SECED 2023 Conference
Earthquake Engineering & Dynamics for a Sustainable Future
14-15 September 2023, Cambridge, UK

MODELLING DIRECTIVITY IN PROBABILISTIC SEISMIC HAZARD ANALYSIS: EXAMPLES FROM NEW ZEALAND AND APPLICATION TO COMPLEX FAULT RUPTURES

Graeme WEATHERILL[1] & Henning LILIENKAMP[2]

***Abstract**: Near-fault ground motions from large magnitude earthquakes present a particular challenge to seismic hazard modellers and earthquake engineers. Pulse-like motions from rupture directivity have been seen in many earthquakes, leading to amplification at longer periods and can cause increasing damage in mid- to high-rise structures. Despite this, its explicit characterization of directivity in PSHA creates additional challenges and application in regional or national scale seismic hazard models has not yet become established practice. The update to the national seismic hazard model of New Zealand provided an opportunity to explore the feasibility of implementing directivity into national scale PSHA and to understand its implications for engineering design. The existing 2012 national seismic hazard model contains over 500 well-mapped shallow crustal faults, while the national seismic design code (NZS1170.5) contains near-fault factor, $N(T, D)$, to increase the design acceleration close to fast slipping faults. We implement state-of-the art directivity models into the OpenQuake calculation software and adapt the seismic hazard framework to incorporate aleatory variability in rupture hypocentre location. With this we can understand impact of incorporating directivity into PSHA for New Zealand and compare the changes in hazard with the current NZS1170.5 near-fault factor. For the 2022 national seismic hazard model, the adoption of a fault system modelling approach adds complications to directivity modelling, however, as fault ruptures are represented by geometrically complex multi-segment and multi-fault surfaces. As current directivity models are not well suited to such fault ruptures, we explore the application of a convolutional neural network to adapt existing directivity models to this more complex framework. The increasing usage of connected fault systems as seismogenic sources highlights the need for future directivity models to capture the potential range of complex multi-fault surfaces that can emerge and to find innovative and efficient approaches to implement them in PSHA.*

Introduction

Strong ground motions at sites near an earthquake rupture can present a serious threat to buildings in these regions, even those with some degree of seismic resistant design. These near-fault ground motions are often not only large in amplitude, owing to their close proximity to the seismic source, but may also have a clear pulse-like character under certain commonly observed conditions. These pulse-like characteristics are commonly attributed to *directivity,* which results at a site when a fault rupture propagates towards the site in question at velocities near (or equal) to that of the shearwave velocity. Under these conditions the wavefront arrives as an impulsive motion, that is usually strongest in the fault-normal direction of wave motion (Sommerville et al., 1997; Baker, 2007). Pulse-like motion can induce especially strong seismic demands in structures whose fundamental periods are close to the period of the pulse (T_p), which can be especially damaging for engineered tall structures or structures with base isolation (Hall et al., 1995).

Given their importance for seismic design of engineered structures, it would be desirably that effects of rupture directivity are accounted for in the characterisation of ground motions used as inputs to seismic design. Probabilistic Seismic Hazard Analysis (PSHA) is our standard and most effective tool to define the level of seismic shaking (including amplitude and spectral content) that a site may be subject to within a time T, yet *explicit* modelling of directivity in PSHA introduces challenges that have limited its application in practice (Donahue et al., 2019). Among these are the computational challenges associated with integrating variability in the parameters controlling directivity (described in the following section), combined with the limited numbers of ground

[1]Doctor, Seismic Hazard Researcher, GFZ German Research Centre for Geosciences, Potsdam, Germany, graeme.weatherill@gfz-potsdam.de
[2] Doctoral Candidate, GFZ German Research Centre for Geosciences, Potsdam, Germany

motion observations and available models of near-fault ground motion characterisation, which contribute to large epistemic uncertainty in the scale and spatial extent of the amplification resulting from this phenomenon. Nevertheless, modern PSHA applications in regions of moderate to high seismicity regularly contain models of active faults whose three-dimensional geometry and models of earthquake recurrence are well constrained by observed earthquake geology. This allows hazard modellers to identify regions where directivity may present a particular threat to a structure (or structures), meaning that the potential for this to occur should be addressed. How we do so is not simple, however.

New Zealand is one such region where near-fault ground motion amplification effects may have the potential to exacerbate the earthquake hazard and risk. Sandwiched by active subduction zones in the north and south of the country, New Zealand is traversed by numerous shallow faults, many of which are located close to populated regions and have especially high seismogenic potential. More than 9,500 active fault traces have been mapped across the country and 712 fault segments fully parameterised for use in PSHA (Stirling et al., 2012; Gerstenberger et al., 2022a, b). The relevance of near-fault directivity amplification effects for engineered structures has been long recognised in New Zealand and was observed in recent earthquakes such as the 2010 Darfield event. The current seismic design loading standard for New Zealand (NZS 1170.5) is the only such standard worldwide to incorporate an explicit spectral period (T) dependent and distance (D) dependent near-fault factor $N(D,T)$, to be applied to increase the design load for sites near known faults with high seismogenic potential. This near-fault factor is calibrated based on an early empirical directivity model of Somerville et al. (1997) and is applied deterministically to increase the design ground motion (itself determined from PSHA).

