PERFORMANCE OF RESPONSE SPECTRAL MODELS AGAINST NEW ZEALAND DATA

Chris Van Houtte[1,2]

(Submitted *May 2016*; Reviewed *August 2016;* Accepted *November 2016*)

ABSTRACT

An important component of seismic hazard assessment is the prediction of the potential ground motion generated by a given earthquake source. In New Zealand seismic hazard studies, it is commonplace for analysts to only adopt one or two models for predicting the ground motion, which does not capture the epistemic uncertainty associated with the prediction. This study analyses a suite of New Zealand and international models against the New Zealand Strong Motion Database, both for New Zealand crustal earthquakes and earthquakes in the Hikurangi subduction zone. It is found that, in general, the foreign models perform similarly or better with respect to recorded New Zealand data than the models specifically derived for New Zealand application. Justification is given for using global models in future seismic hazard analysis in New Zealand. Although this article does not provide definitive model weights for future hazard analysis, some recommendations and guidance are provided.

INTRODUCTION

Empirical ground-motion models (also known as attenuation relations or ground-motion prediction equations) relate ground-motion intensity measures that are of interest to the engineering and risk communities, such as peak ground acceleration (PGA), peak ground velocity (PGV) and spectral acceleration (S_A), to a set of variables that describe the effects of the source, wave propagation path and local site conditions. These models are a key ingredient of probabilistic seismic hazard assessment (PSHA), and despite the recent rise in popularity of simulated ground motions, are currently the most accepted way of modelling ground motions within a probabilistic framework. Therefore, the performance of ground-motion models against New Zealand seismic data is of strong interest to seismic hazard and risk assessments.

There are an ever-increasing number of site-specific PSHA studies being undertaken in New Zealand, but very little guidance on appropriate ground-motion models that can be used in a New Zealand context. It has been common practice to only apply New Zealand specific ground-motion models in local PSHA studies, despite the availability of numerous published models that are deemed suitable for global application [1]. Previous seismic hazard studies in New Zealand have largely avoided using global models, presumably because it was unclear how these models perform with respect to recorded data in New Zealand. Instead, hazard analysts have preferred to use New Zealand specific ground motion models, in case New Zealand earthquakes are in some way systematically different from other recorded international earthquakes.

The consequence of this practice is that only a very small number of models are available for application. As a result, many seismic hazard assessments within New Zealand, even today (e.g. [2-4]), are undertaken using a single ground-motion model. The pitfalls of the single-model approach have been known for more than 30 years [5]. As ground-motion model parameters cannot be estimated with certainty, alternative parametric forms and parameter values must be considered. This uncertainty associated with model forms and

parameters, typically referred to as 'epistemic uncertainty' in earthquake engineering practice, is a key consideration if the estimates of seismic hazard are to be robust. By applying only one model, or at most two models, New Zealand seismic hazard assessments give little consideration to epistemic uncertainty, and hence are falling behind the international state of the art.

The objective of this article is to take a first step towards addressing this problem, using the New Zealand Strong Motion Database described in this issue [6, 7]. Nine empirical ground-motion models are assessed against the New Zealand data, and justification is given for using internationally-derived models in New Zealand PSHA. Additionally, recommendations for a ground-motion model logic tree are provided, to improve how epistemic uncertainty is managed in New Zealand PSHA.

HISTORY OF EMPIRICAL GROUND MOTION MODELLING IN NEW ZEALAND

This article begins with a summary of previous efforts to model the behaviour of ground motions from large New Zealand earthquakes. While there have been several hundred ground-motion models developed around the world [8], there have been surprisingly few developed specifically for use in New Zealand. To my knowledge, the first New Zealand model for an instrumental ground motion measure was the peak ground acceleration (PGA) model of Matuschka (1980) [9]. The first response spectral model to have widespread use in New Zealand was the Katayama [10, 11] Japanese model, modified for New Zealand conditions by Peek (1980) [12], Mulholland (1982) [13] and McVerry (1986) [14] to account for the frequency response of Japanese strong-motion instruments, the apparent lower variability observed in New Zealand data, and different rates of path attenuation respectively. This modified model was adopted by Matuschka *et al.* (1985) [15] in the first response-spectrum based seismic hazard model for New Zealand, which was then used for the development of the earthquake loadings standard in NZS4203:1992 [16].

---

[1] *Corresponding Author, Engineering Seismologist, GNS Science, Lower Hutt, c.vanhoutte@gns.cri.nz (Member)*
[2] *Fellow, Harvard University, Massachusetts, United States*

The first attempt at deriving a statistical response spectrum model using New Zealand data was by Matuschka and Davis (1991) [17], although the dataset at the time was far too sparse to obtain a robust, well-behaved model. Several years later, Zhao et al. (1997) [18] developed the first high quality database of New Zealand strong-motion data from 1966-1995, supplemented with some short-distance international data, to develop a peak ground acceleration (PGA) expression for New Zealand crustal and subduction earthquakes. This database was expanded by Cousins et al. (1999) [19] with weak-motion recordings from the New Zealand National Seismograph network, and an alternative PGA equation was developed. It is an interesting historical note that the Cousins et al. (1999) equation may have been the first published equation worldwide to model regional attenuation effects, by including a separate term for ray paths within the Taupo volcanic zone.

A ground-motion model with a long history in recent New Zealand seismic hazard assessment is the McVerry et al. (2006) [20] model, which provides response spectra and PGA predictions for crustal and subduction zone earthquakes. This model was derived using the Zhao et al. (1997) dataset, and uses the form of the Abrahamson and Silva (1997) [21] model as the base for the crustal model, and the Youngs et al. (1997) [22] equation as the base for the subduction model. The model has been through numerous developmental stages over the years. An early prototype was used in the 1998 version of the National Seismic Hazard Model [23], before its first publication as a brief conference paper in McVerry et al. (2000) [24]. The full model was eventually published in 2006, in this journal. Since then, there have been three modified or alternative models published. The first modification was to better model earthquakes from the subducting Pacific plate beneath the North Island [25]. A variant of the model was presented in McVerry (2011) [26], which allows the site effect terms to vary continuously with site period, and finally an alternative subduction model was developed to incorporate simulated motions from the Hikurangi subduction interface for Wellington [27]. A common criticism of the McVerry et al. (2006) model is its lack of reproducibility, due to an incorrect term in the published equation. In the nonlinear part of the site term for Class D sites, the median rock PGA should be the median rock spectral acceleration at the spectral period of interest (G. McVerry, pers. comm.). This unpublished amendment to the equation is available from the author upon request.

A suite of models of the Arias Intensity [28] for New Zealand crustal earthquakes was developed by Stafford et al. (2009) [29], also using a New Zealand dataset supplemented with foreign recordings. Although the Arias Intensity has been demonstrated to be well correlated with damage to short-period structures, it is unclear how often this model is used in New Zealand seismic hazard assessment. Additionally, Stafford (2006) [30] developed an empirical model for the Fourier amplitude spectrum of New Zealand crustal earthquakes using the same dataset.

The most recent model for New Zealand data is that of Bradley (2013) [31], for crustal earthquakes. Like the McVerry et al. (2006) model, this study considered the New Zealand dataset too sparse to directly derive an equation, and hence adopted an overseas equation as a base model. The model itself is largely the same as the Chiou et al. (2010) [32] modification of the Chiou and Youngs (2008) [33] equation, with some additional modifications to correct for residual biases in small magnitude scaling, class A site response, anelastic attenuation in the New Zealand crust, normal faulting events and volcanic path attenuation. Due to deficiencies in modelling Christchurch data, particularly at long periods, Bradley (2015) [34] developed some Christchurch-specific modifications for the Bradley (2013) model, to better represent systematic source and site effects for various sub-regions within and around the city.

The current published version of the National Seismic Hazard Model (NSHM) [2] only utilises the McVerry et al. (2006) equations to model the behaviour of ground-motion within New Zealand. However, the Canterbury seismic hazard model of Gerstenberger et al. (2014) [35] uses versions of both the McVerry et al. (2006) and Bradley (2013) models in a logic tree framework, with weights on the logic tree branches determined by expert elicitation. This logic tree has been updated since its 2014 publication, with the branches corresponding to the Bradley (2013) model replaced by branches with the Bradley (2015) modified model, although the branch weights remain unchanged (M. Gerstenberger and G. McVerry, pers. comm.).

CONSIDERED MODELS

The objective of this article is to assess the performance of these New Zealand models, and also to test internationally-developed models against New Zealand data. Six models for shallow crustal events in active regions are considered here, hereafter referred to as 'crustal models', as well as four models for subduction interface and subducted slab events, hereafter referred to as 'subduction zone models'.

The first two crustal models are the McVerry et al. (2006) and Bradley (2013) equations, as these are the primary response spectral models currently used in New Zealand seismic hazard and risk studies. Four global models from the Next Generation Attenuation West2 (NGA-W2) project [36] are also included in this comparison, as these four are widely considered to be the best ground-motion models currently available for shallow crustal earthquakes in tectonically active regions. These models are the Abrahamson et al. (2014) [37], Boore et al. (2014) [38], Campbell and Bozorgnia (2014) [39] and Chiou and Youngs (2014) [40] models. The fifth NGA-W2 model, Idriss (2014) [41], is only intended for use at sites with $V_{S30} \geq$ 450 m/s, hence is not considered in this study due to its limited range of applicability.

For subduction zones, the first considered model is the New Zealand specific McVerry et al. (2006) equation. Additionally, the three subduction zone models recommended for global application by Stewart et al. (2015) [1] are also included, namely the Abrahamson et al. (2016) [42], Atkinson and Boore (2003) [43, 44] and the Zhao et al. (2006) [45] models.

The considered models are not intended to be an exhaustive list of all published equations. The reason for limiting the number of models is a practical one, to ensure simple implementation across the various open-source PSHA software currently available. Additionally, preference was given to models that are derived from global datasets, rather than from specific, foreign, regional datasets.

METHOD FOR COMPARISON

Performance Evaluation of Models

There are many ways to test the performance of the six crustal and four subduction zone models. In the past decade, numerous studies have been undertaken to test the goodness-of-fit of ground motion models against strong-motion datasets, usually by analysing model residuals (observed data minus model prediction). For example, Scherbaum et al. (2004) [46] analysed the total model residual i.e.

$$y = \mu(\mathbf{X} \mid \beta) + \varepsilon \tag{1}$$

where y is a ground-motion variable e.g. the logarithm of the peak ground acceleration (PGA), $\mu(\mathbf{X} \mid \beta)$ is the logarithm of the model prediction given the predictors $\mathbf{X}$ and model parameters β, and ε is the total model residual. The Scherbaum et al. (2004) study developed a set of criteria for ranking the performance of models, which depend on a residual likelihood parameter denoted as the LH value, the absolute value of the mean and median of the normalised residuals, and the normalised residual standard deviations. This approach was adopted by Douglas et al. (2006) [47] and Douglas and Mohais (2009) [48] to investigate the behaviour of strong-motion data in the French Antilles, and by Allen and Brillon (2015) [49] to evaluate models against data from British Columbia.

This concept was extended by Stafford et al. (2008) [50], to account for the fact that most ground-motion models assume that the total residual may be partitioned into an event-specific component and a record-specific component i.e.

$$y_{ij} = \mu(\mathbf{X}_{ij} \mid \beta) + \eta_i + \varepsilon_{ij} \qquad (2)$$

where η_i is the between-event residual for earthquake i and ε_{ij} is the within-event residual for record j from a given earthquake i. In this method, the LH value is calculated for both the between-event and within-event residuals.

Scasserra et al. (2009) [51] introduced a new method to examine the performance of global ground-motion models against an Italian dataset with respect to the models' magnitude, distance and site effects scaling. Alternatively, Allen and Wald (2009) [52] tested ground-motion models against a global dataset by analysing residuals against key model parameters.

