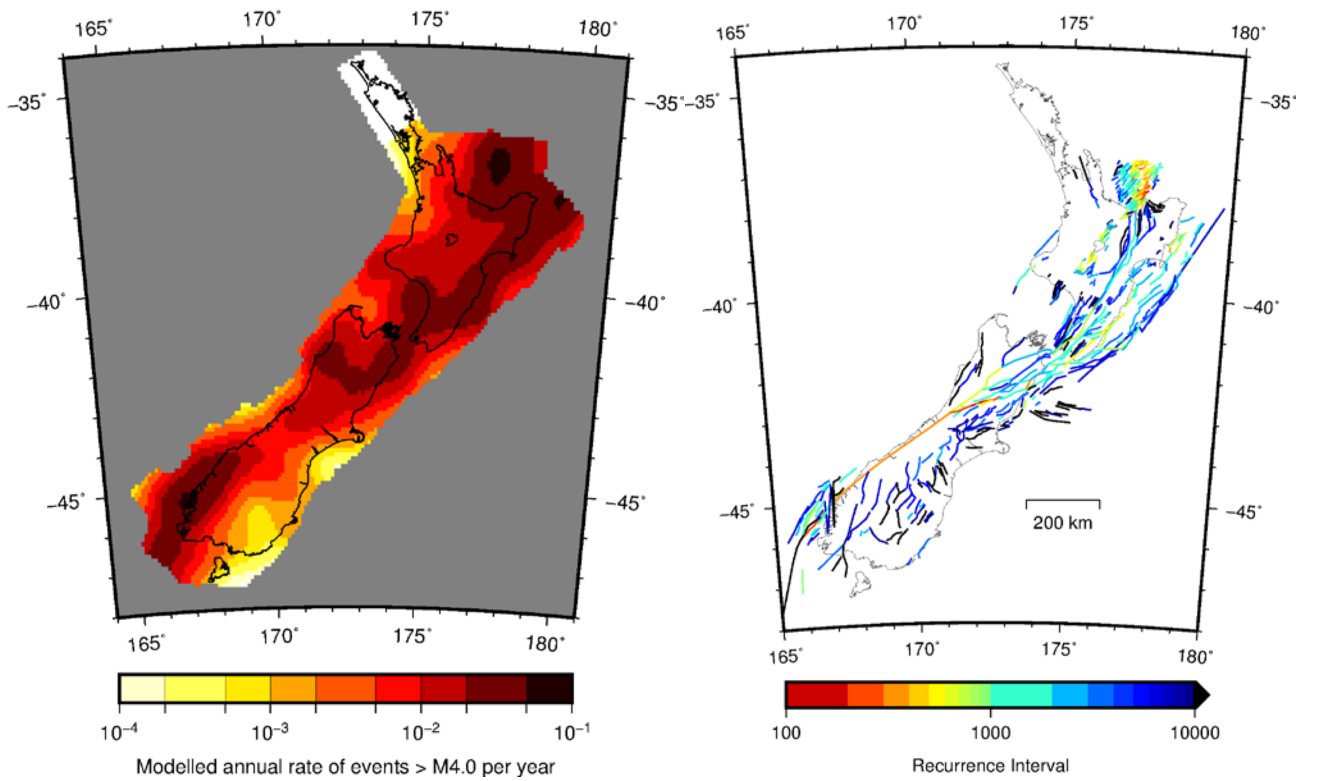


Fig 1 - Seismic source models used in the NZ Seismic Hazard Model. Left: background seismicity model showing the annual rate of M>4.0 earthquakes per year. Right: fault sources with their average recurrence interval in years.