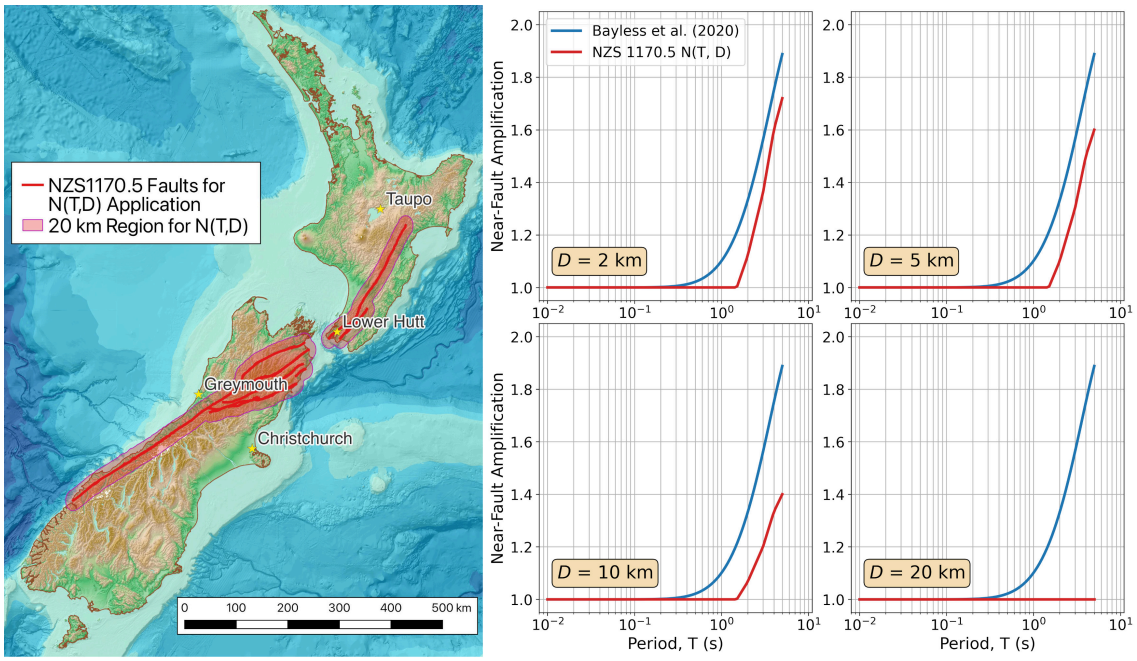

Figure 1. NZS1170.5 Near-Fault Factor $N(T,D)$ region of application (left) and scaling with period, T, and distance to fault, D (right).

In 2022 a new update to the National Seismic Hazard Model (NSHM) was published (Gerstenberger et al., 2022a), which introduced a radical change of modelling approach compared to its 2012 predecessor (Stirling et al., 2012) by adopting an *earthquake rupture forecast* (*ERF*) approach to seismogenic source characterisation, taking the innovative fault system modelling framework introduced for the Uniform California Earthquake Rupture Forecast (UCERF) v3 by Field et al. (2014) and adapting it to the New Zealand context. ERF-based source characterisation aims to reconcile observations from seismicity, geology and geodesy into a model of the full description (or inventory) of ruptures and their probabilities of occurrence across a complete fault-system, constrained by established fault rupture physics. These models present new challenges for ground motion characterisation in PSHA, however, as they produce ruptures of greater geometric complexity than previous models, including multi-segment and multi-fault

case that extrapolate beyond the types of ruptures for which current models of near-fault ground motion were originally designed.

We explore the feasibility and impact of incorporating rupture directivity explicitly into PSHA for New Zealand. Initially our work focuses on application to the 2012 NSHM (Stirling et al., 2012), which is used as a test bed application for integration of explicit directivity modelling in PSHA at a national scale. This will demonstrate implementation process and the potential impacts on ground motion for seismic design in New Zealand. We then turn our attention to the 2022 NSHM and the challenges presented by the ERF source models. Here, the more conventional directivity modelling approaches have not proven feasible to implement at scale, which has motivated us to seek an alternative strategy using machine learning for efficient prediction of spatial patterns of near-fault amplification calibrated on existing empirical directivity amplification models. Though explicit directivity modelling was not ultimately integrated into the 2022 New Zealand NSHM, the framework we present here has the potential to facilitate efficient implementation of directivity in future seismic hazard models, particularly those that adopt the ERF framework.

Implementing Explicit Directivity Models in PSHA

The conditions for rupture directivity to manifest as strongly impulsive motions depend on both the proximity of the site to a rupture and the alignment with respect to the direction of propagation. Directivity effects are strongest when sites are located close to the ends of the ruptures, and the strength itself will depend on the distance travelled by the rupture front between its point of nucleation (the hypocentre) and the termination of the rupture, with longer distances producing stronger pulses. Therefore, not only is a detailed characterisation of fault geometry critical for modelling this phenomenon in PSHA, but one must also take into account the uncertainty in the the position of the hypocenter in the rupture plane, and in some cases also the variability in the local direction of rupture slip. Both properties are aleatory processes that differ from rupture to rupture. As near-fault pulse-like ground motions are present within the databases of observed ground motions used to calibrate a ground motion model (GMM) directivity can be said to be *implicitly* captured within its total aleatory variability σ_T of the predicted ground motion. Considering response spectral period T at sites with a distance R from an earthquake of magnitude M, given additional source, path and site constraints θ, directivity amplification can vary significantly among different sites that are otherwise equidistant from the fault. Implicit modelling within σ_T would not reflect this difference and the variability at a site would reflect that of both the pulse-like and non-pulse-like records across all ground motions for similar scenarios. We therefore talk of *explicit* directivity modelling as that in which the expected ground motion $\mu(M,R,\theta,T)$ from an empirical GMM and its aleatory variability $\sigma_T(M,R,\theta,T)$ is modified by a factor $\Delta\mu_{dir}(M,R,\theta,T)$ and $\Delta\sigma_{T,dir}(M,R,\theta,T)$ respectively, where the modifications reflect the change in ground motion and its variability *with respect to a directivity neutral condition*.