Around the same time, Scherbaum et al. (2009) [53] introduced a new framework for data-driven testing of ground-motion models, derived from information theory. This approach uses a measure known as the Kullback-Leibler (KL) divergence, to measure the difference between continuous distributions P and Q, denoted as $D_{KL}(P\|Q)$. In the context of ground motion, recorded observations can be interpreted as realisations of a complex process, described by a continuous random variable with distribution P. The KL divergence represents the loss of information when a ground-motion model Q is used to approximate P, and can be represented by:

$$D_{KL}(P \| Q) = -E_P \log_2(P) - E_P \log_2 Q \qquad (3)$$

in units of bits, where E_P is the statistical expectation with respect to P. In ground-motion modelling, P is unknown, hence the first term in equation (3) cannot be calculated. However, given that the objective is to compare different ground-motion models against P, this term can be considered a constant that cancels out. The second term can be approximated by calculating the negative average sample log-likelihood (LLH), given the observations $\mathbf{y}$:

$$-LLH(Q, \mathbf{y}) = -\frac{1}{N} \sum_{i=1}^{N} \log_2[q(y_i)] \qquad (4)$$

where $q(y)$ is the density of Q and $i = 1, \ldots, N$ are the samples. Given that ground motion models are derived by assuming the logarithm of the spectral acceleration is normally distributed:

$$q(y_i) = \frac{1}{\sigma_{\bar{y}} \sqrt{2\pi}} \exp\left[\frac{-\left(y_i - \bar{y}\right)^2}{2\sigma_{\bar{y}}^2} \right] \qquad (5)$$

where $\bar{y}$ is the median model prediction, y_i is the sample and $\sigma_{\bar{y}}$ is the model standard deviation. This equation assumes that each sample is independent. The advantage of this approach is that the output is a distance measure with a reasonably simple interpretation, and the method was adopted by Beauval et al. (2012) [54], Delavaud et al. (2012) [55], Mousavi et al. (2012) [56], Edwards and Douglas (2013) [57] and Haendel et al. (2015) [58] to evaluate models against a French, global, Iranian, Australian and Chilean dataset respectively.

Bradley (2013) [31] took the same approach as Allen and Wald (2009) to select the Chiou and Youngs (2008) NGA equation as the basis for his ground-motion model, but avoided using statistical performance parameters for model evaluation, with the justification that the strong-motion dataset was too sparse. Since that study, a significant amount of large-magnitude, short-distance data have been recorded from the Canterbury and Cook Strait sequences. In this author's opinion, this now enables a data-driven statistical evaluation of the performance of published ground-motion models against New Zealand data, as presented in later sections of this paper.

To evaluate the performance of the nine considered ground-motion models against the New Zealand dataset, the methods of Allen and Wald (2009) and Scherbaum et al. (2009) are adopted. As a first step, the model residuals are calculated and analysed with respect to the model predictor. The –LLH value, as a proxy for the KL divergence, is subsequently calculated, and this information is used to examine the performance of the nine ground-motion models compared to New Zealand data. The –LLH approach is valid when models are independent of the test dataset. This condition is mostly satisfied in this study. Some of the older events were used to derive the McVerry et al. (2006) model, although the total number of common events in the overall test dataset is small. Similarly, the Darfield and Christchurch earthquakes were part of the dataset used to derive the NGA-W2 models, although these are only two of many large magnitude earthquakes used to derive these models.

Consistent Measures of Horizontal Ground Motion

Direct comparison of ground-motion models can be complicated, given that they predict a variety of different intensity measures. For example, the McVerry et al. (2006) model, which underpins the NZS1170.5:2004 code spectra [59] has separate coefficients for predicting geometric mean response spectra, and the larger of the two as-recorded horizontal components of ground-motion. On the other hand, the Bradley (2013) model is based on an alternative intensity measure, GMRotI50 [60], which was used in the NGA West 1 project [61] as an orientation-independent measure of horizontal seismic intensity, but was replaced in the NGA-W2 project by an alternative measure called RotD50 [62]. The four NGA-W2 global models are all in terms of an orientation-independent measure of the horizontal response spectrum, RotD50, as is the New Zealand database. The global subduction models are in terms of the geometric mean, except for Atkinson and Boore (2003), which uses 'both horizontal components'.

The comparison in this paper is solely in terms of RotD50 (the calculation of which can be found in [6], this issue). This necessitates some models being converted from alternative intensity measures, using previously-published conversion correlations. To convert the three geometric mean subduction models and the McVerry et al. (2006) geometric mean crustal model to RotD50, a Christchurch-specific equation has been derived by Bradley and Baker (2015) [63]. Figure 1a plots the

model of Bradley and Baker (2015) against Canterbury data from the New Zealand Strong Motion Database. In general, the Bradley and Baker (2015) model is similar to the Canterbury data (from which it was derived), except at long periods. The discrepancy at long periods is likely to arise from different processing of the waveforms. The wider New Zealand dataset is systematically different from the Bradley and Baker (2015) model for periods greater than around 0.4 s (Figure 1b), and a different conversion is applied (shown as the blue lines in Figure 1b).

The Bradley (2013) ground motion model is converted from GMRotI50 to RotD50 using the following equation for the median (approximated from Figure 3 of Boore, 2010 [62]):

$$\frac{RotD50}{GMRotI50} = \begin{cases} 1 & \text{for } T < 0.1 \\ 1 + 0.05 \times \left(\dfrac{\log(T/0.1)}{\log(100)} \right) & \text{for } 0.1 \leq T < 10 \\ 1.05 & \text{for } T = 10 \end{cases}$$

(6)

The Boore (2010) standard deviation results are approximated as a linear trend in $\log(T)$ space between 0.01 and 10 s, varying from 0.05 to 0.08 ln units. In addition to the different horizontal component definitions, the McVerry *et al.* (2006) model is different from the others considered here in that it predicts 5%-damped absolute acceleration response spectra, while the other models in this study predict 5%-damped pseudo acceleration response spectra. While the difference between the two types of spectra is not necessarily negligible, the difference has been ignored for the purposes of this comparison.

Subset of Database

The selection of data is a key issue in model comparison and will influence the results, particularly given that the New Zealand strong-motion dataset is imbalanced with respect to magnitude and distance. The method for data selection is different for crustal and subduction zones.

To evaluate the performance of crustal models, a subset of the New Zealand database with moment magnitude $M_W \geq 5$ and rupture distance $R_{rup} \leq 200$ km is utilised, with the added constraint that each event must have at least three recordings. The purpose of this constraint is to obtain a more reliable partitioning of the total residual into between-event and within-event components. For crustal earthquakes, the number
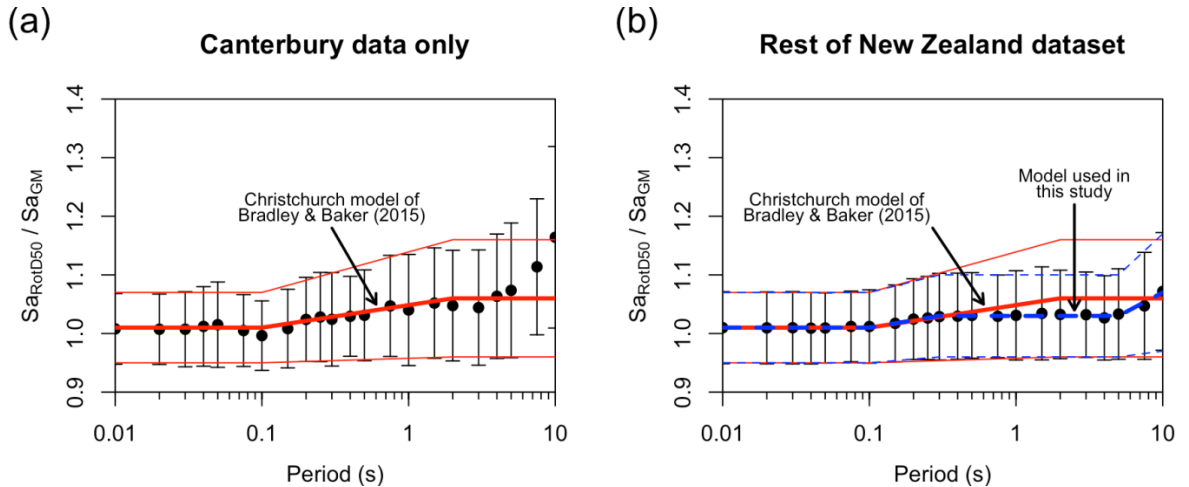
of records selected for comparison against response spectral period is shown in Figure 2a, with the magnitude-distance distribution of records for PGA and T = 10 s in Figure 2b. This subset of the New Zealand database contains data from the recent Canterbury and Cook Strait earthquake sequences, as well as older events such as the 1968 Inangahua and the 1994 Arthur's Pass earthquakes. There are significantly fewer data at T = 10 s than for PGA, due to the amplitudes of many earthquake recordings in the database falling to the amplitude of the long-period noise by 10s period.

To select data for subduction model comparisons, it was decided to separate the data from the Fiordland and Hikurangi subduction zones, given their very different interface orientations and crustal properties [64, 65]. While there are larger magnitude data available for the Fiordland subduction zone (up to M_w7.8), for brevity, the comparison of this paper is solely in terms of Hikurangi data, given the closer proximity of the Hikurangi subduction zone to infrastructure. The subset of the New Zealand database utilised for model comparison is all Hikurangi interface or Hikurangi slab events with $M_W \geq 5$ and $R_{rup} \leq 200$ km. Using this subset, the number of available records against oscillator period are shown in Figure 2c, and the magnitude-distance distribution of the dataset is shown in Figure 2d. The largest events in this subset of the database are the 2007 Gisborne M6.6 and 2014 Eketahuna M6.3 earthquakes, both events that occurred within the subducting Pacific plate, hence the data is unable to constrain the large magnitude scaling for Hikurangi subduction interface earthquakes.
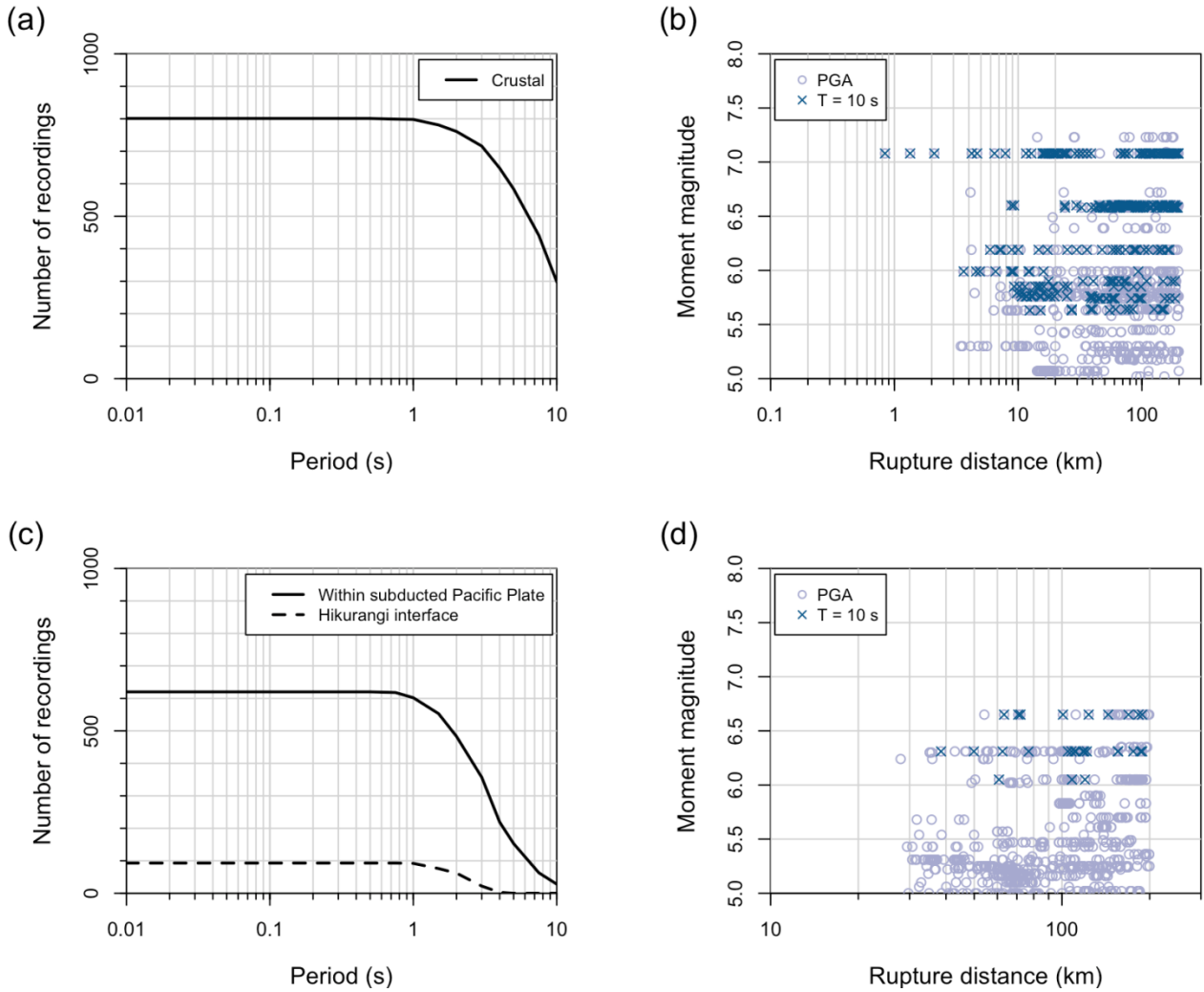
ANALYSIS OF RESIDUALS

For the data in Figure 2, this section examines the residuals of the six crustal models and four subduction zone models against the model predictors M_w and R_{rup}. The residuals are separated into between-event and within-event components, as demonstrated in equation (2). In a first step, the between-event residuals are calculated using the following equation from Abrahamson and Youngs (1992) [66]:

$$\eta_i = \frac{\tau^2 \sum_{j=1}^{n_i} (y_{ij} - \mu_{ij})}{n_i \tau^2 + \phi^2}$$

(7)

Figure 1: The ratio of the RotD50 spectral acceleration intensity measure to the geometric mean spectral acceleration against period. Results from the Canterbury data (a) and the wider New Zealand database (b) are represented by circles, with error bars representing standard deviations. Thick and thin red lines represent mean and ±1 standard deviation of the Bradley and Baker (2015) model respectively, while the blue dashed line represents the conversion used in this study.