2.3 Ground Motion Prediction Equations (GMPEs)

In the past, New Zealand has lacked a suitable amount of strong ground motion data to derive a New Zealand GMPE. Instead, the functional form of global models have been used as the basis for a New Zealand GMPE [6,7,8,9], and the model coefficients have been updated using the relatively small New Zealand strong motion dataset. The New Zealand NSHM uses a single GMPE model [8] referred to here as McVerry2006 that has different sub-models for active shallow crust, subduction interface, subduction intraslab, and earthquakes with a volcanic source-to-site path. Table 2 shows how these sub-models are mapped to different seismic sources in the OQ-NSHM. The NSHM GMPE model [8] required some modifications of the variables to be compatible for use in OpenQuake. First, the McVerry2006 model has three terms for faulting type; normal faulting (CN), reverse faulting (CR) and oblique reverse (CR). These have been mapped to rake as follows; CN = -1 for -146 < rake < -33, CR = 1 for 67 < rake < 123, and CR = 0.5 for 33 < rake < 66 where the rakes are converted from the original fault sense types in the seismic source model as defined in Table 1. The McVerry2006 model uses the New Zealand Site Class definitions to apply a soil term. These terms defined as δ_C and δ_D define whether the site is Site Class B (rock), Class C (shallow soil) or Class D (deep soil). Note the McVerry2006 model does not have a term for Class E. In the OpenQuake implementation of McVerry2006, $\delta_C = 1$ if 360 <= Vs30 < 760 and $\delta_D = 1$ if Vs30 < 360. Otherwise these coefficents are zero. The McVerry2006 model also applies a volcanic path term (r_{vol}) that adds additional attenuation of ground motions if the path travels through the Taupo Volcanic Zone [8]. In the published description of the McVerry2006 model [8] the term r_{vol} is equal to the three dimensional distance of the source-to-site path within the Taupo Volcanic Zone. However in the implementation of the McVerry2006 model in the NSHM [5] is slightly different than intended in the original paper. In the NSHM for any source that is considered volcanic (i.e. that is located inside the Taupo Volcanic Zone, Table 2), r_{vol} is equal to r_{rup}.

Table 2 – Ground Motion Models

OpenQuake GMPE Class Name	**Tectonic Region Type and Notes**	**NSHM Source Code**
McVerry2006Asc	Active Shallow Crust: All crustal faults and background sources with a depth less than 30 km.	All others beside NV, IF, DS
McVerry2006Volc	Volcanic: All background and fault sources inside the Taupo Volcanic Zone	NV
McVerry2006SInter	Subduction Interface	IF
McVerry2006SSlab	Subduction Intraslab	DS, or if no code is given and depth is greater than 30 km.

A recent version of the NSHM that accounts for time-dependence in Canterbury caused by the Canterbury Earthquake Sequence [9] adopts an additional GMPE [10,11] for seismic hazard analysis in that region. This model, referred to as Bradley2010 [10,11], has also been implemented within OpenQuake and is available in the GMPE library.

2.4 PSHA Calculation Parameters

To reproduce the New Zealand NSHM a number of parameters need to be configured in the control file to compare to the NSHM results. These are outlined in Table 3.

Table 3 – OQ-NSHM Configuration Parameters

Parameters	**Value**
Site Spacing	0.1° x 0.1°
GMPE Maximum Distance	400 km
GMPE Uncertainty Truncation	3σ
Rupture Mesh Spacing	1 km
Reference Vs30 Value	560 m/s (NZS1170.5 site class C)

Investigation Time	50 years

4. Model Comparison

The results from NSHM and OQ-NSHM are compared to ensure the OQ-NSHM implementation is producing similar results. These are presented as hazard curves and hazard maps.

4.1 Hazard Maps

Hazard maps for peak ground acceleration (PGA) for a 10% probability of exceedence in 50 years on NZS1170.5 site class C (shallow soil, Vs_{30}=560 m/s) are shown in Fig. 2. The maps are nearly identical apart form discrepancies in the low hazard areas in northern New Zealand and south eastern New Zealand. The mismatch between the two maps is quantified using a hazard map error (HME) term, which is defined as (NSHM – OQ) / NSHM. A HME of 1.0 is a 100% underestimation by OpenQuake model, and -1.0 is a 100% overestimation by OpenQuake. Fig. 3 shows the hazard map error for two maps in Fig. 2.