Directivity amplification is dependent not only upon the fault rupture and the site location but also on relative position of the hypocentre position within the fault ($h = [X,Z]$, where X and Z are the relative along-strike hypocenter and down-dip hypocentre positions respectively), and the local direction of slip, γ. Both are aleatory variables described by probability distributions $f_H(h)$ and $f_\gamma(\gamma)$ respectively. For the explicit incorporation of directivity into PSHA we can consider two alternative approaches, which are outlined by Donohue et al. (2019): 1) *full hypocentre randomisation* and 2) *modified moments from a randomised hypocentre.* Full hypocentre randomisation is the natural extension to the PSHA integral to incorporate the additional aleatory variabilities introduced by $f_H(h)$ and $f_\gamma(\gamma)$. This is presented here in the form of the *earthquake rupture forecast* formulation of the PSHA integral (Field et al., 2003; Pagani et al., 2014):

$$P(a \geq A|T) = 1 - \prod_{i=1}^{N_{SRC}} \prod_{j=1}^{N_{RUP,i}} \prod_{k=1}^{N_h} \prod_{l=1}^{N_\gamma} (1 - P(\gamma_l) \cdot P(h_k) \cdot P_{rup,ij}(n \geq 1|\mathrm{T})) \, {}^{P(a \geq A|rup_{ij},h_k,\gamma_l)} \quad (1)$$

Where $P(a \geq A \vee T)$ is the probability of ground motion a at a site exceeding a given threshold A in T years, $P_{rup,ij}(n \geq 1 \vee T)$ is the probability of one or more occurrences in T years of rupture j from seismogenic source i, $P(a \geq A|rup_{ij},h_k,\gamma_l)$ is the probability of ground motion a at the site exceeding A given the occurrence of rup_{ij} with hypocentre position h_k and local slip angle γ_l (i.e. the directivity adjusted GMM), $P(h_k)$ the probability of occurrence hypocentre position h_k from N_h

hypocenter positions and $P(\gamma_l)$ the probability of occurrence of local slip angle γ_l from N_γ possible slip angles. Collectively, this describes the probability of exceedance of ground motion from all the possible ruptures in the source model *and* all hypocentre positions and slip angles for each rupture from their distributions $f_H(h)$ and $f_\gamma(\gamma)$.

The *modifier of moments from a randomised hypocentre* approach is an alternative in which only the GMM is modified by defining distributions of $\Delta\mu_{dir}$ and σ_{dir} calibrated outside of the hazard calculation by modelling the expected amplification and its variability due to hypocentre position and slip angle from a large number of rupture and site configurations (e.g. Watson-Lamprey, 2018). These terms modify the mean and total aleatory variability of the ground motion model using often simpler parametric models, removing the need to integrate over $P(\gamma_l)$ and $P(h_k)$ in the PSHA calculation. The result is a more efficient seismic hazard calculation that can still capture the spatial patterns in ground motion amplification from directivity and its associated variability, albeit with a potential loss of accuracy owing to the parametric representation of $\Delta\mu_{dir}$ and σ_{dir}. In addition, $\Delta\mu_{dir}$ and σ_{dir} are conditional upon the directivity amplification model used and the assumed hypocentre and local slip angle distributions, requiring that the GMM is adapted for each of the considered combinations of directivity model and hypocentre distribution.

Application to New Zealand using the 2012 NSHM

Description of the Seismic Hazard Input Model

In the first stage of this work, we use the 2012 New Zealand NSHM (Stirling et al., 2012) as the basis for implementing directivity in PSHA at a regional scale using the *full hypocentre randomization* approach. The aim of this implementation is to assess the feasibility of explicit directivity modelling on a realistic national seismic hazard model for New Zealand and to identify where and by how much the hazard might be expected to change for the inclusion of directivity in PSHA. A complete description of the implementation and model development is provided in Weatherill (2022). Stirling et al. (2012) is particularly suited to such a test case application as the model incorporates the active shallow faults as single- or multi-segment fault surfaces, each represented by a single plane or connected set of planes and associated to a single characteristic magnitude M_{char} with a fixed annual rate of occurrence. The model contains 530 ruptures from shallow crustal and volcanic fault sources, and a further 11 ruptures from subduction interface sources. Off-fault shallow seismicity and deep subduction in-slab seismicity are modelled as gridded background sources, for which no explicit fault characterization is available. Directivity is modelled explicitly only for shallow crustal and volcanic fault sources, but not for the subduction or background sources. The complete set of fault sources for all three tectonic region types (active shallow crust, volcanic and subduction interface) is shown in the left-hand side of Figure 2.

For the selection of the ground motion model, we are confronted with a limitation in the 2012 NSHM which is the exclusive adoption of McVerry et al. (2006). In contrast to the current generation of GMMs and empirical directivity models, McVerry et al. (2006) models ground motions for spectral periods $T \leq 3$ s and assumes and adopts the envelope of the two horizontal response spectra as the definition of horizontal motion. We also lack sufficient information to determine whether the McVerry et al. (2006) GMM (M06 hereafter) is directivity centered (i.e., that $\mu(M, R, \theta, T)$ represents a directivity neutral condition). This limits the extent to which we can understand the impact of directivity on long-period seismic hazard. Instead, we undertake two analysis: the first applies the selected directivity amplification model to M06 (regardless of inconsistencies in horizontal component definition), the second replaces M06 with the GMM of Chiou & Youngs (2014) [CY14 hereafter] with its original direct-point parameter directivity term (f_{DPP}) turned off. This latter option is a GMM that is known to represent a directivity neutral condition (when $f_{DPP} = 0$), is broadband ($0.01 \leq T \leq 10$ s) and is potentially more informative as to the impacts on directivity when applied to the 2022 NSHM GMM selection, among which CY14 is adopted directly in one GMM logic tree branch, and its adaptation by Stafford (2022) constitutes a scaled backbone GMM in a different set of branches in the logic tree (Bradley et al., 2022).

