Nationwide investigation of systematic site effects in New Zealand: Residual analysis of ground-motion simulations

A. Tiwari, B.A. Bradley, C.A. de la Torre & R.L. Lee
University of Canterbury, Christchurch.

ABSTRACT

This study examines systematic site effects from the prediction residuals of physics-based ground motion simulations using a dataset of small magnitude ($3.5 \leq M_W \leq 5.0$) active shallow crustal earthquakes recorded in New Zealand (NZ). A significant amount of total uncertainty in ground-motion modelling comes from within-event residuals, highlighting the need for a comprehensive study of the site characteristics that contribute to this uncertainty. A diverse range of sedimentary basins and sites in distinct geomorphic categories are considered in this study with the primary objective of improving physics-based ground motion simulations in NZ. Advancing ground-motion predictability through ground motion simulations is a continually iterative process and requires addressing fundamental questions like: Which geographic regions have predictions that significantly deviate from observations and why? Which predictor variables show dependence with the site-to-site residuals? This study examines these questions by classifying 212 NZ sites using Nweke et al. (2022) geomorphic categories and hierarchical clustering of site-to-site residuals. Using these categories and data-driven approaches, this study explores the geospatial variation of site-to-site residuals. The utilization of hierarchical clustering of site-to-site residuals in conjunction with geomorphic categorization of sites has facilitated the understanding of diverse shapes of site-to-site residuals throughout the country. Site-to-site residuals for the hill sites within the selected clusters of the country were primarily influenced by the relative elevations of the ground motion recording sites. Presently, an iterative process of utilizing site characterization and clustering is underway to gain insights into the underlying causes of site-specific biases in ground-motion modelling.

1 INVESTIGATION OF SYSTEMATIC SITE EFFECTS

Comprehensive validation of physics-based ground motion simulations is crucial to assess their predictive capability for use in seismic hazard assessment. Hence, this research study specifically seeks to answer: Can ground motion recording sites with strong systematic effects in the country (NZ) be identified, and can the attributes influencing these effects be determined to advance ground motion simulations?

Prediction residuals of ground motion simulations from Lee et al. (2022) are thoroughly examined in this study. The ground-motion database includes 5,218 ground motions from 479 earthquakes and 212 strong motion stations (Figure 1). The site conditions are quantified using site metrics, such as 30-m time-averaged shear wave velocity, V_{S30}, and fundamental site period, T_0, from Wotherspoon et al. (2022).

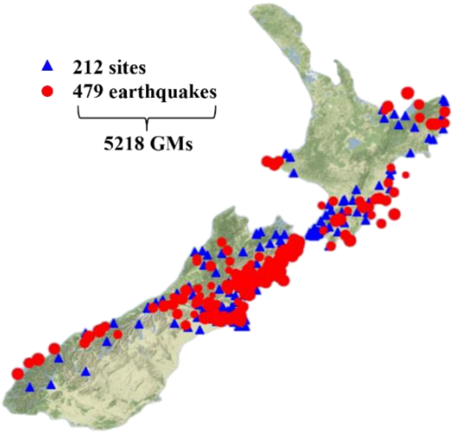


Figure 1: Nationwide investigation of systematic site effects for advancing physics-based ground motion simulations

The hybrid broadband ground-motion simulation methodology adopted in Lee et al. (2022) was developed by Graves and Pitarka (2015) and uses two different approaches for simulating the low- and high-frequency components (LF and HF, respectively). A LF-HF transition frequency of 1.0 Hz and a minimum shear wave velocity of 500 m/s was adopted. For the LF component, the NZ Velocity Model (NZVM), v2.02 (Thomson et al., 2020) was utilized, considering a finite difference grid spacing of 100 m. The prediction residuals were partitioned following the Al Atik et al. (2010) notation into various components associated with source, path, and site terms using mixed-effects regression. The total prediction residual, Δ, between the predicted ground motion from simulations and the observed ground motion is given by Equation (1). This can be further decomposed into fixed and random effects:

$$\Delta = ln\, IM_{es} - f_{es} = a + \delta B_e + \delta S2S_s + \delta W_{es}^0 \quad (1)$$

where ln IM_{es} is the natural logarithm of the observed intensity measure (IM) for earthquake e and site s, f_{es} is the mean of the predicted logarithmic IM (from ground motion simulations); a is the model bias, δB_e, $\delta S2S_s$, and δW_{es}^0 are the residuals with zero mean and variances τ^2, ϕ_{S2S}^2, and ϕ_{SS}^2 respectively.