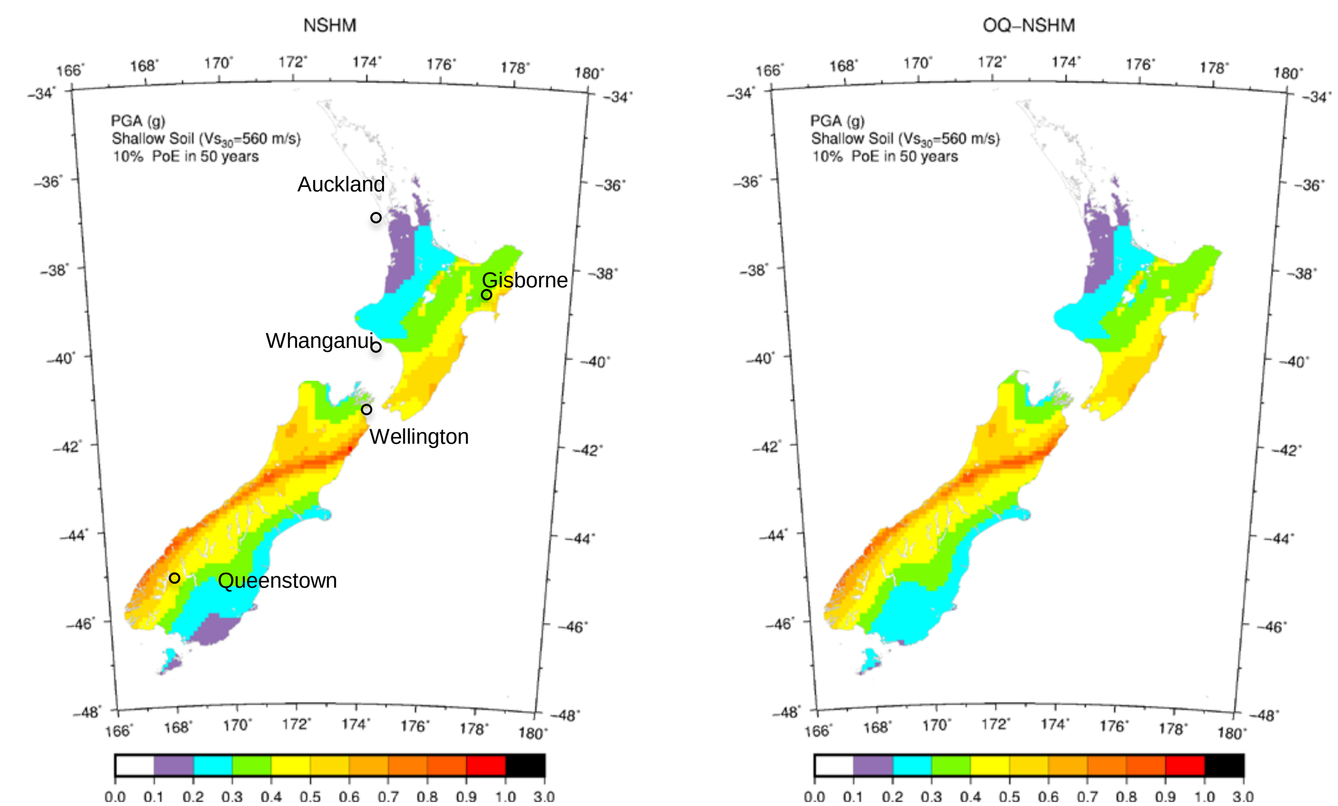
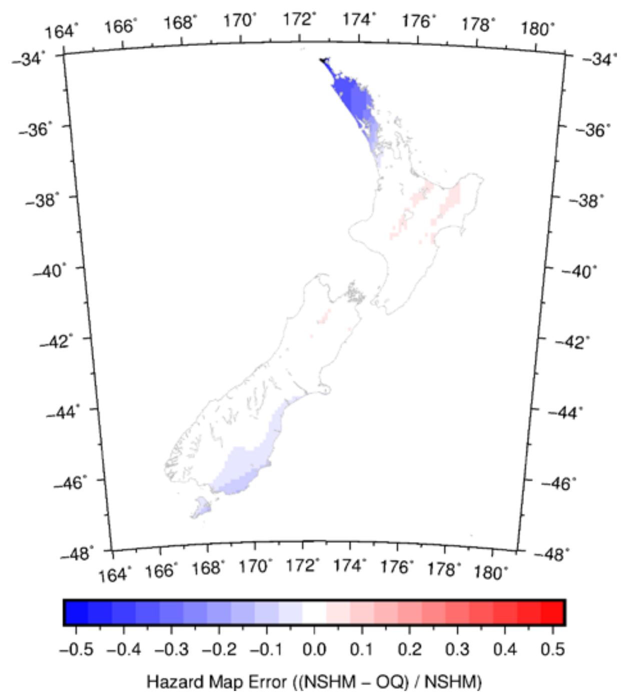


Fig. 2 – Hazard map comparison for peak ground acceleration 10% probability of exceedence in 50 years, on soft soil (NZS1170.5 Site Class C / Vs_{30} 560 m/s) using the original NSHM code (left) and the OpenQuake implementation of the NSHM (right).