Figure 2: (a) The number of crustal earthquake recordings used for the subset of the New Zealand database, and (b) comparison of magnitude-distance distribution of data for PGA (circles) and a spectral period of 10 s (squares) for the crustal database in (a). (c) The number of Hikurangi subduction zone records used in this comparison, separated into those from slab and interface events, and (d) the magnitude-distance distribution of the subduction zone records in (c), for PGA (circles) and T = 10 s (squares).

where n_i is the number of recordings associated with the earthquake i, and τ and ϕ are the given ground-motion model's expression for the between- and within-event variability respectively. Where applicable, τ and ϕ correspond to an average of the record-specific values associated with event i. The purpose of this equation is to ensure a stable partitioning of residuals in the case where an event has few recordings. If an earthquake only has one recording, the between-event component of the residual is given as $\tau^2/(\tau^2+\phi^2)$ times the total residual, but if the event has a large number of recordings then η_i tends to the mean residual. The within event residuals ε_{ij} are then calculated as per equation (2). While equation (7) doesn't exactly represent the inter-event term for the models considering nonlinear site response [67, 68], it is considered an acceptable approximation for the purposes of the exploratory analyses in this section.

Figures 3, 4 and 5 show residual plots for the six crustal models at PGA, T = 1 s and T = 5 s respectively. While an exhaustive comparison is possible for each response spectral period, only these three oscillator periods are shown for brevity. These periods are selected as examples of the short, mid-range and long period behaviour of the six models. Additionally, Figure 6 shows the PGA within-event residuals against NZS1170.5 site classifications. Figures 7 and 8 show PGA and T = 1 s spectral acceleration residuals for the four
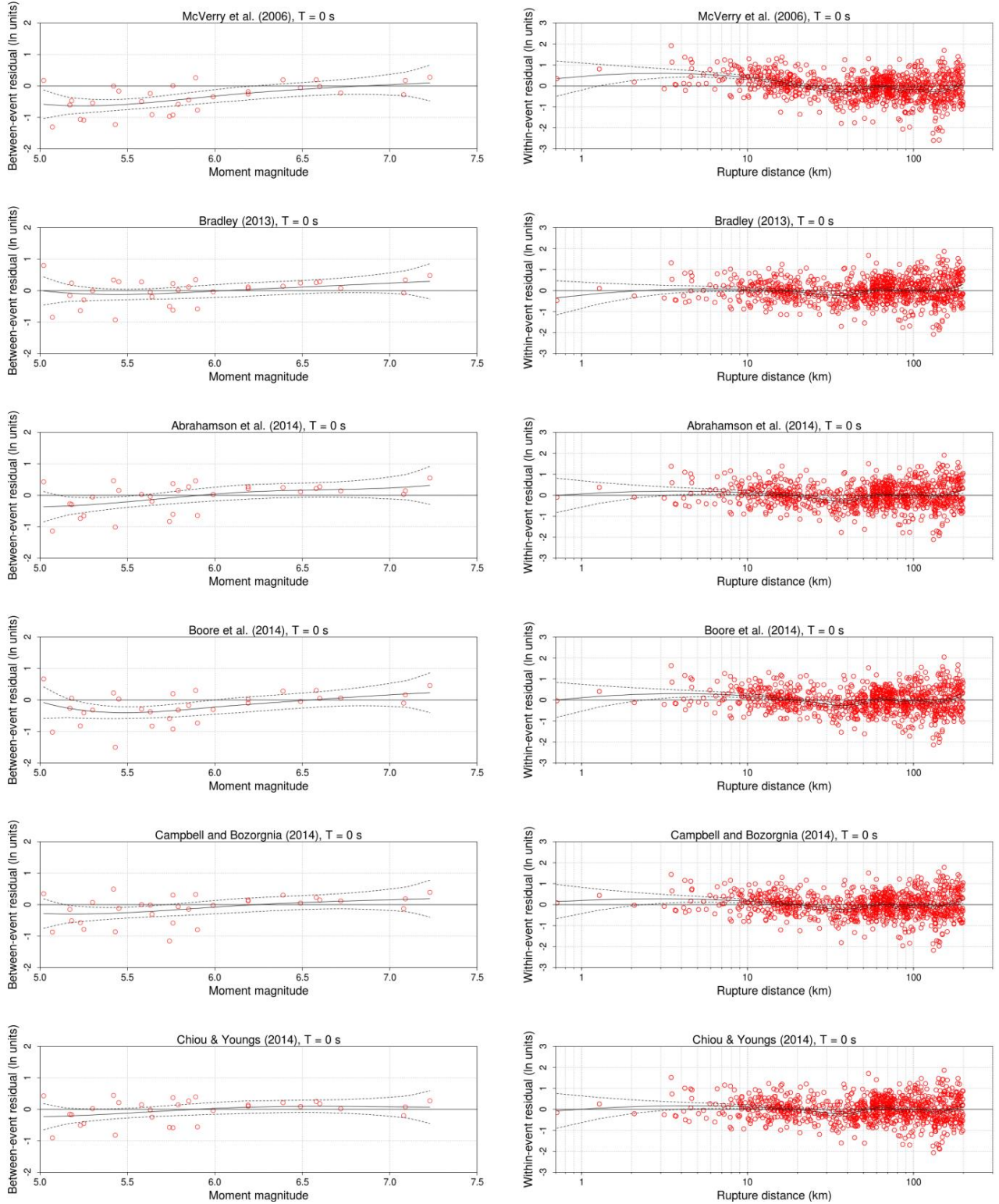
subduction zone models against the Hikurangi subduction zone dataset. There are too few data in the Hikurangi dataset at long periods to warrant a comparison here.

Crustal Models

Behaviour at PGA

Figure 3 shows how the six crustal models represent New Zealand PGA data, with between-event residuals against M_w in the left column, and within-event residuals against R_{rup} in the right column. With respect to M_w, the McVerry *et al.* (2006) model performs very well for magnitudes greater than 6, but significantly overpredicts the smaller magnitude data by around 80% on average. Conversely, the Bradley (2013) model provides good fits for the smaller magnitudes, but underpredicts the large magnitude data by around 20%. The Abrahamson *et al.* (2014), Boore *et al.* (2014) and Campbell and Bozorgnia (2014) models have similar behaviour to the McVerry *et al.* (2006) model, but the overprediction at magnitudes 5 to 6 is not as pronounced. These over-predictions of the small magnitude data may be related to the fact that mainshocks and aftershocks are not distinguished between in the New Zealand database, however the NGA-W2 models model lower motions for aftershocks. The model with
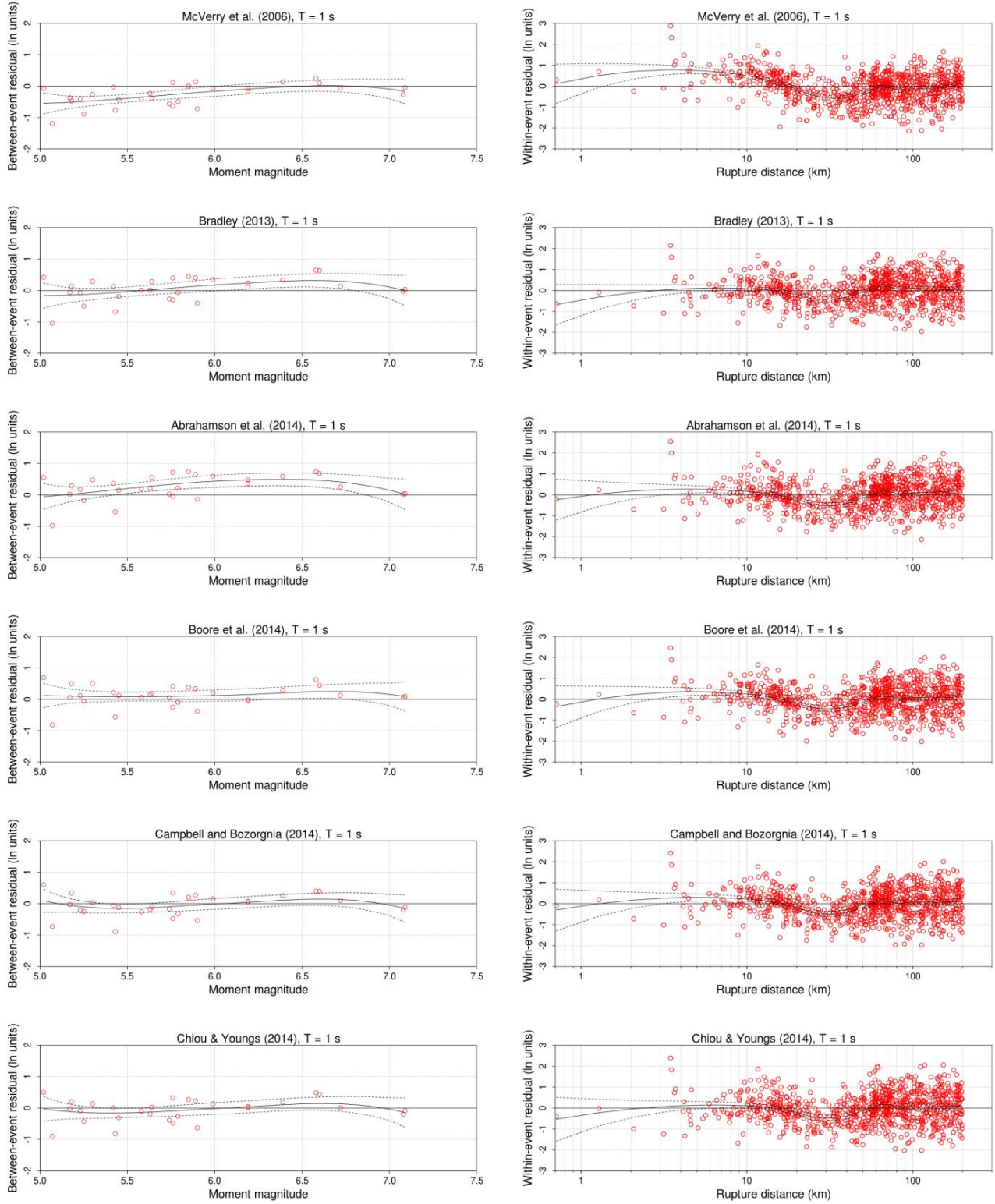
Figure 3: Residuals for the six crustal models considered in this study, with respect to the New Zealand PGA dataset. The left column is the between-event residual against magnitude, and the right column is the within-event residual against distance. Also indicated on each plot is loess fit to the residuals, with 95% confidence intervals.

the smallest bias with respect to magnitude is the Chiou and Youngs (2014) model, which fits the data well, particularly for the large magnitudes. All models have little bias in the within-event residuals with respect to the rupture distance, R_{rup}, although there is a hint of underprediction at short distances and over-prediction at 20-50 km distance. This effect is most pronounced for the McVerry *et al.* (2006) model, with underpredictions of around 50% on average, for $R_{rup} < 10$ km.