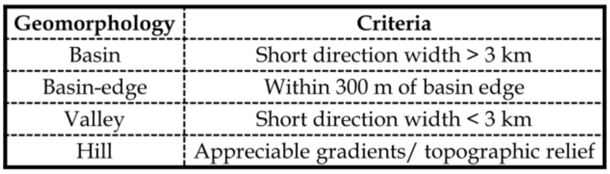
This work focuses on the nationwide identification of systematic site effects and the attributes influencing them. Hence, the study thoroughly scrutinizes the site-to-site residual, $\delta S2S_s$ for all NZ ground motion recording sites. This nationwide study examines systematic site effects through:

(a) Geomorphic classification of sites

(b) Hierarchical clustering of site-to-site residuals

2 GEOMORPHIC CLASSIFICATION OF SITES

Nweke et al. (2022) proposed a method, illustrated in Figure 2, for classifying the geomorphology of sedimentary basins to model site amplification for Southern California. In this study, this classification was applied in a NZ context, covering the 212 ground motion recording sites. Specifically, Tiwari et al. (2023) assessed systematic site effects for sites in the Canterbury and Wellington regions, which have multiple observed strong-motion records and higher quality site classification. In summary, the classification for these two regions resulted in 88 Basin, 20 Basin-edge, 14 Valley, and 49 Hill sites, as shown in Figure 3.



Geomorphology	Criteria
Basin	Short direction width > 3 km
Basin-edge	Within 300 m of basin edge
Valley	Short direction width < 3 km
Hill	Appreciable gradients/ topographic relief

Figure 2: Proposed geomorphic classification (Nweke et al., 2022)