The HME is less than 2.5% (white areas in Fig. 3) for most of New Zealand including all the major cities. However in the low seismic hazard areas OpenQuake is overestimating by over 5-40% (Fig 3 – right). This is thought to be due to how the floor level of seismicity are handled and is being investigated.



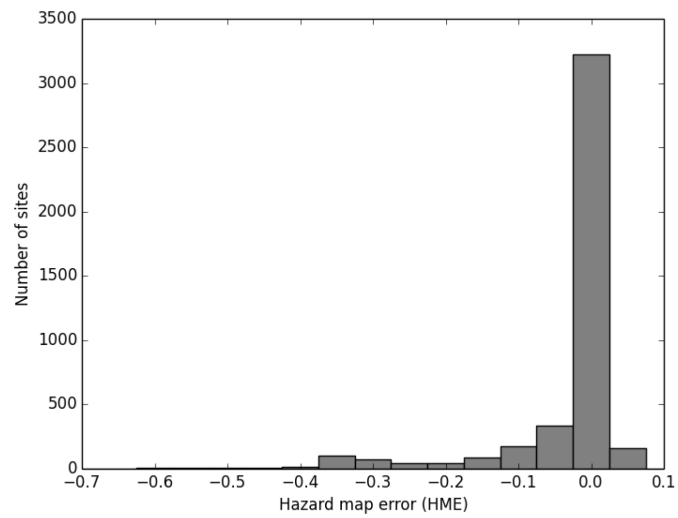


Fig. 3 – Hazard map error (HME) represented as a map (left) and histogram of sites (right) for peak ground acceleration 10% probability of exceedence in 50 years, on soft soil (NZS1170.5 Site Class C / Vs_{30} 560 m/s). A positive value indicates NSHM is higher, and a negative value indicates OpenQuake is higher.

4.2 Hazard Curves

The hazard curves generated by each model are compared in Fig 4 to Fig 8 for a selection of cities around New Zealand on NZS1170.5 site class C (Vs_{30} = 560 m/s). The cities have been selected to encompass a range of hazard levels and dominant source types. For example, Gisborne and Wellington are located within 20-25 km above the subduction interface and are both also dominated by other nearby large crustal faults. Queenstown is dominated by background seismicity but also has the Alpine Fault nearby, whereas Whanganui and Auckland are dominated primarily by background sources and are moderate to low hazard levels respectively. Refer to Fig. 2 for locations of the cities.

The comparison results show a good match within 5% error for most cities. There is a slight mismatch at the low annual probability of exceedence (~10^{-3} to 10^{-4}) and this is systematic across most cities. It is unclear what is causing this discrepancy, but it is thought it could be due to slight differences in numerical integration of the probability density functions in the PSHA calculation between the two codes, which is highlighted when integrating very small probabilities. Auckland, which has the lowest hazard level of all the cities referred to here, also shows a mismatch at the high POE, this again could be a numerical integration error, or to do with sensitivities around how the floor level is applied the two models. This issue is still being explored at the time of writing.