For the directivity model we will be using Bayless et al. (2020) [Ba2020 hereafter]. After reviewing several available models in the literature, this model was selected as it fulfills the criteria of being directivity centered in its empirical calibration, narrowband, efficient to calculate and suitable for multi-segment ruptures owing to its use of the Generalised Coordinate System (GC2). Furthermore, Ba2020 is not explicitly dependent upon the angle of local slip (γ), meaning that $P[\gamma] = 1$ throughout, removing the need for this additional loop in the PSHA calculation. Ba2020

was implemented into the OpenQuake-engine seismic hazard calculation software and the resulting amplifications fields verified against those produced by code supplied directly from the author. Furthermore, the directivity fields from specific earthquake ruptures of interest in New Zealand were inspected to verify coherence of the expected patterns.

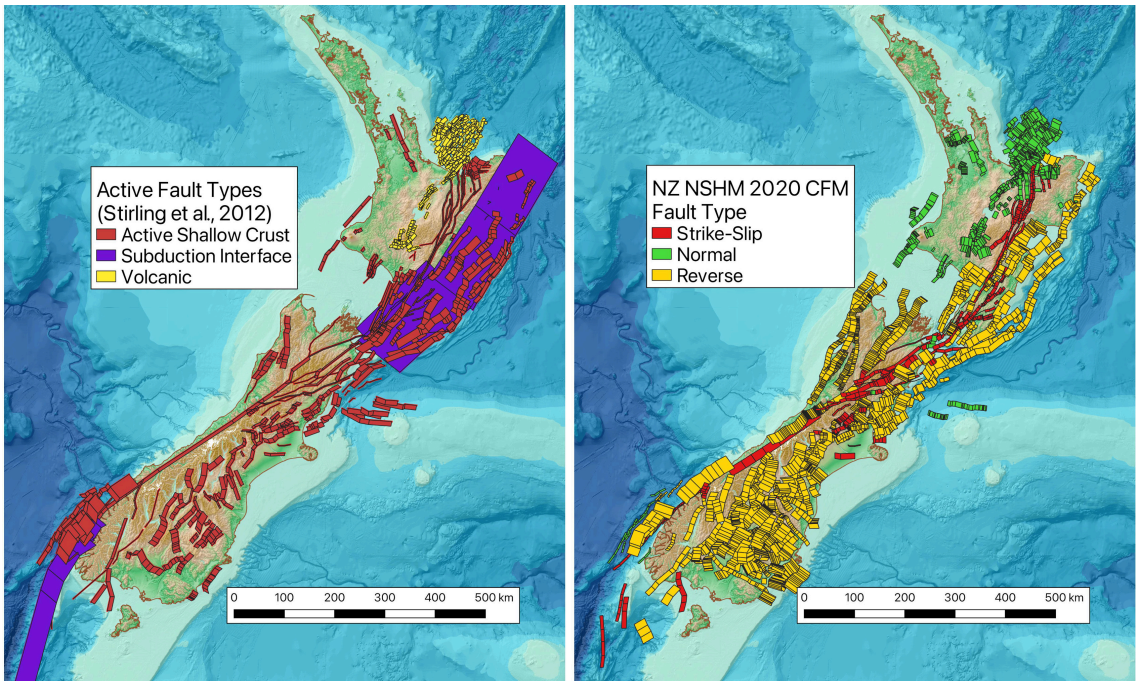

Figure 2. Active faults in New Zealand according to the 2012 NSHM (Stirling et al. 2012) (left) and the Community Fault Model constructed as part of the 2022 NSHM (right).

Although Ba2020 does not have a dependence on γ, it does required integration over the distribution of relative hypocenter positions within the rupture, so a hypocenter model is required. For this we adopted the preferred hypocenter distribution of Watson-Lamprey (2018) for along-strike hypocenter position on strike-slip and dip-slip faults, combined with the down-dip hypocenter position distributions proposed by Melgar & Hayes (2019).

Results of the Directivity Applications

PSHA is calculated using the full hypocentre randomization approach with directivity on the shallow fault sources. The percent changes in seismic hazard for a period of Sa (3.0 s) is shown for the 475 year and 2475-year return period in Figures 3 and 4 respectively. The greatest increase in seismic hazard is seen around the Alpine fault, on the western side of the South Island, and its branches onto the Clarence and Hope faults across the northern South Island. Increases are on the order of 10 – 15 % for the 475-year return period and up to 30 % for the 2475-year return period, though this number is much less at shorter spectral periods and diminishes to 2 – 3 % and 10 % for the same return periods for Sa (1.0 s). Small increases can also be seen in the Lake Taupo region of the North Island. For other areas such as the Canterbury plains (eastern South Island) and the northern coast of the North Island we observe a net decrease in seismic hazard when including directivity.