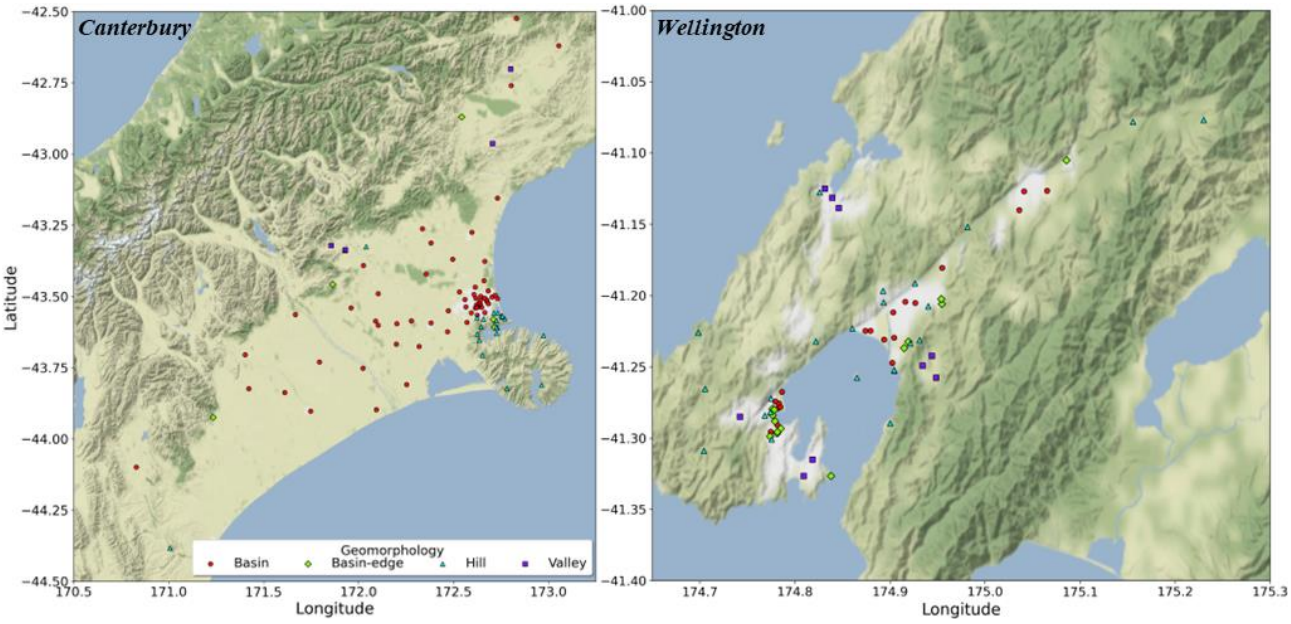


Figure 3: Categorization of stations based on geomorphology in the Canterbury and Wellington regions (Tiwari et al., 2023)

2.1 Canterbury region

Figure 4 (Tiwari et al., 2023) shows the corresponding site-to-site residuals for the two major geomorphic categories in the Canterbury region. The average $\delta S2S_S$ is almost zero at all periods for basin sites in the Canterbury region which is indicative of the high-resolution basin model used in these physics-based ground motion simulations and generally better site characterization (Lee et al. 2022).

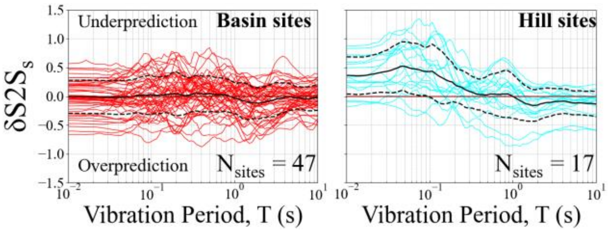


Figure 4: Site-to-site residuals as a function of vibration period from physics-based ground motion simulations for basin and hill sites in the Canterbury region. The black solid and dashed lines are the mean and mean ± 1 standard deviation respectively

There is a clear underprediction of short period amplifications for the Canterbury hill sites, primarily due to their underlying loess soils and/or dominant topographic effects. Topographic effects are not currently modelled in physics-based ground motion simulations.

2.2 Wellington region

Figure 5 (Tiwari et al., 2023) illustrates the site-to-site residuals from physics-based ground motion simulations in each proposed geomorphic category. The average $\delta S2S_S$ is generally unbiased for all categories, especially at long periods (i.e., T>1s). Basin sites, on average, are overpredicted at short vibration periods (T<0.2s) and underpredicted in the range of T=0.5-3s. These trends are consistent with prior observations of ground motions in Wellington, attributable to basin amplification at moderate periods (Bradley et al. 2018). The fact that they are present despite the use of a 3D ground-motion simulation prediction indicates that the 3D velocity models may not be refined enough to capture the full site amplification, and/or the adopted spatial resolution is too coarse.

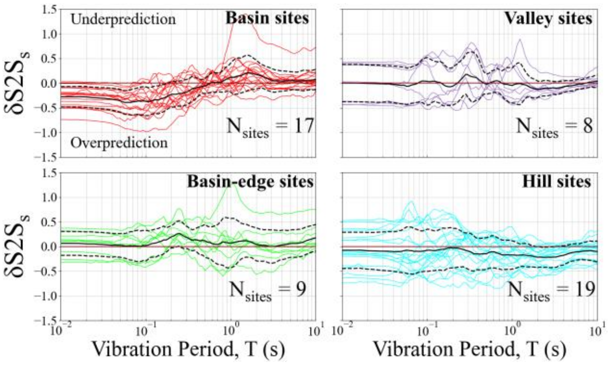


Figure 5: Site-to-site residuals as a function of vibration period from physics-based ground motion simulations for each geomorphic category in the Wellington region. The black solid and dashed lines are the mean and mean ± 1 standard deviation respectively (Tiwari et al., 2023)