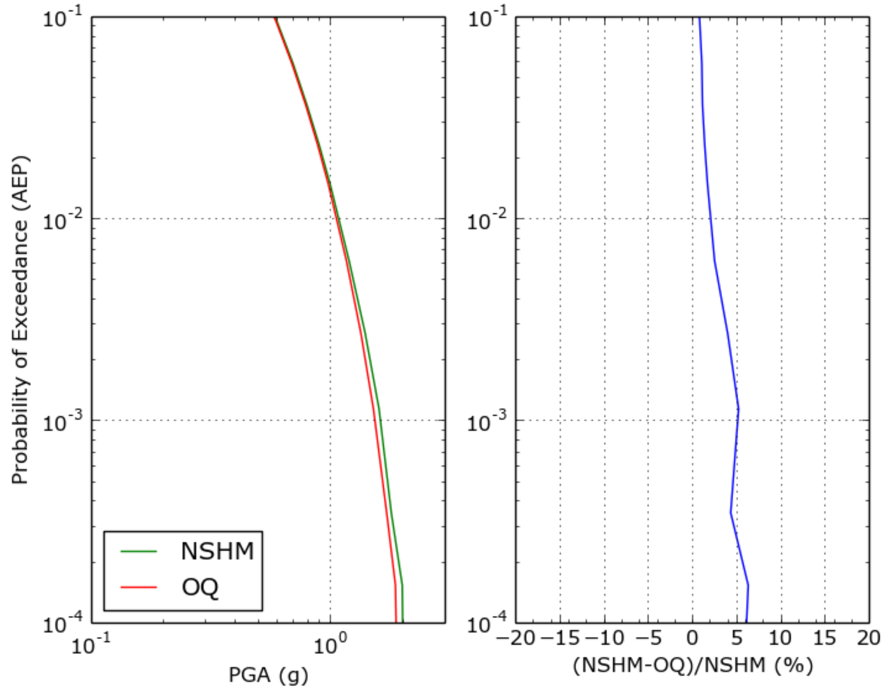


Fig. 4 – Hazard curve comparison for Gisborne

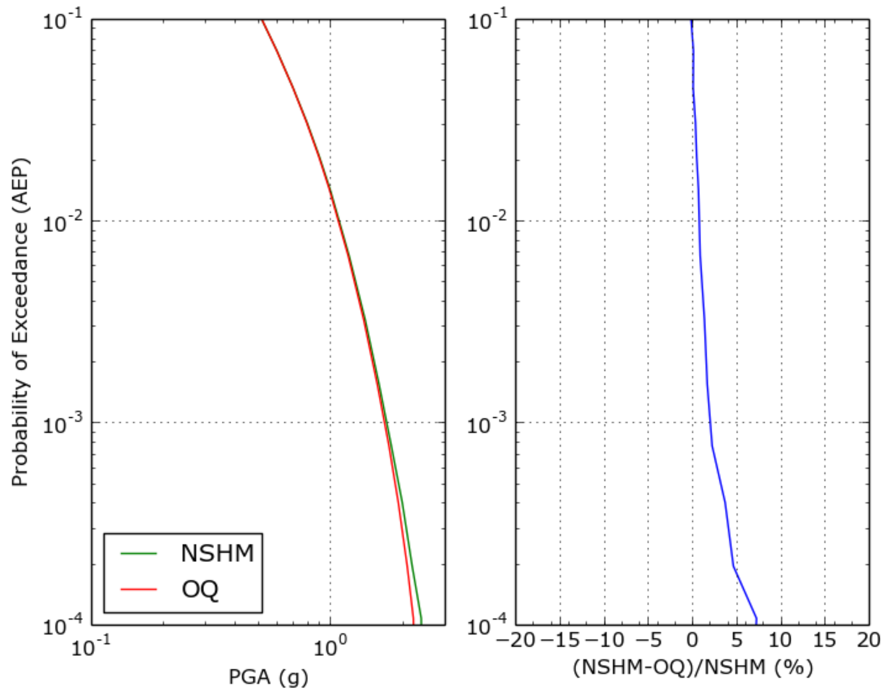


Fig. 5 – Hazard curve comparison for Wellington

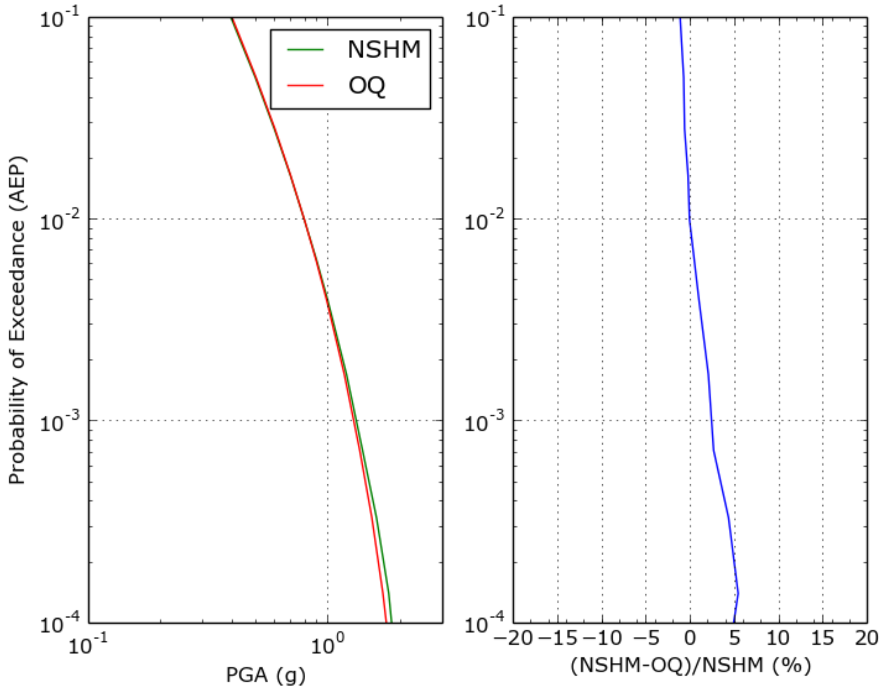


Fig. 6 – Hazard curve comparison for Queenstown

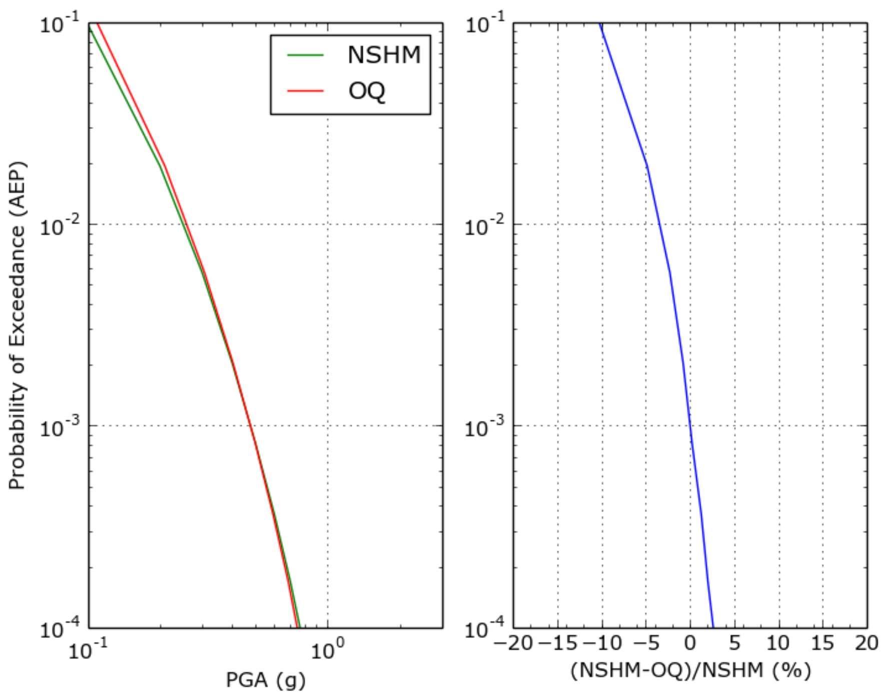


Fig. 7 – Hazard curve comparison for Auckland

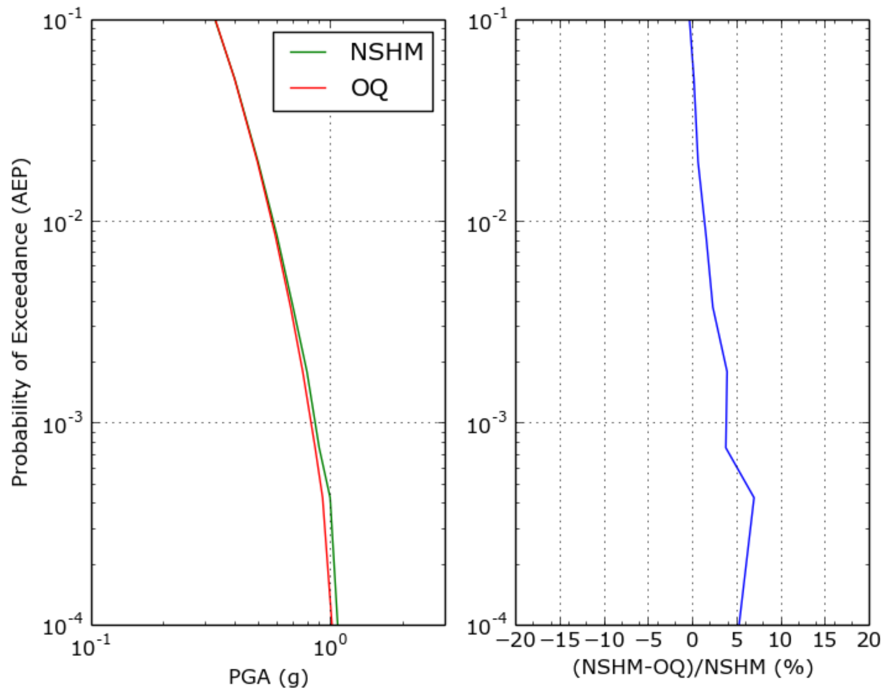


Fig. 8 – Hazard curve comparison for Whanganui

5. Challenges and Opportunities in using OpenQuake

The process of implementing the NSHM into OpenQuake and calibrating the results has taken hundreds of hours. The majority of this time has been reconciling differences between the results, which involves an iterative manual process of 1) comparing results, 2) attempting to identify an potential issue, 3) reconfiguring the models, 4) rerunning the code, 5) comparing results and repeating. Through this process the most effective way of calibrating the models has been by breaking down each model into the smallest components. For example, the models were run with background only, and fault only sources then compared for all the cities. When there were discrepancies at a given city, the dominant source (i.e. a given fault source) was isolated through deaggregation and the calculation was then repeated for a single source at a time. This process helped identify discrepancies between how some fault source types (e.g. subduction interface) were being handled by the different models and were then easily corrected. However, this process is very time consuming and hence the significant amount of time that has been spent on calibrating the OQ-NSHM model.