An initial conclusion that one might draw from this analysis is that the overall impact of modelling directivity explicitly in PSHA in New Zealand is relatively modest, with increases on the order of at most 20 – 30 % and largely limited to long spectral periods. Some results may even seem counter intuitive. In Christchurch, for example, where the seismic hazard decreases even though directivity is included on the Greendale fault, which produces amplification in Christchurch regardless of the hypocenter location, as was observed in the 2010 Darfield earthquake. Many factors influence the seismic hazard at a site, however, and several of these moderate the influence of directivity in the present case. The magnitude frequency distribution (MFD) is one major factor, which as we are considering purely characteristic faults with a single magnitude and rate of recurrence means that the fault in question may only impact on seismic hazard at very long return periods; hence, certain faults only become visible in the PSHA difference map in Figure 3 because their annual rates of recurrence are less than 1 / 475 years. As seismic hazard

at a site represents the net effect of potentially many different sources, in some cases positive increases in hazard from directivity in one fault are offset by potential decreases in hazard on other faults with higher annual rates. This is the case in Christchurch for the 2012 NSHM where the nearby Greendale fault, which should produce directivity amplification, has an annual rate of occurrence for M 7.1 of 4.0×10^{-5}, while other reverse faults to the north of the city have higher annual rates of occurrence and produce slight de-amplification.

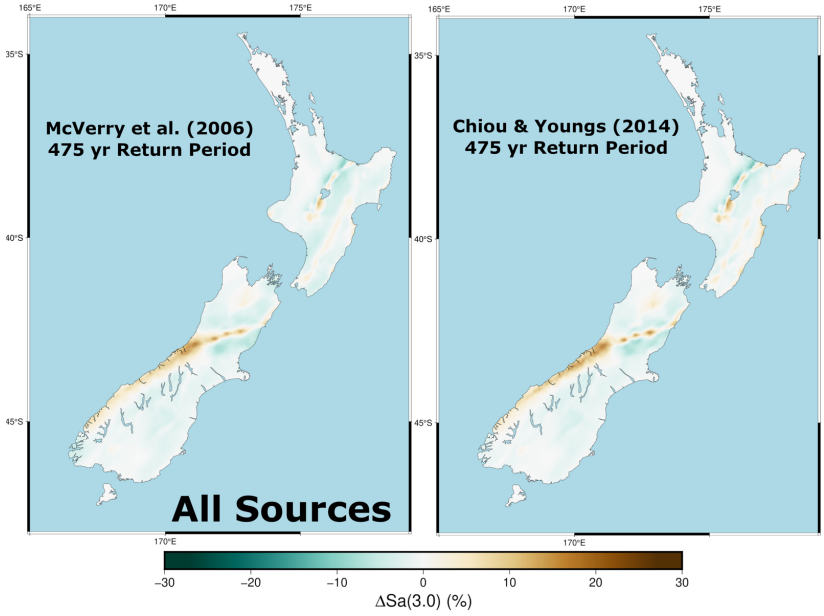

Figure 3. Change in 475-yer return period seismic hazard for Sa (3.0 s) when including explicit directivity modelling. McVerry et al. (2006) (left) and Chiou & Youngs (2014) (right).

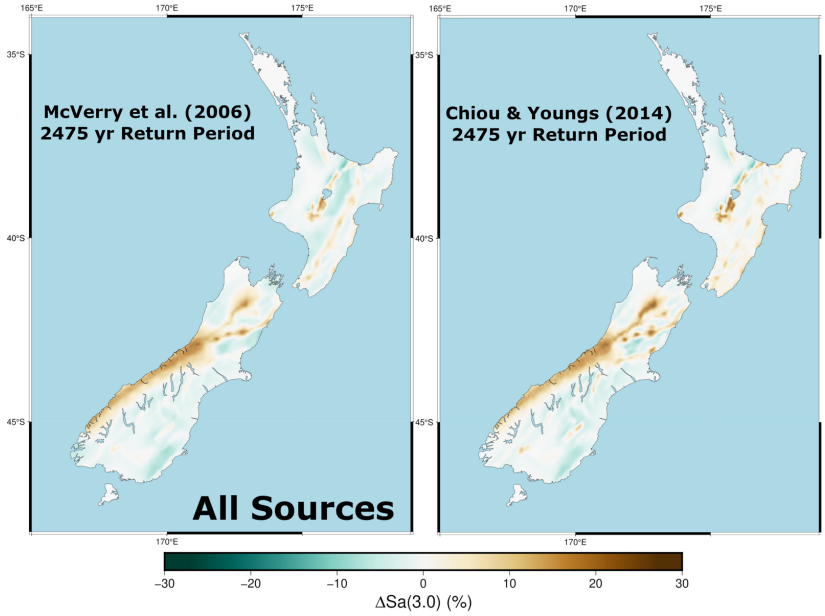

Figure 4. As Figure 3, for the 2,475-year return period.

The other major influencing factor is the symmetry of hypocenter distribution $f_H(h)$, which means that both directions of rupture propagation (toward and away from a site) are equally weighted for strike-slip ruptures, with bilateral rupture propagation more common than unilateral. The northern end of the Alpine fault section where it splits into several strike-slip faults (the Hope, Clarence, Awatere and Wairu faults) emerges as the region most influenced by directivity because this is the endpoints of the respective faults and would likely experience a degree of forward directivity regardless of the hypocentre location.

The regions of highest directivity increase are in relatively good agreement with both the area of application and scale of $N(T, D)$ in the South Island (comparing against Figure 1). In the North Island $N(T, D)$ is intended for application on the Wellington, Wairapapa and Mohaka faults, though these regions are not found to show such an increase in seismic hazard. Here the influence of the Hikurangi subduction becomes more significant for the hazard, for which no explicit directivity is modelled. Given this, one might argue that the current provisions for $N(T, D)$ in NZS1170.5, while not based on PSHA, remain generally adequate and may benefit only from modest revision to slightly increase the distance of influence of near-fault amplification from the current 20 km limit to possible 50 – 60 km. This conclusion is tentative, however, and from the analysis shown it is possible that both the pattern and scale of the change in hazard may depend strongly on the source model, which means we must turn our attention to the 2022 NSHM.