Tiwari et al. (2023) shows the observed underprediction at moderate periods is better understood, particularly for valley sites, when the $\delta S2S_S$ residuals are interpreted in terms of normalized vibration periods (i.e., the x-axis is normalized with respect to individual site's fundamental period). This is primarily because valleys are much smaller in length scale and generally not appropriately represented in the 100m grid spacing of the low-frequency portion of the 3D simulations. While geomorphic classification helped understand general trends in systematic site residuals, it was inadequate in explaining various other inaccuracies in ground-motion modelling. Hence, clustering based on the shape of site-to-site residuals was performed to understand the site characteristics that affect this shape

3 HIERARCHICAL CLUSTERING OF SITE-TO-SITE RESIDUALS

Hierarchical clustering approach is performed on the site-to-site residuals across the entire country. Sites characterized by smaller Euclidean distances between their residuals across all vibration periods are grouped into the same cluster. Selected clusters from this method are shown in Figure 6. The objective of applying this clustering approach is to identify various underlying physical mechanisms that contribute to the diverse shapes of site-to-site residuals in these clusters. Hierarchical clustering of site-to-site residuals, along with site geomorphology and various other site characterization parameters, is utilized to identify systematic site effects across the country, for improving forward prediction through physics-based ground motion simulations. Figure 7 shows the geomorphology distribution of the selected clusters from Figure 6.

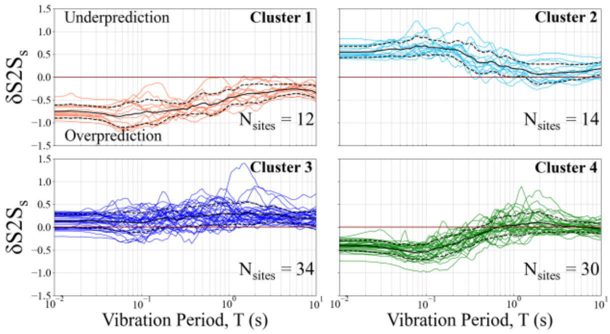


Figure 6: Selected clusters from the hierarchical clustering performed on the site-to-site residuals across all NZ sites. The black solid and dashed lines are the mean and mean ± 1 standard deviation respectively

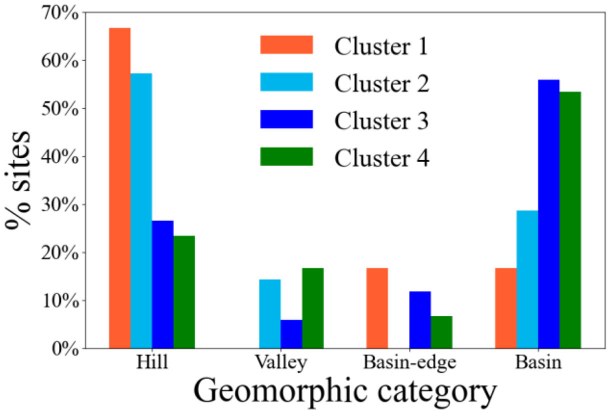


Figure 7: Distribution of geomorphology across the selected clusters

Significant trends with geomorphology from Figure 6 & Figure 7 include:

(a) Cluster 1 shows systematic overprediction at short periods (T < 0.2s) with most sites containing hill sites (~70%).

(b) Cluster 2 shows systematic underprediction at short periods (T < 0.2s) with most sites containing hill sites (~60%).

(c) Cluster 3 shows systematic underprediction between moderate to long periods (0.5 < T < 3s), with most sites containing basin sites (~60%).

(d) Cluster 4 shows systematic underprediction at moderate periods (T~1s), with most sites containing basin sites of Wellington (~60%).

To understand the causes of observed systematic site effects within these clusters, this nationwide investigation examines sites in each geomorphic category. This involves assessing sub-groups within each category, and their diverse site responses that result in distinct site-to-site residuals. The remainder of the paper will focus on examining hill sites within these selected clusters.