Behaviour at T = 1 s

At T = 1 s, the McVerry *et al.* (2006) model has similar trends to those at PGA. The Abrahamson *et al.* (2014) model tends to underpredict the *M* > 6 data, by around 60-70% on average. The Bradley (2013) and Boore *et al.* (2014) models also tend underpredict the larger magnitudes. The Campbell and Bozorgnia (2014) and Chiou and Youngs (2014) models have
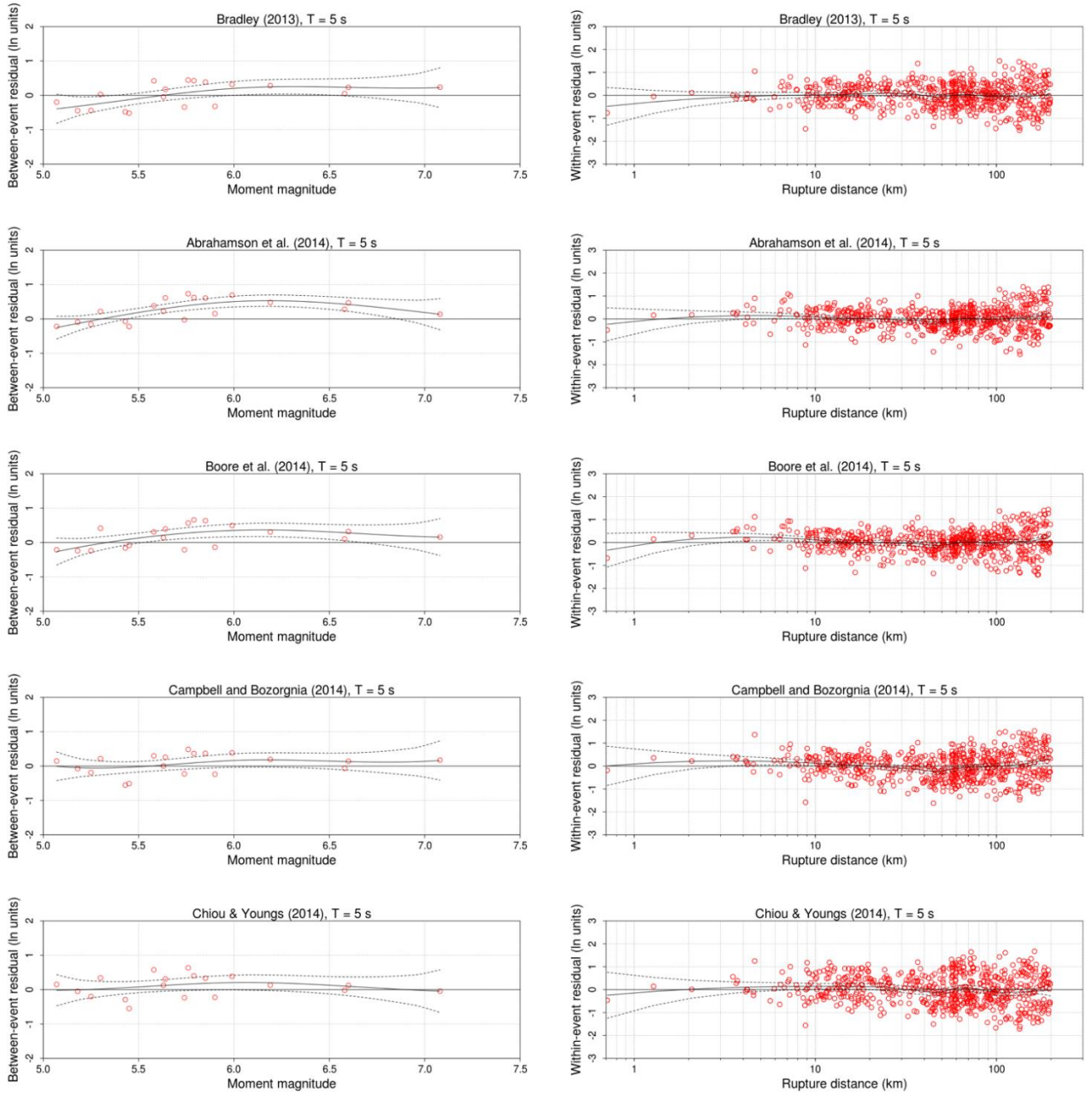
Figure 4: Residuals for the six crustal models considered in this study, with respect to the New Zealand spectral acceleration dataset at T = 1 s.

little bias with respect to earthquake magnitude for the T = 1 s New Zealand data.

With respect to distance, all models overpredict the New Zealand dataset at 20-50 km, the reasons for which are not currently clear. Of note is that there are two recordings with R_{rup} approximately equal to 3.5 km, which all models greatly underpredict by up to an order of magnitude. These recordings are from the June 2011 M6 earthquake from the Canterbury sequence, recorded at the Godley Drive and Panorama Drive rock sites in the Port Hills. Both of these sites are highly

influenced by topographic amplification, which is an effect that is not currently considered in empirical ground-motion models. In particular, the Godley Drive record has the largest horizontal PGA in the database ($PGA_{RotD50} = 1.48$ g), but has been shown to have strong topographic amplification at T = 1 s [69]. This explains the very large discrepancy between these data and the model predictions, particular in Figure 4. The lack of consideration for topographic effects is a shortcoming of empirical ground-motion models, and further research is necessary.

Figure 5: Residuals for the six crustal models considered in this study, with respect to the New Zealand spectral acceleration dataset at T = 5 s.

Behaviour at T = 5 s

The T = 5 s abscissa has been included in this paper (Figure 5) to illustrate the long-period behaviour of the ground-motion models, which is important for tall buildings and base-isolated structures.

The McVerry *et al*. (2006) model is not pictured in Figure 5 as the maximum period it is defined for is 3 s. The other models all underpredict the data, particularly for $M_w > 6$. The underprediction is strongest for the Bradley (2013), Abrahamson *et al*. (2014) and Boore *et al*. (2014) models, in the order of 20-50%. The Campbell and Bozorgnia (2014) and Chiou and Youngs (2014) models appear to be the best representations of the data at this oscillator period.
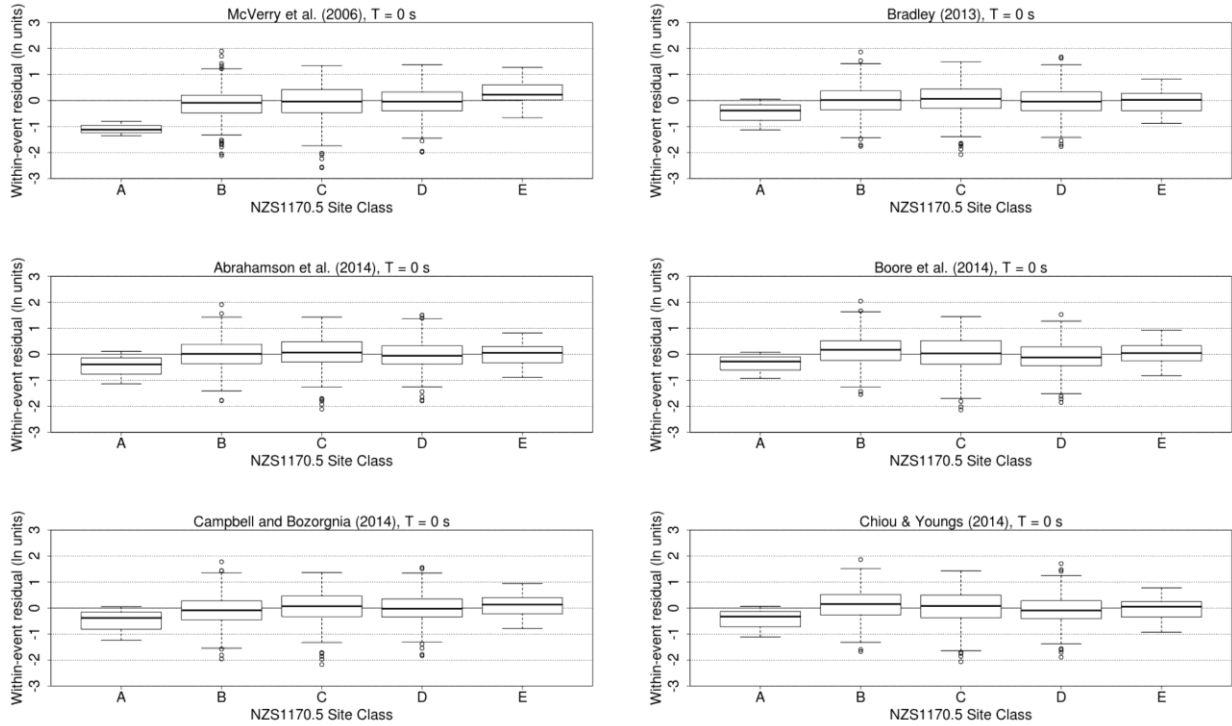
It appears that the models do not have major biases against distance, which suggests that they are reasonably representing the average path effects at long periods.
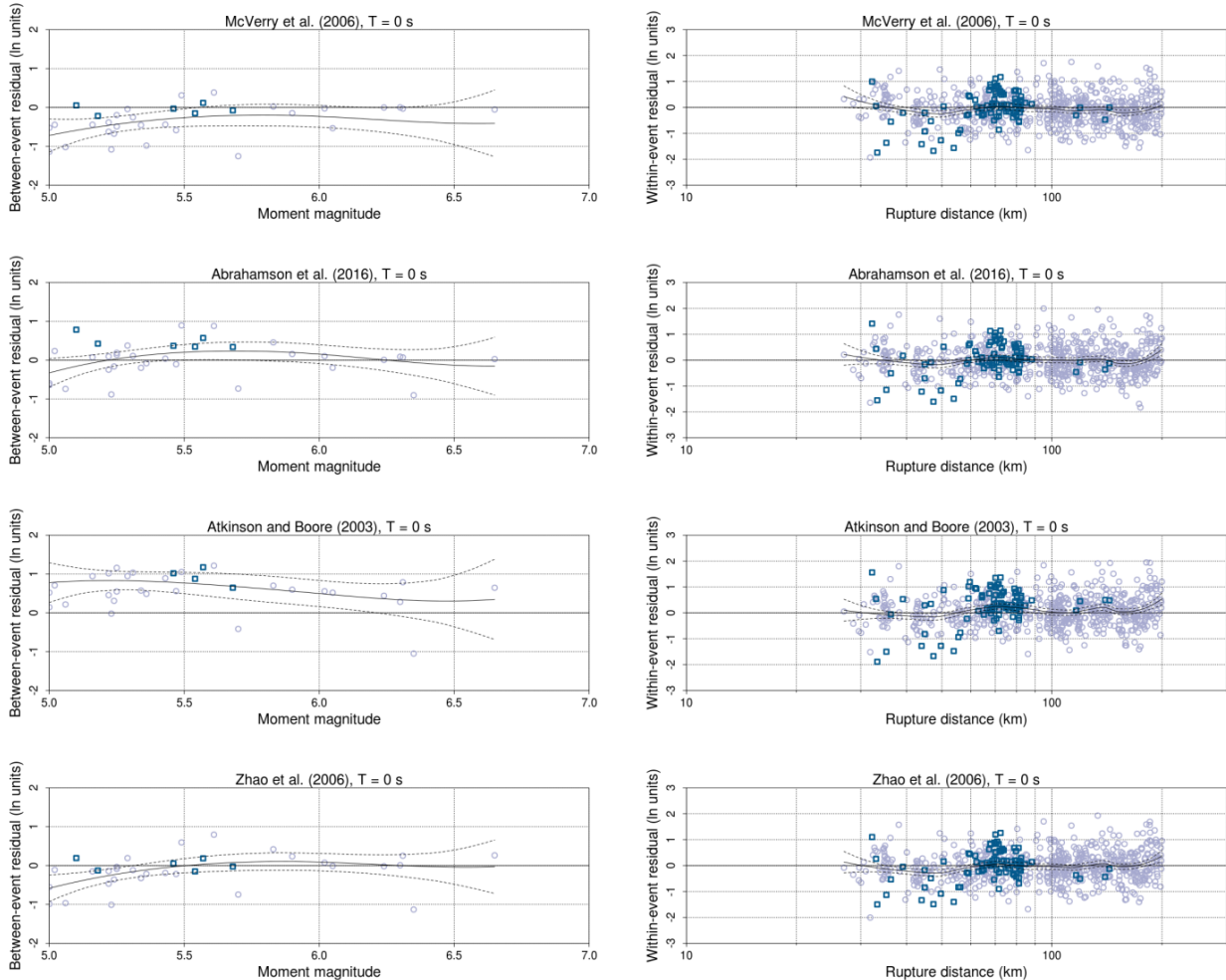
Site Class Dependence

Figure 6 shows the variation of the within-event residuals against NZS1170.5 site class [59]. Only PGA is shown here for brevity. The McVerry *et al.* (2006) model does not provide predictions for class E sites, hence for the purpose of these plots, the class D site terms were applied to class E sites. The McVerry *et al.* (2006) model generally performs well for classes B to D, but significantly overpredicts class A sites. This is not a surprising result, given that the model does not discriminate between class A and class B.