The 2022 NZ NSHM: Directivity for an ERF-Based Model

The 2022 NSHM sees a radical change in source modelling by adopting an *ERF* source model based on the UCERF 3.0 approach (Field et al., 2013). Starting with the description of the geometry of the entire fault system (the community fault model, CFM), faults are divided into 15 x 15 km micro-segments for crustal sources. Given a target MFD inferred from seismicity, and constrained by observations of geological slip, models geodetic deformation and magnitude-to-area scaling, an inversion process is applied that yields complete ERFs describing the inventory of possible ruptures in the system alongside their probabilities of occurrence. Details are provided in Gerstenberger et al. (2022b). Critical for our assessment of directivity, however, is the ERF itself and the nature of the ruptures contained therein.

In the ERF source model of the 2022 NSHM earthquake ruptures are no longer characteristic traces for each individual fault source (as they were in the 2012 NSHM) but are instead combinations of microsegments constrained by fundamental assumptions of earthquake physics. The earthquake sources can become substantially more complex, allowing for ruptures comprising segments from different fault sources, geometries with discontinuities and step-overs, and several hundred km long ruptures with large magnitudes (M > 8). Among the ruptures are many cases that include participation in the earthquake rupture from segments of faults with substantially different dips and/or rakes. These are considered “multi-fault” ruptures, as they represent multiple *different* earthquakes (each with their own individual hypocentre) combining into a single slip event. They not only stretch limits of applicability of current directivity models, but they also require a different approach to implementation than the full hypocenter randomization shown previously for the 2012 NSHM.

Recent work by Al Atik and Gregor (2020, personal communication) attempted an implementation of such a model to the original UCERF3 sources. We leverage on several aspects of their approach but adopt a different strategy for efficient implementation in PSHA at regional scale by developing a machine learning based *modifier of moments* directivity model to predict spatial patterns of amplification for set of ERFs using a convolutional neural network (CNN), a type of machine learning model that has been proven to perform well at learning complex 2D spatial trends in fields of ground motion data (Lilienkamp et al., 2022).

The key idea of the modifier of moments approach is to calculate the mean expected directivity related amplification and the corresponding variability for each rupture prior to the actual hazard calculation and utilize the results to modify the ground motion model during usage in the hazard integral. The most straightforward implementation of this approach would be to simply store the results for $\Delta\mu_{dir}$ and σ_{dir} for all considered ruptures, periods, and sites in a lookup table that could then be accessed during the hazard calculations. However, with a size of about 480 MB, such a lookup table would be far too large for an efficient implementation in a full scale and parallelized seismic hazard calculation for the New Zealand NSHM. Alternatively, Watson-Lamprey (2018) suggests to fit and use a simplified equational form to the modifiers $\Delta\mu_{dir}$ and σ_{dir} that could be used far more efficiently in PSHA calculations, although loss of detail in the spatial patterns of $\Delta\mu_{dir}$ and σ_{dir} must be accepted. In this study we want to suggest a new approach utilizing an artificial neural network (ANN) that acts like a strongly compressed and efficient lookup table, that preserves the finer nuances in the spatial patterns of $\Delta\mu_{dir}$ and σ_{dir}.

Generating the Mean and Variance of Directivity Amplification for the Ruptures

The CNN-based modifier of moments model requires that for each rupture in the ERF we generate fields describing the mean amplification for directivity ($\Delta\mu$) and its corresponding standard

deviation (σ_{dir}). For this we generate 200 fields of directivity amplification using Ba2020, varying the hypocentre position randomly according to the same $f_H(h)$ model described previously. Multi-fault cases are identified according to three criteria specified by Al Atik and Gregor (2020, *personal communication*): 1) a change in dip angle between consecutive sub-sections of the fault exceeding 50˚, 2) a change in first quadrant rake angle between consecutive sub-sections exceeding 30˚, 3) a gap of more than 10 km is between successive sub-sections. From the ERF provided we find 2,184 ruptures with M ≥ 6.8, of which 64 are multi-fault cases. For the single fault cases the directivity is calculated for each hypocentre, but in the multi-fault case each sub-fault has its own hypocentre. Following the approach of Al Atik and Gregor (2020, *personal communication*), the hypocentre is sampled independently for each sub-fault and the maximum directivity parameter at each site from each sub-fault is taken as the directivity parameter for determination of the field. Altogether 436,800 directivity fields are generated for a target mesh of 5 km spacing covering onshore New Zealand. These are grouped per rupture to produce the 2,184 fields of $\Delta\mu$ and σ_{dir} unique to each in the ERF, which form the data to train the ANN.

Implementing a neural network based version of the modifier of moments approaches

The general target in deep learning is to teach an artificial neural network (ANN) to learn a relationship between a set of input parameters, e.g., the properties of a seismic rupture, and a target parameter, e.g., the spatial distribution of $\Delta\mu_{dir}$ in the vicinity of the rupture (LeCun et al., 2015). This target is achieved via training the ANN, i.e., by showing a large amount of pairs of input and target parameters (a training dataset) to the ANN, such that the ANN can autonomously learn the underlying relationship. Every few steps during this training phase, the performance of the ANN is evaluated using data examples from a second dataset, the validation set. If the performance of the ANN improves on these "unknown" data examples as the training proceeds, this indicates that the ANN is indeed learning a universally valid relationship in the dataset and that it does not overfit to the specific examples in the training dataset.