Other discrepancies that were identified and may be relevant for others undertaking a similar process include:

- Numerical precision: some small errors were introduced due to the original NSHM code being implemented with 32 bit precision where as OpenQuake is 64 bit. When numbers are rounded by 32 bit precision then carried through further calculations, the error can compound and then become significant;
- Numerical integration: OpenQuake uses discrete PDFs of variables where as NSHM uses continuous integration. This can lead to some small errors depending on the method used;
- Errors introduced during conversion of parameters: OpenQuake requires defined parameters such as the way the fault geometries are defined as being the surface trace, whereas NSHM fault sources are defined as the location of the upper seismogenic zone of the fault. which for non-surface rupturing fault sources, requires conversion to OpenQuake surface fault traces, potentially introducing some error.

As well as presenting a number of challenges, using OpenQuake also creates opportunities for increased functionality. The NSHM was designed around a single GMPE and single source model. This was hard-coded and has precluded the ability to include epistemic uncertainty in these PSHA components without major refactoring of the legacy code. By using OpenQuake this functionality is easily applied, which is important as it is now becoming a requirement for many site-specific seismic hazard studies of important sites, such as dams, in New Zealand and elsewhere. Further, through OpenQuake a large repository of GMPEs are available for use. While this project has contributed to this repository by adding the two New Zealand specific GMPEs [8,10], it also allows the recent NGA models to be included in logic trees, which have been shown to be a good fit to New Zealand strong motion data [12]. Another benefit of using OpenQuake for the NSHM is for reproducibility of results. The NSHM source models are publically available for download[1], however, how they are implemented as described in Section 2 is not provided beyond what is available in the original paper [5]. By implementing the model in OpenQuake, the entire OpenQuake project including the source models, GMPEs, logic tree structures (if implemented in future models), calculation configurations, and site models can be provided allowing entirely reproducible results, which means less burden on the user of the model since they will not having to undertake a similar process as described here. Finally, the ultimate goal of the New Zealand model is to have an end-to-end hazard-to-risk national model. This will allow consistency between seismic hazard and risk calculations. The risk components of the RiskScape loss model [13] have also been converted to OpenQuake format and have been used for some initial risk calculations [14]. Following completion of the hazard model calibration, national scale risk modeling in OpenQuake will be feasible.