While the global models and the Bradley (2013) model do not include NZS1170.5 site class as a predictor, the model residuals are assessed against the site classification to examine any bias. All of these models show a strong over-prediction of class A sites, but do not have significant biases against other site classes. While Bradley (2013) modified the Chiou and Youngs (2008) model to account for class A site response, the
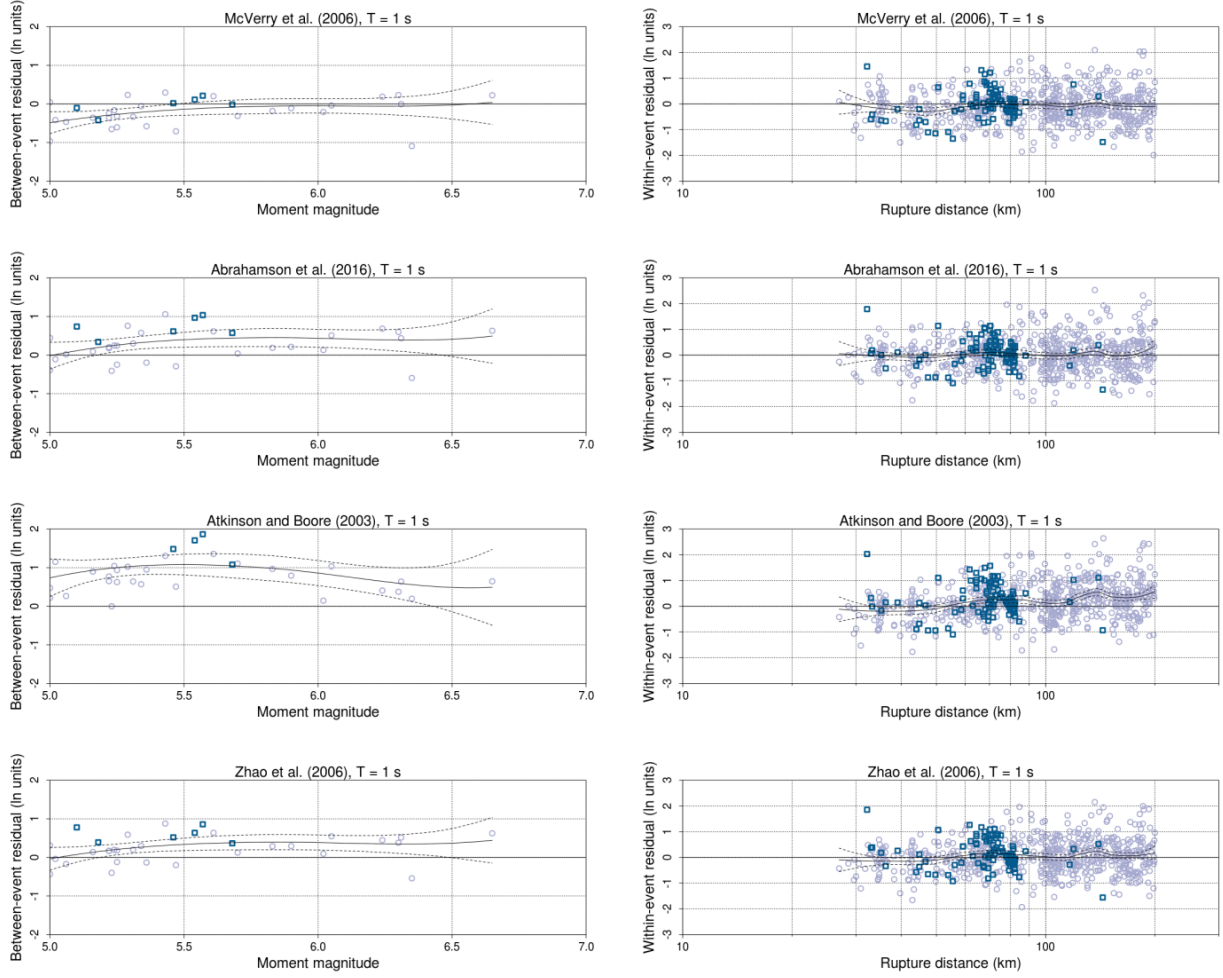
Figure 6: Model within-event residuals for shallow crustal events against NZS1170.5 Site Classifications, for PGA. The global models do not use NZS1170.5 site class as a predictor, but the residuals are calculated using the models' V_{S30} and Z_1 site response models before comparing these residuals against NZS1170.5 site class.

Figure 7: Residuals for the four subduction models considered in this study, with respect to the Hikurangi subduction zone spectral acceleration dataset at PGA. Squares and circles represent data from interface and slab events respectively.

Figure 8: Residuals for the four subduction zone models considered in this study, with respect to the Hikurangi zone spectral acceleration dataset at T = 1 s. Squares and circles represent data from interface and slab events respectively.

modified model still shows a strong overprediction with respect to class A sites. However, this result has limited interpretability, due to the lack of availability of measured site parameters for any class A sites in New Zealand [7].

Subduction Zone Models

Behaviour at PGA

Figure 7 shows how the four subduction zone models represent Hikurangi PGA data, with squares representing recordings from interface events and circles representing recordings from events within the subducted slab. For slab events, the Abrahamson *et al.* (2016) and Zhao *et al.* (2006) models have little bias in the between-event residual against M_w, while the McVerry *et al.* (2006) model overpredicts the small magnitudes, and the Atkinson and Boore (2003) model underpredicts nearly all of the Hikurangi slab data. For the interface events, it is difficult to make any inferences on model performance, given that the magnitude range of interest is primarily greater than 8.

For the within-event residuals, the four models show little bias with respect to distance for the slab events. There may be indications that the four models are not capturing systematic path effects for the Hikurangi interface data, although the data are too few to conclude this with any certainty.

Behaviour at T = 1 s

For spectral accelerations at *T* = 1 s, Figure 8 shows that all models except for Atkinson and Boore (2003) are able to reasonably represent the Hikurangi slab data. One event of M_w

= 6.35 is significantly overpredicted by the models, but this event had a focal depth greater than 150 km, so is beyond the models' range of applicability. It is again difficult to make conclusions on the model performance for interface events, with such few data at large magnitudes.

With respect to distance, all models significantly underestimate one data point at R_{rup} = 32 km, which corresponds to the Waipawa recording of the 1993 Tikokino earthquake. It is unlikely that this can be attributed to a site effect, because a collective analysis of all recordings at this site does not reveal a systematically-large positive residual. As such, it may be that the path terms of the subduction models are not entirely appropriate for predicting the distance-dependence of ruptures from the Hikurangi subduction zone, although it is difficult to make this conclusion from a single data point and further investigation is necessary.

Summary

Many plots similar to Figures 3-8 can be made, against a variety of model predictors and for all spectral periods. However, they become increasingly difficult to interpret, particularly for the crustal models and their strong observed biases against M_w and R_{rup}. Therefore, the following section addresses the performance of models in a more objective way, by assessing the overall fit of the models to the New Zealand data.

EVALUATION OF MODEL PERFORMANCE

Mean Bias

To summarise the overall behaviour of the six crustal models against the New Zealand dataset, Figure 9a plots the mean total residual for all models, against period. Of the two New Zealand equations, the McVerry *et al*. (2006) equation overpredicts the data on average across all periods. Comparing to Figures 3-5, this is a general reflection of the McVerry et al. (2006) equation overpredicting the small magnitudes. The Bradley (2013) model underpredicts on average at mid-range and very long periods, but is generally unbiased at the other periods. Overall, the bias of the international models is of similar order to the Bradley (2013) model, but slightly larger. However, it should be noted that on average, all models underpredict the data for T ≥ 3 s and overpredict the data for *T* ≤ 0.2 s.

Figure 9b shows the mean total residual against period for the four subduction zone models considered in this study. The Abrahamson *et al*. (2016) and Zhao *et al*. (2006) models have the least bias across all periods with respect to the Hikurangi dataset. The McVerry *et al*. (2006) and Atkinson and Boore (2003) models overpredict and underpredict the data respectively, across all their predicted periods.

LLH Values

While the mean bias gives an overview of the median model performance, another informative way of comparing models is using the negative average sample log-likelihood (–LLH) value, as a proxy for the KL divergence, D_{KL}. Period-dependent –LLH values are calculated for each model using equation (4). To avoid sampling bias from well-recorded events, the –LLH values in this section are derived using a bootstrap procedure. 200 sets of random samples are taken from the database with replacement, and the mean and standard deviation of the –LLH value is calculated. To check that the computations of the parameters are stable, the database size is divided in two, and then in four, before the –LLH is recalculated. A similar procedure was adopted by Edwards and Douglas (2013) [57].

Table 1 shows the crustal model –LLH values for PGA, T = 1 s and T = 5 s, for the three bootstrapped databases. Table 2 is similar to Table 1, but is for the subduction zone models. The mean –LLH values are very stable, and dividing the dataset

into four has almost no influence. The standard deviation of the –LLH increases as the size of the dataset decreases, however it still remains reasonably low. As such, the –LLH values derived here can be considered robust. Figures 10a and 10b show the –LLH values against period for the six crustal models and four subduction zone models respectively. As the –LLH value can be thought of as a measure of distance, the smaller the –LLH, the closer the model to the observed data.
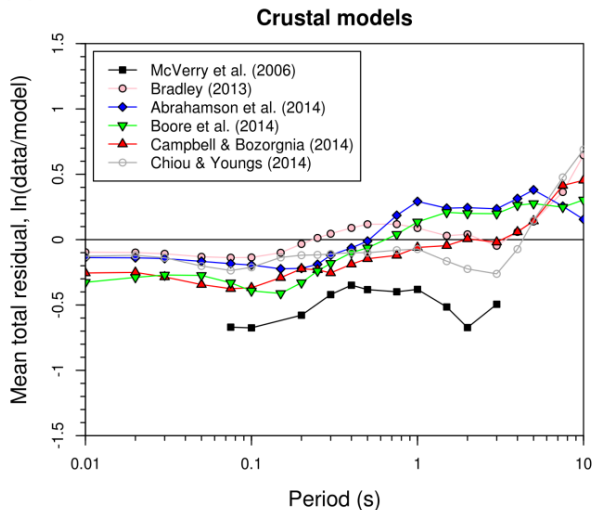
Of the six crustal models, the McVerry *et al*. (2006) has the largest –LLH value across all of its predicted periods. The Bradley (2013) model has the smallest –LLH values across the short periods, and still performs well at all but the longest periods. For the mid-range periods, the Campbell and Bozorgnia (2014) model is generally the closest to the New Zealand data, while for the longest periods the Abrahamson et al. (2014) model is closest to the data. The four NGA-W2 and the Bradley (2013) models are all reasonably close to the New Zealand data.

For the subduction zone models, Abrahamson *et al*. (2016) and Zhao *et al*. (2006) are inseparable as the closest models to the Hikurangi dataset. The McVerry *et al*. (2006) model is further from the New Zealand dataset, but is closer than the Atkinson and Boore (2003) model for periods less than 2 s. While there are no relevant data that can inform model performance for a Hikurangi mega-thrust event, this plot does indicate that the Abrahamson *et al*. (2016) and Zhao *et al*. (2006) models might be good representations of ground motion from events within the subducting Pacific Plate.
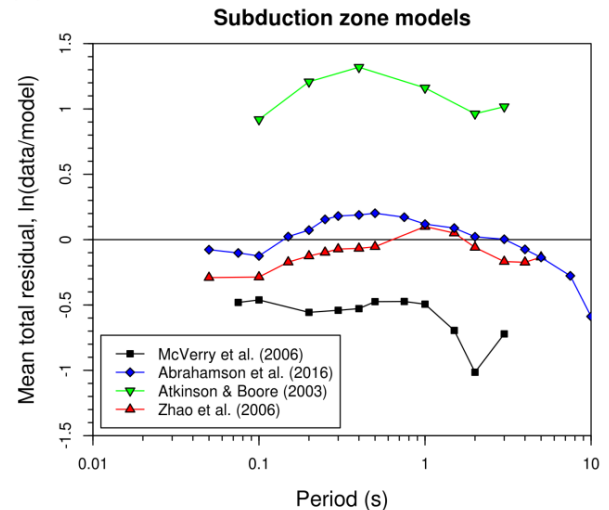
Effect of Data Selection

The results of this analysis are strongly dependent on the subset of data from which the performance metrics are calculated. For example, Figure 11 plots the mean bias and –LLH values of the crustal models, for a subset of the crustal database that is most relevant for PSHA, where all data is from an event with $M_W \geq 6$ and the recordings are at rupture distances $R_{rup} < 100$ km. The results are quite different for this subset of the database. The McVerry *et al*. (2006) model becomes one of the closer models to the data at mid-range oscillator periods. The Bradley (2013) model is still one of the best performing models at short periods, but becomes one of the more distant at periods of most engineering interest, and underpredicts the data on average by 20-40%. Of the global models, the Campbell and Bozorgnia (2014) model appears to generally be the closest model, except for *T* > 5 s.