In the specific task at hand, we want to train an ANN to predict spatial patterns of $\Delta\mu_{dir}$ and σ_{dir} at multiple periods given a representation of a seismic rupture from the New Zealand NSHM. While avoiding overfitting is usually an absolutely crucial objective when training an ANN, the fact that the number of 2,184 ruptures is fix in the hazard model lifts the necessity for generalization to new ruptures, and allows us to intentionally overfit a small-sized ANN to learn with high detail the fields of $\Delta\mu_{dir}$ and σ_{dir} for all the ruptures in the New Zealand NSHM.

We use a 2D convolutional neural network as this type of architecture is known to work very well for spatial problems. The input layer consists of a single numerical value that takes the unique ID of the fault. Via various mathematical operations (e.g., convolution, embedding, and up-sampling) an output matrix of the shape (384 x 320 x 22) is generated, where the first two dimensions refer to the shape of the regular 5km x 5km grid covering New Zealand, and the 22 channels provide the resulting estimates $\widehat{\Delta\mu_{dir}}$ and $\widehat{\sigma_{dir}}$ of $\Delta\mu_{dir}$ and σ_{dir}, respectively, for 11 different periods.

The ANN is trained as follows: In a first step we provide a fault ID as input to the model forward-run the ANN. Next, we quantify the misfit between the generated and the desired output as the mean squared error. Finally, we use backpropagation and gradient descent to update the model parameters in the neural network to provide an output similar to the desired one. After repeating this procedure 5000 times for every rupture in the dataset, the training phase is completed, and the model can henceforth be utilized to retrieve maps of $\widehat{\Delta\mu_{dir}}$ and $\widehat{\sigma_{dir}}$ for each rupture in the New Zealand NSHM. A selection of typical examples of the actual maps and reproductions by the ANN are presented in Figure 5.

From these examples one can clearly see that the ANN can reproduce the spatial amplification patterns with very little loss of accuracy. This quality holds true for all 2,184 ruptures in the dataset. Due to the intentional overfitting, the model is applicable only to the exact set of ruptures that are considered in the New Zealand NSHM, only to the Ba2020 directivity model, and only to the set of 11 periods that are considered in this study. As the ERF models tend to retain the same rupture sets with different probabilities on different source model branches, however, the ANN approach is easy to apply even for larger numbers of source model logic tree branches. While the amplification patterns $\widehat{\Delta\mu_{dir}}$ and obtained $\widehat{\sigma_{dir}}$ from the ANN are barely distinguishable from the original fields $\Delta\mu_{dir}$ and σ_{dir}, the size of the ANN on hard disk is almost 125 times smaller than that of the corresponding lookup table. Such a compact ANN is efficient and scalable in a fully parallelized, regional-scale PSHA calculation.

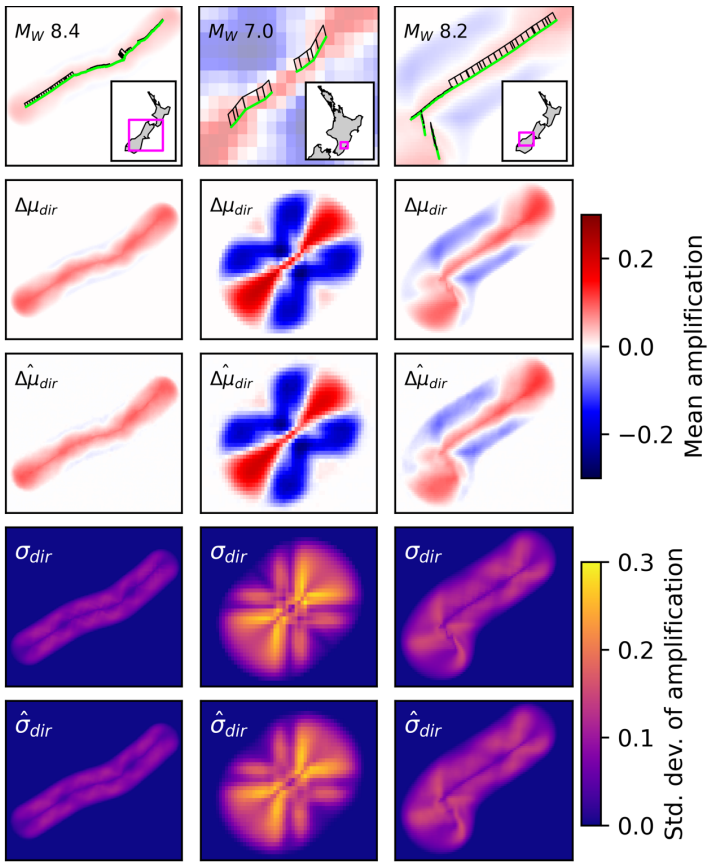

Figure 5: Comparison of precalculated fields of $\Delta\mu_{dir}$ and σ_{dir} with their reproductions $\widehat{\Delta\mu_{dir}}$ and $\widehat{\sigma_{dir}}$ from the decoder network for three example ruptures.