6. Conclusions

The New Zealand National Seismic Hazard Model has been ported from a legacy software code and implemented in OpenQuake. This process has been challenging, in terms of calibrating the results due to a number of reasons discussed in this paper, but on the other hand it has created many opportunities for increased functionality of the seismic hazard model.

7. Acknowledgements

Macro Pagani, Vitor Silva and the OpenQuake development team are thanked for their support in implementing the New Zealand OpenQuake model.

8. References

[1] Pagani, M., Monelli, D., Weatherill, G., Danciu, L., Crowley, H., Silva, V., Henshaw, P., Butler, L., Nastasi, M., and Panzeri, L. (2014). OpenQuake engine: An open hazard (and risk) software for the global earthquake model, Seismol. Res. Lett. 85, no. 3, 692–702, doi: 10.1785/0220130087.

[2] Silva V., Crowley H., Pagani M., Monelli D., Pinho R. (2013). Development of the OpenQuake engine, the Global Earthquake Model's open-source software for seismic risk assessment, Natural Hazards, Vol. 66.

[3] Matuschka, T., Berryman, K.R., O'Leary, A.J., McVerry, G.H., Mulholland, W.M. and Skinner, R.I. 1985. New Zealand seismic hazard analysis. Bulletin of the New Zealand National Society for Earthquake Engineering, Vol 18: 313-322.

[4] Stirling M.W., McVerry G.H. & Berryman K.R. (2002). A new seismic hazard model for New Zealand. Bulletin of the Seismological Society of America 92,1878-1903.

[5] Stirling, M. W., McVerry, G. H., Gerstenberger, M. C., Litchfield, N. J., Van Dissen, R. J., Berryman, K. R., Barnes, P., Beavan, R. J., Bradley, B., Clark, K. J., Jacobs, K., Lamarche, G., Langridge, R. M., Nicol, A., Nodder, S., Pettinga, J., Reyners, M. E., Rhoades, D. A., Smith, W. D., Villamor, P., Wallace, L. M. (2012). National seismic hazard model for New Zealand: 2010 update. Bulletin of the Seismological Society of America, 102, 1514–1542.

[6] Katayama, T. (1982) "An Engineering Prediction Model of Acceleration Response Spectra and its Application to Seismic Hazard Mapping", Earthquake Engineering and Structural Design, 10, pp. 149-163.

[7] McVerry, G., Zhao, J., Abrahamson, N. and Somerville, P. (2000). "Crustal and subduction zone attenuation relations for New Zealand earthquakes". In Proceedings of Twelfth World Conference on Earthquake Engineering, Paper Number 1834, Auckland, New Zealand, 30 January - 4 February 2000

[1]http://www.gns.cri.nz/Home/Our-Science/Natural-Hazards/Earthquakes/Earthquake-hazard-modelling/National-Seismic-Hazard-Model

[8] McVerry, G. H., J. X. Zhao, N. A. Abrahamson, and P. G. Somerville (2006). Response spectral attenuation relations for crustal and subduction zone earthquakes, Bull. New Zeal. Soc. Earthquake Eng. 39, 1–58.

[9] Gerstenberger, M.C.; McVerry, G.H.; Rhoades, D.A.; Stirling, M.W. 2014. Seismic hazard modeling for the recovery of Christchurch, New Zealand. *Earthquake Spectra, 30(1): 17-29; doi: 10.1193/021913EQS037M*

[10] Bradley, BA. (2013). A New Zealand-specific pseudo-spectral acceleration ground-motion prediction equation for active shallow crustal earthquakes based on foreign models. Bulletin of the Seismological Society of America. Vol. 103(3), 1801-1822

[11] Bradley, B. (2015). "Systematic ground motion observations in the Canterbury earthquakes and region-specific non-ergodic empirical ground motion modeling". Earthquake Spectra, 31(3), 1735-1761

[12] Van Houtte, C., Performance of response spectral ground motion models against New Zealand data, *Bulletin of the New Zealand Society for Earthquake Engineering*, submitted.

[13] King, A.B., Bell, R.G. (2009). RiskScape Project: 2004 - 2008. NIWA Science Report 2009/75. 172p.

[14] Horspool, N., King, A.B., (2015). What drives seismic risk in New Zealand? Insights from a next generation loss model. Proceedings of the Pacific Conference on Earthquake Engineering, November 2015, Sydney, Australia.