(a)

(b)

Figure 9: The mean bias of (a) the six crustal ground-motion models against the New Zealand data and (b) the four subduction zone models against data associated with the Hikurangi subduction zone. In both, plots positive values represent underprediction and negative values represent overprediction.
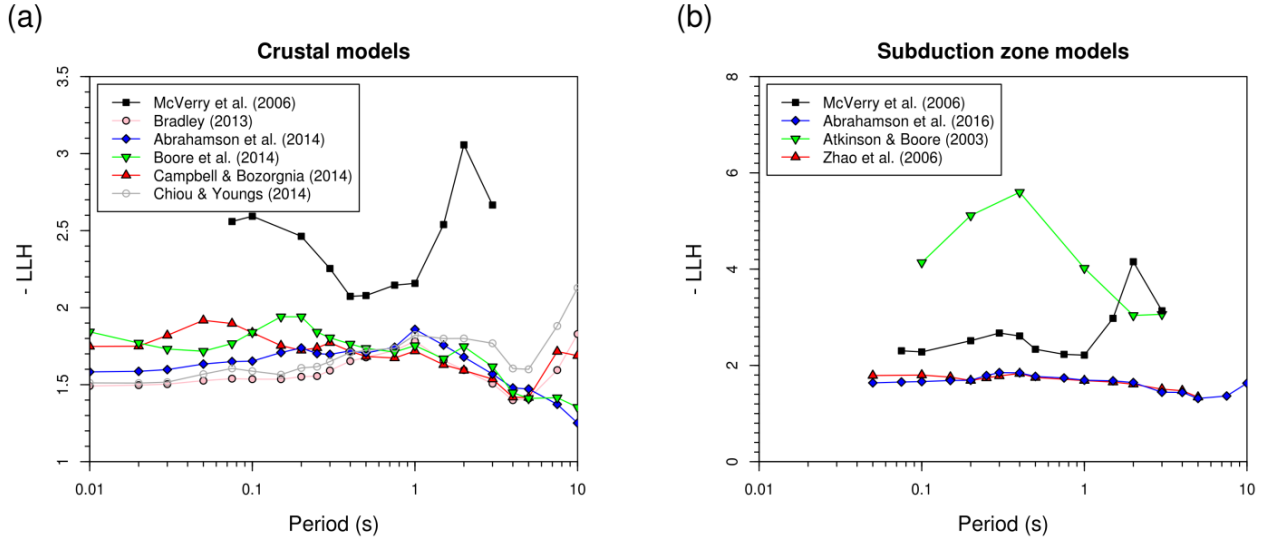
(a)

(b)

Figure 10: The mean -LLH of the (a) six considered crustal ground-motion models and (b) four considered subduction zone models, against the bootstrapped New Zealand database. The smaller the –LLH value, the closer the model is to the data.
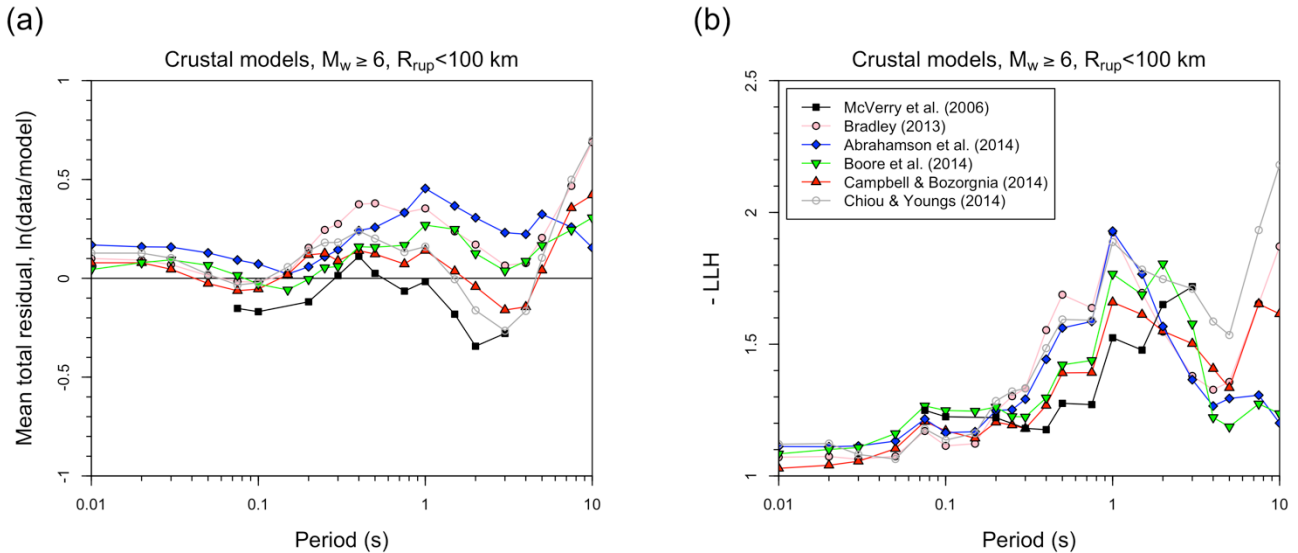
Table 1: Stability of the –LLH values calculated for the six crustal models.

Crustal model	Period	-LLH (all data)	-LLH (half of dataset)	-LLH (quarter of dataset)	Rank
McVerry *et al*. (2006)	PGA	2.20 ± 0.07	2.22 ± 0.10	2.22 ± 0.16	6
	1 s	2.11 ± 0.07	2.13 ± 0.10	2.12 ± 0.13	6
	5 s	-	-	-	-
Bradley (2013)	PGA	1.48 ± 0.04	1.50 ± 0.05	1.49 ± 0.08	1
	1 s	1.79 ± 0.05	1.79 ± 0.06	1.79 ± 0.08	3
	5 s	1.43 ± 0.02	1.44 ± 0.04	1.44 ± 0.06	3
Abrahamson et al. (2014)	PGA	1.57 ± 0.04	1.58 ± 0.05	1.57 ± 0.07	3
	1 s	1.88 ± 0.04	1.87 ± 0.06	1.87 ± 0.08	5
	5 s	1.47 ± 0.03	1.48 ± 0.04	1.47 ± 0.06	4
Boore *et al*. (2014)	PGA	1.82 ± 0.05	1.84 ± 0.08	1.84 ± 0.11	5
	1 s	1.77 ± 0.04	1.77 ± 0.06	1.77 ± 0.07	2
	5 s	1.42 ± 0.03	1.42 ± 0.04	1.42 ± 0.06	1
Campbell & Bozorgnia (2014)	PGA	1.73 ± 0.05	1.74 ± 0.07	1.73 ± 0.11	4
	1 s	1.73 ± 0.04	1.73 ± 0.05	1.73 ± 0.06	1
	5 s	1.43 ± 0.03	1.43 ± 0.04	1.44 ± 0.06	2
Chiou & Youngs (2014)	PGA	1.51 ± 0.04	1.52 ± 0.05	1.51 ± 0.07	2
	1 s	1.84 ± 0.04	1.84 ± 0.06	1.84 ± 0.07	4
	5 s	1.62 ± 0.03	1.63 ± 0.05	1.62 ± 0.07	5

Table 2: Stability of the –LLH values calculated for the four subduction zone models.

Subduction zone model	Period	-LLH (all data)	-LLH (half of dataset)	-LLH (quarter of dataset)	Rank
McVerry *et al*. (2006)	PGA	2.32 ± 0.13	2.34 ± 0.19	2.34 ± 0.25	3
	1 s	2.15 ± 0.08	2.17 ± 0.11	2.15 ± 0.15	3
Abrahamson *et al*. (2006)	PGA	1.57 ± 0.04	1.57 ± 0.05	1.58 ± 0.08	1
	1 s	1.68 ± 0.04	1.68 ± 0.06	1.68 ± 0.09	2
Atkinson & Boore (2003)	PGA	4.56 ± 0.22	4.52 ± 0.36	4.57 ± 0.46	4
	1 s	4.03 ± 0.14	4.03 ± 0.19	4.05 ± 0.28	4
Zhao *et al*. (2006)	PGA	1.67 ± 0.05	1.69 ± 0.08	1.67 ± 0.11	2
	1 s	1.67 ± 0.04	1.67 ± 0.06	1.66 ± 0.09	1

Figure 11: The mean bias (a) and –LLH values (b) of the six crustal ground-motion models against a larger magnitude, shorter distance subset of the New Zealand database, with all recordings from events with $M_w \geq 6$ and having rupture distances $R_{rup} < 100$ km.

RECOMMENDATIONS FOR HAZARD ASSESSMENT IN NEW ZEALAND

Figures 9, 10 and 11 show that in general, the global models tend to represent the New Zealand data just as well as the New Zealand specific Bradley (2013) model, particular at the long periods, and better than the McVerry et al. (2006) model at all periods. This result shows that it is justifiable, at least in the short term, to use the NGA-W2 crustal models and global subduction zone models in New Zealand seismic hazard analyses.

For PSHA, the typical next step is to derive a logic tree that accounts for the epistemic uncertainty in the model predictions. The use of logic trees in ground motion characterisation is a highly controversial topic that has been heavily debated over the last decade [70-77]. This section attempts to provide guidance to seismic hazard analysts for a ground-motion model logic tree in New Zealand PSHA.

It is acknowledged that obtaining consensus on a ground-motion logic tree is very difficult, hence the subsequent content of this section only represents the opinion of this author. The recommendation in this section is not intended to be strictly applied in all site-specific seismic hazard studies in New Zealand, however the proposed logic tree may be considered a null hypothesis position that can be altered as required.

Crustal Models

Model Selection

The first challenge in logic tree development is the selection of candidate models. Cotton *et al*. (2006) [78], and later Bommer *et al*. (2010) [79] provide a set of performance criteria to include, or exclude, ground motion models for PSHA. One criterion common to both studies, the exclusion of models that have been superseded by more recent publications, has particular importance for New Zealand PSHA. Recall that the McVerry *et al*. (2006) model is based heavily on the form of Abrahamson and Silva (1997), which has since been superseded by Abrahamson and Silva (2008) [80] and now Abrahamson *et al*. (2014). The same applies for Bradley

(2013), which is very similar to the Chiou and Youngs (2008) and Chiou *et al*. (2010) models that have now been superseded by Chiou and Youngs (2014). While it may be argued that the models are optimised to different datasets and hence are independent of their base models, it should be noted that many of the most important parameters of the New Zealand models are fixed to be the same as their superseded base model, for example the coefficients that control large magnitude scaling, near-source distance saturation, hanging wall effects and nonlinear soil response.

The issue with including a superseded model is that the superseded and updated models tend to be quite similar. Figures 9a and 10a show that there are similarities between the Bradley (2013) and Chiou and Youngs (2014) models. Examination of the models' forms shows that there are some model terms that are unchanged from the Chiou and Youngs (2008) to the Chiou & Youngs (2014) model. The logic tree framework requires each branch to be independent, which should be adhered to at least in a practical sense. However, in this case the Bradley (2013) and Chiou and Youngs (2014) models cannot be considered practically independent. By having duplicate models in a logic tree, the weight to that model's form is disproportionately increased, hence including both the Bradley (2013) and Chiou & Youngs (2014) models is controversial. The Bradley (2013) and Chiou and Youngs (2014) models perform almost identically at short periods, and the better performing model at longer periods depends on the period of interest. It is therefore difficult to select the preferred model of the two. If both models are to be included in PSHA, care needs to be taken to ensure that these models are weighted appropriately. This is addressed further in the 'model weights' section.

The McVerry *et al*. (2006) model is slightly more independent of Abrahamson *et al*. (2014), particularly the version where the site effects are modelled as a continuous function of the fundamental site period [26]. However, it greatly overpredicts the small magnitudes, primarily due to a lack of magnitudes less than 5.75 in the database from which it was derived. Bommer *et al*. (2010) [79] recommend excluding models where the range of applicability of the model is too small to be useful in the range required by PSHA, which in the NSHM is a minimum magnitude of 5.25 [2]. If this criterion is to be

strictly applied, the McVerry *et al.* (2006) model should be removed from the NSHM. It is my opinion that this criterion can be relaxed if, and only if, there is quantification of the uncertainty in a model's parameters due to limited data. This uncertainty can then be included as additional epistemic uncertainty, using the method detailed in Al Atik & Youngs (2014) [81], and would reflect that the uncertainty is much larger at M5 than at M6. Should this information become available for the McVerry *et al.* (2006) model, then it can be included on a ground-motion model logic tree with the increased epistemic uncertainty, but without it, in my opinion it should be assigned zero weight.

While the inclusion of global models will strongly reduce the influence of the two New Zealand specific models, both New Zealand models have some useful behaviour that is not well constrained by the data. In particular, both the McVerry *et al.* (2006) and the Bradley (2013) models have terms to represent the strong attenuation in the Taupo Volcanic Zone (TVZ). The TVZ attenuation terms of the two New Zealand models are reasonably similar, hence either term (or a weighted combination of both) can be attached to the four NGA-W2 models for modelling crustal sources around the TVZ. This will allow the NGA-W2 models to be used for estimating the seismic hazard from TVZ sources. Likewise, the Christchurch-specific ground-motion model modifications proposed by Bradley (2015) improve ground-motion modelling with respect to Canterbury data. These modifications are specific to the Bradley (2013) model. While similar Christchurch-specific modifications to the NGA-W2 models are not available in the literature, the public availability of the New Zealand Strong Motion Database [6] enables reasonably straight-forward derivation of similar modifications for the NGA-W2 models.