Conclusions

The relevance of directivity for the seismic response of tall and/or flexible structures has been recognized for several decades and pulse-like ground motions have been observed in many large earthquakes including the recent February 2023 earthquake sequence in Turkey. We have shown here that using recent empirical models for directivity amplification from the literature it is possible to implement this in PSHA at a regional scale for New Zealand. While this is found to produce an increase in seismic hazard in areas close to large active faults, the influence on the results of the 2012 NSHM is somewhat modest overall. As incorporation of directivity can incur considerable computational demand to the PSHA calculation, and while current models in the literature are themselves quite divergent, the cost of modelling directivity explicitly must be weighed against its influence on the seismic hazard. For New Zealand, the analysis here may be sufficient to demonstrate where and to what extent directivity influences the resulting seismic hazard. The results suggest that the current $N(T, D)$ might only benefit from minor re-calibration to obtain better agreement between the design spectrum and the UHS obtained with directivity in PSHA.

The 2022 NSHM is the second regional-scale PSHA to adopt the UCERF approach, but such models may become commonplace in the future. Hazard modellers are confronted with the need to characterize near-fault effects with respect to geometrically complex ruptures, which are under-represented in databases of observed ground motion. We demonstrate how machine learning via a neural network can facilitate application of existing directivity models to such sources in the framework of PSHA using an enhanced modifier of moments approach. In the current application this makes directivity modelling for ERF models feasible inside the PSHA calculation. In the longer term, however, CNNs may open several possibilities for efficient characterization of near-field ground motion in a PSHA framework. One possibility is to be able to capture detailed characteristics of near-fault ground motion from databases ground motion simulations, which can then be implemented into non-ergodic GMMs in a hazard and risk calculation framework. As a non-parametric form of predictive model, this type of neural network can potentially form a bridge toward integrating physics-based simulations into PSHA in a feasible and efficient manner.

References

Baker JW (2007), Quantitative Classification of Near-Fault Ground Motions Using Wavelet Analysis, *Bulletin of the Seismological Society of America,* 97(5): 1486 – 1501

Bayless J, Somerville P, Skarlatoudis A (2020), A rupture directivity adjustment model applicable to the NGA West2 ground motion models and complex fault geometries, *U. S. Geological Survey Technical Report*, USGS Award G18AP0092

Bradley BA, Bora S, Lee RL, *et al*. (2022), Summary of the ground-motion characterization model for the 2022 New Zealand National Seismic Hazard Model, *GNS Science Report* 2022/46 doi:10.21420/9BMK-ZK64

Chiou BS -J, Youngs RR (2014), Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra, *Earthquake Spectra, 30(3): 1117 - 1153*

Donahue JL, Stewart JP, Gregor N, Bozorgnia Y, (2019), Ground-motion directivity modeling for seismic hazard applications, *Pacific Earthquake Engineering Research Centre,* PEER Report No 2019/03

Field EH, Jordan TH, Cornell CA (2003), OpenSHA: a developing community-modeling environment for seismic hazard analysis, *Seismological Research Letters*, 74(4): 406 – 419

Field EH, Arrowsmith RJ, Biasi GP, *et al*. (2014) Uniform California earthquake rupture forecast, version 3 (UCERF3) – the time-independent model, *Bulletin of the Seismological Society of America*, 104(3): 1122 - 1180

Gerstenberger MC, Bora S, Bradley BA *et al*. (2022a), New Zealand National Seismic Hazard Model 2022 Revision: model, hazard and process overview, *GNS Science Report* 2022/57, doi:10.21420/TB83-7X19

Gerstenberger MC, van Dissen RJ, Rollins C, *et al.* (2022b), The Seismicity Rate Model for the 2022 New Zealand National Seismic Hazard Model, *GNS Science Report* 2022/47, doi:10.21420/2EXG-NP48

Hall JF, Heaton TH, Halling MW, Wald DW (1995), Near-source ground motion and its effects on flexible buildings, *Earthquake Spectra,* 11(4): 569 – 605

LeCun Y, Bengio Y, Hinton G (2015), Deep learning, *Nature*, 521: 436 – 444 https://doi.org/10.1038/nature14539

Lilienkamp H, von Specht S, Weatherill G, Caire G, Cotton F (2022), Ground-Motion Modeling as an Image Processing Task: Introducing a Neural Network Based, Fully Data-Driven, and Nonergodic Approach. *Bulletin of the Seismological Society of America*, 112(3): 1565 - 1582

McVerry GH, Zhao JX, Abrahamson NA, Somerville PG (2006), New Zealand acceleration response spectrum attenuation relations for crustal and subduction zone earthquakes. *Bulletin of the New Zealand Society for Earthquake Engineering,* 39(1): 1 – 58.

Melgar D, Hayes GP (2019), The correlation lengths and hypocentral positions of great earthquakes. *Bulletin of the Seismological Society of America*, 109(6): 2582 - 2593

Pagani M, Monelli D, Weatherill G, Danciu L, Crowley H, Silva V, Henshaw P, Butler L, Mastasi M, Panzeri L, Simionato M, Vigano D (2014), OpenQuake Engine: an open hazard (and risk) software for the Global Earthquake Model, *Seismological Research Letters*, 85(5): 692 - 702

Somerville PG, Smith NF, Graves RW, Abrahamson NA (1997), Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity, *Seismological Research Letters*, 68(1): 199 – 222

Stirling MW, McVerry GH, Gerstenberger MC, *et al.* (2012), National seismic hazard model for New Zealand: 2010 update, *Bulletin of the Seismological Society of America*, 102(4): 1514 – 1542

Watson-Lamprey JA (2018), Capturing directivity effects in the mean and aleatory variability of the NGA-West2 ground-motion prediction equations. *Pacific Earthquake Engineering Research Centre,* Report No. 2018/04

Weatherill G (2022), Impact of directivity on probabilistic seismic hazard calculations in New Zealand, *GNS Science Report* 2022/01, doi:10.21420/RETZ-D556