The four NGA-W2 models satisfy all the criteria of Cotton *et al.* (2006) and Bommer *et al.* (2010), and models for their coefficient uncertainties are quantified in Al Atik & Youngs (2014) [81]. As such, the four NGA-W2 models are considered of sufficient quality to include on a logic tree for New Zealand PSHA.

Model Weights

Assigning objective weights to the ground motion models is fraught with difficulty. The process is somewhat aided by the –LLH values illustrated in Figure 10a. As discussed by Scherbaum *et al.* (2009) [53], logic tree weights can in principle be directly derived from the –LLH values illustrated in Figure 10a, using the following formula:

$$w_i = \frac{2^{-LLH_i}}{\sum_{i=1}^{I} 2^{-LLH_i}}$$
, (8)

where w_i is the KL weight for model i. However, this criterion is difficult to apply, given that the Bradley (2013) and Chiou and Youngs (2014) models are closely related. Additionally, the –LLH weighting approach has been criticised because, by normalising by the number of samples in equation (4), the weights do not converge to the 'true model' with increasing evidence [82]. A Bayesian framework may therefore be preferable, where prior weight distributions are updated by the model likelihood function to obtain posterior weight distributions. However, the posterior weights are highly dependent on the subjective definition of the prior distribution.

Internationally, there are different approaches for assigning model weights. For the recent European Seismic Hazard Model, the KL weights are considered a starting point only, with final model weights determined by expert elicitation [55, 83]. In the Western United States, equal weights are applied to the five NGA-W2 models [81]. For New Zealand, the best way forward might be to formally engage international and local experts to collectively derive some consensus model weights, or to collectively define a consensus prior distribution for the weights.

Unfortunately, this recommendation does not immediately address the current issue, where seismic hazard assessments are performed in New Zealand using only one or two ground-motion models. I therefore propose an interim solution. Until consensus model weights are derived, my opinion is that the Abrahamson et al. (2014), Boore et al. (2014) and Campbell and Bozorgnia (2014) models should be assigned equal weights of 0.25, with the remaining 0.25 being shared equally between the Chiou and Youngs (2014) and Bradley (2013) models. In addition, the epistemic uncertainty model proposed by Al Atik and Youngs (2014) [81], which accounts for the epistemic uncertainty that arises from the limited dataset, should be applied. This consideration of epistemic uncertainty beyond the between-model uncertainty acknowledges that less is known about large magnitudes events, given the global scarcity of large magnitude data. While the Al Atik and Youngs (2014) epistemic variance model doesn't directly apply to the Bradley (2013) model, nor is it directly applicable for New Zealand conditions, in my opinion it is an acceptable approximation until a more robust New Zealand specific model is derived.

This recommendation does not immediately apply to the Canterbury seismic hazard model [35], where a ground motion model logic tree developed from expert elicitation already exists. However, it is nevertheless recommended that hazard analysts consider deriving Canterbury-specific modifications to the four NGA-W2 models using the New Zealand Strong Motion Database, and including these on a ground-motion logic tree. This would allow greatly improved PSHA for Canterbury, while still adequately managing and quantifying the epistemic uncertainty. In the interim, the published logic tree for the Canterbury seismic hazard model, as modified to include the work of Bradley (2015) [34], should be adopted.

Subduction Zone Models

Given that there are relatively few global subduction zone models, and also given the sparse Hikurangi dataset, it is difficult find compelling evidence for excluding subduction zone models in New Zealand PSHA. The Atkinson and Boore (2003) model is furthest from the New Zealand dataset, but the dataset is comprised of predominantly small magnitude events and does not evaluate the model's behaviour at magnitudes of interest in hazard. Stewart *et al.* (2015) [1] recommend that Abrahamson *et al.* (2016), Atkinson and Boore (2003) and Zhao *et al.* (2006) models be used in PSHA globally, although the authors acknowledge that the inclusion of the Atkinson & Boore (2003) model is controversial. Given that the Hikurangi dataset is insufficient to evaluate the models' behaviour where it is of most interest for hazard, I recommend adopting the Stewart *et al.* (2015) model set, but with the addition of the McVerry *et al.* (2006) model, given its reasonable fit to the Hikurangi dataset, its special consideration for Hikurangi-specific deep slab events [25], and the ability to consider information from Hikurangi subduction zone simulations in the hazard calculations [27].

An additional recommendation of the Stewart *et al.* (2015) study is to replace the linear site terms of Zhao *et al.* (2006) with the Abrahamson *et al.* (2016) site terms, to consider nonlinear soil response. However, applying the alternative site terms to the Zhao *et al.* (2006) model led to a poorer fit with respect to the Hikurangi dataset, hence it may be preferable to retain the Zhao *et al.* (2006) model in its published form.

The Stewart *et al*. (2015) study only provides recommendations on model selection, hence model weights must be determined by the hazard analyst. Given that, even globally, there are few data at the magnitudes of interest to New Zealand PSHA (up to M_w9), defining weights on subduction zone models is a difficult task. For subducted slab events, a proposed interim solution is equal weights on the four subduction zone models considered in this study. For Hikurangi subduction interface events, the standard framework of assigning model weights may not be appropriate, as the four selected models are unlikely to collectively capture the full range of epistemic uncertainty. Some guidance on potential alternative frameworks, based on recent international experience, is available in Abrahamson *et al*. (2014) [84].

DISCUSSION AND CONCLUSION

Although the modern view is that it is ill-advised, New Zealand has a history of only using a single ground-motion model in seismic hazard analysis, thereby ignoring the epistemic uncertainty [2-4]. This article recommends that the global NGA-W2 models and the Bradley (2013) model should be incorporated into New Zealand PSHA, applied with equal weights and the additional epistemic variance model of Al Atik and Youngs (2014). This logic tree is not expected to be applicable to all site-specific hazard analyses in New Zealand, but represents a null hypothesis position that can be deviated from if there is sufficient evidence that it is inappropriate. Given that there is a significant research effort underway to better account for epistemic uncertainty in New Zealand PSHA, the life expectancy of this recommendation is likely to be short. While a more robust epistemic uncertainty model will be developed in the near future, the purpose of this article is to improve New Zealand seismic hazard analysis in the interim, by better acknowledging the epistemic uncertainty in ground motion modelling. In Canterbury, the logic tree of Gerstenberger *et al*. (2014) should be adopted, modified to allow use of the work of Bradley (2015) [34], although the inclusion of global models should also be considered.

Some additional consideration will be necessary for calculating hazard at long periods. As illustrated in Figures 9a and 11a, all crustal models consistently underpredict the New Zealand dataset at periods greater than four seconds. In this case, no weighted combination of the models will be a good representation of New Zealand data. Instead, a scaled backbone approach [77] to epistemic uncertainty is likely to be a more appropriate method for quantifying the long period seismic hazard in New Zealand. This approach should be considered for site-specific hazard assessment for tall buildings and base-isolated structures.

For subduction zone models, the New Zealand database cannot inform model selection for predicting the ground motion from a Hikurangi mega-thrust event. However for modelling ground motion from events in the subducted slab, in my opinion the McVerry *et al*. (2006), Abrahamson *et al*. (2016), Atkinson and Boore (2003) and Zhao *et al*. (2006) models should be used with equal weights.

ACKNOWLEDGMENTS

REFERENCES

1   Stewart J, Douglas J, Javanbarg M, Bozorgnia Y, Abrahamson N, Boore D, Campbell K, Delavaud E, Erdik M and Stafford P (2015). "Selection of ground motion prediction equations for the Global Earthquake Model". *Earthquake Spectra,* **31**(1): 19-45.

2   Stirling M, McVerry G, Gerstenberger M, Litchfield N, Van Dissen R, Berryman K, Barnes P, Wallace L, Villamor P, Langridge R, Reyners M, Bradley B, Rhoades D, Smith W, Nicol A, Pettinga J, Clark K and Jacobs K (2012). "National seismic hazard model for New Zealand: 2010 update". *Bulletin of the Seismological Society of America,* **102**(4): 1514-1542.

3   Skarlatoudis A, Somerville P and Bayless J (2015). "Probabilistic Response Spectra for Christchurch CBD Ground Motions Incorporating Amplification Factors Derived from the 2010-2011 Canterbury Earthquake Sequence". *In Proceedings of 6th International Conference on Earthquake Geotechnical Engineering*, Christchurch, New Zealand, 1-4 November 2015, Paper number 367.

4   Bradley B (2015). "Benefits of site-specific hazard analyses for seismic design in New Zealand". *Bulletin of the New Zealand Society for Earthquake Engineering,* **48**(2): 92-99.

5   Kulkarni R, Youngs R and Coppersmith K (1984). "Assessment of confidence intervals for results of seismic hazard analysis". *In Proceedings of Eighth World Conference on Earthquake Engineering*, San Francisco, United States, 21-28 July, Volume 1B.

6   Van Houtte C, Bannister S, Holden C, Bourguignon S and McVerry G (2017). "The New Zealand strong motion database". *Bulletin of the New Zealand Society for Earthquake Engineering,* **50**(1): 1-20.

7   Kaiser A, Van Houtte C, Perrin N, Wotherspoon L and McVerry G (2017). "Site characterisation of GeoNet stations for the New Zealand strong motion database". *Bulletin of the New Zealand Society for Earthquake Engineering,* **50**(1): 39-49.

8   Douglas J (2011). "*Ground-motion prediction equations 1964-2010*". PEER Report 2011/102, Pacific Earthquake Engineering Research Center, Berkeley, United States, 442 p.

9   Matuschka T (1980). "*Assessment of seismic hazards in New Zealand*". Department of Civil Engineering Report No. 222, University of Auckland, Auckland, New Zealand, 214 p.

10   Katayama T, Iwasaki T and Seaki M (1978). "Statistical analysis of earthquake acceleration response spectra (in Japanese)". *Proceedings of the Japanese Society of Civil Engineers,* **275**: 29-40.

11   Katayama T (1982). "An engineering prediction model of acceleration response spectra and its application to seismic hazard mapping". *Earthquake Engineering and Structural Dynamics,* **10**: 149-163.

12   Peek R (1980). "*Estimation of seismic risk for New Zealand*". Department of Civil Engineering Research Report 80-21, University of Canterbury, Christchurch, New Zealand, 84 p.

13   Mulholland W (1982). "*Estimation of design earthquake motion for New Zealand*". Department of Civil

Engineering Research Report 82-9, University of Canterbury, Christchurch, New Zealand, 97 p.

14 McVerry G (1986). "Uncertainties in attenuation relations for New Zealand seismic hazard analysis". *Bulletin of the New Zealand National Society for Earthquake Engineering,* **19**(1): 28-39.

15 Matuschka T, Berryman K, O'Leary A, McVerry G, Mulholland W and Skinner R (1985). "New Zealand seismic hazard analysis". *Bulletin of the New Zealand National Society for Earthquake Engineering,* **18**(4): 313-322.

16 Standards New Zealand (1992). "*NZS 4203:1992 - Code of practice for general structural design and design loadings for buildings".* Wellington, New Zealand.

17 Matuschka T and Davis B (1991). "Derivation of an attenuation model in terms of spectral acceleration for New Zealand". *In Proceedings of Pacific Conference on Earthquake Engineering*, Auckland, New Zealand, 20-23 November 1991, Volume 2.

18 Zhao J, Dowrick D and McVerry G (1997). "Attenuation of peak ground accelerations in New Zealand earthquakes". *Bulletin of the New Zealand National Society for Earthquake Engineering,* **30**(2): 133-158.

19 Cousins W, Zhao J and Perrin N (1999). "A model for the attenuation of peak ground acceleration in New Zealand earthquakes based on seismograph and accelerograph data". *Bulletin of the New Zealand National Society for Earthquake Engineering,* **32**(4): 193-220.

20 McVerry G, Zhao J, Abrahamson N and Somerville P (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.

21 Abrahamson N and Silva W (1997). "Empirical response spectral attenuation relations for shallow crustal earthquakes". *Seismological Research Letters,* **68**(1): 94-127.

22 Youngs R, Chiou B, Silva W and Humphrey J (1997). "Strong ground motion attenuation relationships for subduction zone earthquakes". *Seismological Research Letters,* **68**(1): 58-73.

23 Stirling M, Wesnousky S and Berryman K (1998). "Probabilistic seismic hazard analysis of New Zealand". *New Zealand Journal of Geology and Geophysics,* **41**(4): 355-375.

24 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*, Auckland, New Zealand, 30 January - 4 February 2000, Paper Number 1834.

25 Eberhart-Phillips D and McVerry G (2003). "Estimating slab earthquake response spectra from a 3D Q model". *Bulletin of the Seismological Society of America,* **93**(6): 2649-2663.

26 McVerry G (2011). "Site-effect terms as continuous functions of site period and Vs30". *In Proceedings of Ninth Pacific Conference on Earthquake Engineering*, Auckland, New Zealand, 14-16 April 2011, Paper Number 10.

27 McVerry G and Holden C (2014). "Incorporating simulated Hikurangi subduction interface spectra into probabilistic hazard calculations for Wellington". *In Proceedings of Conference of the New Zealand Society of Earthquake Engineering*, Auckland, New Zealand, 21-23 March 2014, Paper Number 054.

8 Arias A (1970). "*Measure of earthquake intensity*", Seismic Design for Nuclear Power Plants, Massachusetts Institute of Technology, Cambridge, United States, 438-483 p.

29 Stafford P, Berrill J and Pettinga J (2009). "New predictive equations for Arias intensity from crustal earthquakes in New Zealand". *Journal of Seismology,* **13**(1): 31-52.

30 Stafford P (2006). "*Engineering seismological studies and seismic design criteria for the Buller region, South Island, New Zealand".* (PhD Thesis), University of Canterbury, Christchurch, New Zealand, 342 p.

31 Bradley B (2013). "A New Zealand-Specific Pseudospectral Acceleration Ground-Motion Prediction Equation for Active Shallow Crustal Earthquakes Based on Foreign Models". *Bulletin of the Seismological Society of America,* **103**(3): 1801-1822.

32 Chiou B, Youngs R, Abrahamson N and Addo K (2010). "Ground-motion attenuation model for small-to-moderate shallow crustal earthquakes in California and its implications on regionalization of ground-motion prediction models". *Earthquake Spectra,* **26**(4): 907-926.

33 Chiou B and Youngs R (2008). "An NGA model for the average horizontal component of peak ground motion and response spectra". *Earthquake Spectra,* **24**(1): 173-215.

34 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.

35 Gerstenberger M, McVerry G, Rhoades D and Stirling M (2014). "Seismic hazard modeling for the recovery of Christchurch". *Earthquake Spectra,* **30**(1): 17-29.

36 Bozorgnia Y, Abrahamson N, Al Atik L, Ancheta T, Atkinson G, Baker J, Baltay A, Boore D, Campbell K and Chiou B (2014). "NGA-West2 research project". *Earthquake Spectra,* **30**(3): 973-987.

37 Abrahamson N, Silva W and Kamai R (2014). "Summary of the ASK14 ground motion relation for active crustal regions". *Earthquake Spectra,* **30**(3): 1025-1055.

38 Boore D, Stewart J, Seyhan E and Atkinson G (2014). "NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes". *Earthquake Spectra,* **30**(3): 1057-1085.

39 Campbell K and Bozorgnia Y (2014). "NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra". *Earthquake Spectra,* **30**(3): 1087-1115.

40 Chiou B and Youngs R (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.

41 Idriss I (2014). "An NGA-West2 empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes". *Earthquake Spectra,* **30**(3): 1155-1177.

42 Abrahamson N, Gregor N and Addo K (2016). "BC Hydro ground motion prediction equations for subduction earthquakes". *Earthquake Spectra*, **32**(1): 23-44.

43 Atkinson G and Boore D (2003). "Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions". *Bulletin of the Seismological Society of America,* **93**(4): 1703-1729.

44 Atkinson G and Boore D (2008). "Erratum to empirical ground-motion relations for subduction zone earthquakes and their application to Cascadia and other regions".

Bulletin of the Seismological Society of America, **98**(5): 2567-2569.

45 Zhao JX, Zhang J, Asano A, Ohno Y, Oouchi T, Takahashi T, Ogawa H, Irikura K, Thio H and Somerville P (2006). "Attenuation relations of strong ground motion in Japan using site classification based on predominant period". *Bulletin of the Seismological Society of America,* **96**(3): 898-913.

46 Scherbaum F, Cotton F and Smit P (2004). "On the use of response spectral-reference data for the selection and ranking of ground-motion models for seismic-hazard analysis in regions of moderate seismicity: The case of rock motion". *Bulletin of the Seismological Society of America,* **94**(6): 2164-2185.

47 Douglas J, Bertil D, Roullé A, Dominique P and Jousset P (2006). "A preliminary investigation of strong-motion data from the French Antilles". *Journal of Seismology,* **10**(3): 271-299.

48 Douglas J and Mohais R (2009). "Comparing predicted and observed ground motions from subduction earthquakes in the Lesser Antilles". *Journal of Seismology,* **13**(4): 577-587.

49 Allen T and Brillon C (2015). "Assessment of ground-motion models for use in the British Columbia North Coast region, Canada". *Bulletin of the Seismological Society of America,* **105**(2B): 1193-1205.

50 Stafford PJ, Strasser FO and Bommer JJ (2008). "An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region". *Bulletin of earthquake Engineering,* **6**(2): 149-177.

51 Scasserra G, Stewart J, Bazzurro P, Lanzo G and Mollaioli F (2009). "A comparison of NGA ground-motion prediction equations to Italian data". *Bulletin of the Seismological Society of America,* **99**(5): 2961-2978.

52 Allen T and Wald D (2009). "*Evaluation of ground-motion modeling techniques for use in Global ShakeMap - a critique of instrumental ground-motion prediction equations, peak ground motion to macroseismic intensity conversions, and macroseismic intensity predictions in different tectonic settings*", Open File Report 2009-1047, United States Geological Survey, 122 p.

53 Scherbaum F, Delavaud E and Riggelsen C (2009). "Model selection in seismic hazard analysis: An information-theoretic perspective". *Bulletin of the Seismological Society of America,* **99**(6): 3234-3247.

54 Beauval C, Tasan H, Laurendeau A, Delavaud E, Cotton F, Guéguen P and Kuehn N (2012). "On the testing of ground-motion prediction equations against small-magnitude data". *Bulletin of the Seismological Society of America,* **102**(5): 1994-2007.

55 Delavaud E, Scherbaum F, Kuehn N and Allen T (2012). "Testing the global applicability of ground-motion prediction equations for active shallow crustal regions". *Bulletin of the Seismological Society of America,* **102**(2): 707-721.

56 Mousavi M, Ansari A, Zafarani H and Azarbakht A (2012). "Selection of ground motion prediction models for seismic hazard analysis in the Zagros region, Iran". *Journal of Earthquake Engineering,* **16**(8): 1184-1207.

57 Edwards B and Douglas J (2013). "Selecting ground-motion models developed for induced seismicity in geothermal areas". *Geophysical Journal International,* **195**(2): 1314-1322.

58 Haendel A, Specht S, Kuehn N and Scherbaum F (2015). "Mixtures of ground-motion prediction equations as backbone models for a logic tree: an application to the subduction zone in Northern Chile". *Bulletin of Earthquake Engineering,* **13**(2): 483-501.

59 Standards New Zealand (2004). "*NZS1170.5 Earthquake actions - New Zealand".* Wellington, New Zealand.

60 Boore D, Watson-Lamprey J and Abrahamson N (2006). "Orientation-independent measures of ground motion". *Bulletin of the Seismological Society of America,* **96**(4A): 1502-1511.

61 Power M, Chiou B, Abrahamson N, Bozorgnia Y, Shantz T and Roblee C (2008). "An overview of the NGA project". *Earthquake Spectra,* **24**(1): 3-21.

62 Boore D (2010). "Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion". *Bulletin of the Seismological Society of America,* **100**(4): 1830-1835.

63 Bradley B and Baker J (2015). "Ground motion directionality in the 2010–2011 Canterbury earthquakes". *Earthquake Engineering & Structural Dynamics,* **44**(3): 371-384.

64 Eberhart-Phillips D and Reyners M (2001). "A complex, young subduction zone imaged by three-dimensional seismic velocity, Fiordland, New Zealand". *Geophysical Journal International,* **146**(3): 731-746.

65 Reyners M, Robinson R, Pancha A and McGinty P (2002). "Stresses and strains in a twisted subduction zone—Fiordland, New Zealand". *Geophysical Journal International,* **148**(3): 637-648.

66 Abrahamson N and Youngs R (1992). "A stable algorithm for regression analyses using the random effects model". *Bulletin of the Seismological Society of America,* **82**(1): 505-510.

67 Al Atik L and Abrahamson N (2010). "Nonlinear site response effects on the standard deviations of predicted ground motions". *Bulletin of the Seismological Society of America,* **100**(3): 1288-1292.

68 Stafford PJ (2015). "Extension of the Random-Effects Regression Algorithm to Account for the Effects of Nonlinear Site Response". *Bulletin of the Seismological Society of America,* **105**(6): 3196-3202.

69 Van Houtte C, Ktenidou O-J, Larkin T and Kaiser A (2012). "Reference stations for Christchurch". *Bulletin of the New Zealand Society for Earthquake Engineering,* **45**(4): 184-195.

70 Abrahamson N and Bommer J (2005). "Probability and uncertainty in seismic hazard analysis". *Earthquake Spectra,* **21**(2): 603-607.

71 McGuire R, Cornell CA and Toro G (2005). "The case for using mean seismic hazard". *Earthquake Spectra,* **21**(3): 879-886.

72 Musson R (2005). "Against fractiles". *Earthquake Spectra,* **21**(3): 887-891.

73 Bommer J and Scherbaum F (2008). "The use and misuse of logic trees in probabilistic seismic hazard analysis". *Earthquake Spectra,* **24**(4): 997-1009.

74 Scherbaum F and Kuehn NM (2011). "Logic tree branch weights and probabilities: summing up to one is not enough". *Earthquake Spectra,* **27**(4): 1237-1251.

75 Bommer J (2012). "Challenges of building logic trees for probabilistic seismic hazard analysis". *Earthquake Spectra,* **28**(4): 1723-1735.

76 Musson R (2012). "On the nature of logic trees in probabilistic seismic hazard assessment". *Earthquake Spectra,* **28**(3): 1291-1296.

77 Atkinson G, Bommer J and Abrahamson N (2014). "Alternative Approaches to Modeling Epistemic Uncertainty in Ground Motions in Probabilistic Seismic-Hazard Analysis". *Seismological Research Letters,* **85**(6): 1141-1144.

78 Cotton F, Scherbaum F, Bommer J and Bungum H (2006). "Criteria for selecting and adjusting ground-motion models for specific target regions: Application to central Europe and rock sites". *Journal of Seismology,* **10**(2): 137-156.

79 Bommer J, Douglas J, Scherbaum F, Cotton F, Bungum H and Fäh D (2010). "On the selection of ground-motion prediction equations for seismic hazard analysis". *Seismological Research Letters,* **81**(5): 783-793.

80 Abrahamson N and Silva W (2008). "Summary of the Abrahamson & Silva NGA ground-motion relations". *Earthquake Spectra,* **24**(1): 67-97.

81 Al Atik L and Youngs R (2014). "Epistemic uncertainty for NGA-West2 models". *Earthquake Spectra,* **30**(3): 1301-1318.

82 Arroyo D, Ordaz M and Rueda R (2014). "On the Selection of Ground-Motion Prediction Equations for Probabilistic Seismic-Hazard Analysis". *Bulletin of the Seismological Society of America,* **104**(4): 1860-1875.

83 Woessner J, Laurentiu D, Giardini D, Crowley H, Cotton F, Grünthal G, Valensise G, Arvidsson R, Basili R, Demircioglu M, Hiemer S, Meletti C, Musson R, Rovida A, Sesetyan K and Stucchi M (2015). "The 2013 European seismic hazard model: key components and results". *Bulletin of Earthquake Engineering,* **13**(12): 3553-3596.

84 Abrahamson N, Addo K, Atkinson G, Chiou B, Gregor N, Silva W and Youngs R (2014). "Ground motion characterization for the BC Hydro SSHAC Level 3 study". *In Proceedings of Tenth US National Conference on Earthquake Engineering, Frontiers of Earthquake Engineering*, Anchorage, Alaska, 21-25 July,

85 R Core Team (2015). "*A language and environment for statistical computing.*", R Foundation for Statistical Computing, Vienna, Austria, https://www.r-project.org/.

86 Pagani M, Monelli D, Weatherill G and Garcia J (2014). "*Testing procedures adopted in the development of the hazard component of the OpenQuake engine*". Global Earthquake Model (GEM) Technical Report 2014-09, Pavia, Italy, 73 p.

87 Global Earthquake Model (2014). *GEM GitHub oq-hazardlib.* https://github.com/gem/oq-hazardlib/tree/ master/openquake/hazardlib/tests/gsim/data. (Accessed December 2015).

88 Baker Research Group (2015). *Ground motion prediction equation functions.* http://web.stanford.edu/~bakerjw/ GMPEs.html. (Accessed December 2015).