3.1 Examination of hill sites

Figure 8 shows the site-to-site residuals explicitly for the hill sites within the selected clusters shown in Figure 6. As shown in Figure 8 (& Figure 7), Clusters 1 (~70%) and 2 (~60%) contains most hill sites, while Clusters 3 and 4 have them in the minority (~20%). The physical mechanisms contributing to the different shapes of these site-to-site residuals (e.g., Cluster 1 and 2) is further explored based on various site characterization parameters.

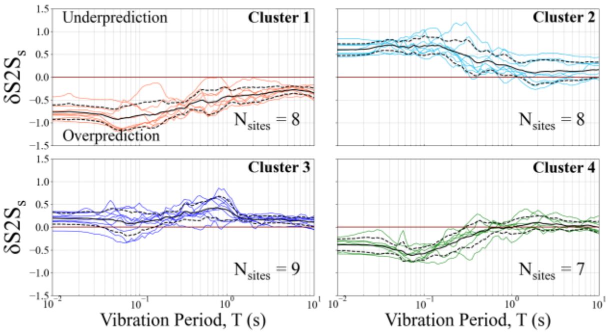


Figure 8: Site-to-site residuals of hill sites within the selected clusters. The black solid and dashed lines are the mean and mean ± 1 standard deviation respectively

Figure 9 shows the distribution of V_{S30} among the hill sites of these selected clusters. The V_{S30} distribution is calculated by considering all the hill sites within each cluster as a percentage of the total sites in that cluster. Clusters 3 and 4 (~20%) have a sparse distribution of hill sites in their total cluster, as opposed to Clusters 1 and 2.

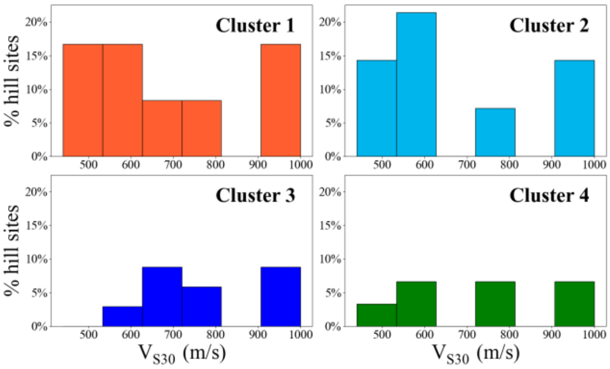


Figure 9: Distribution of V_{S30} among the hill sites as a percentage of the total sites within the selected clusters

As seen from Figure 9, a significant range of V_{S30} variation exists among the hill sites in each of these clusters. However, the V_{S30} distribution between the clusters, especially Clusters 1 and 2, does not show apparent differences, making it challenging to discern the causes of their site-to-site residuals at short periods (T<1s). Moreover, majority of the V_{S30} measurements at these hill sites are not measured directly; rather, they are estimated based on national V_{S30} maps and/or estimated at sites with poor constraints (Wotherspoon et al.,2022). Physics-based ground motion simulations use a V_{S30}-based empirical site amplification model (refer Lee et al.,2022 for detail) to represent shallow site response (T<1s) and deeper site response (T>1s) is inherently captured in the 3D simulation through the NZVM. Furthermore, for the HF component simulations, a generic 1D velocity model is used for the entire country to account for wave propagation and amplification due to the crustal structure and V_{S30}-based site amplification model accounts for shallow soil amplification (Lee et al.,2022).

Figure 10 shows the distribution of fundamental site period, T_0, among the hill sites within the selected clusters. The T_0 distribution is again calculated by considering all the hill sites within each cluster as a percentage of the total sites in that cluster. In contrast to the V_{S30} distribution, T_0 consistently exhibits minimal variation among the hill sites in each of these clusters. Furthermore, unlike the V_{S30} distribution, the majority of T_0 measurements (Wotherspoon et al.,2022) are directly measured at sites with well-constrained conditions. However, T_0 distribution is not sufficient to understand the apparent differences between these clusters (e.g., Clusters 1 and 2) and additional investigation is required.

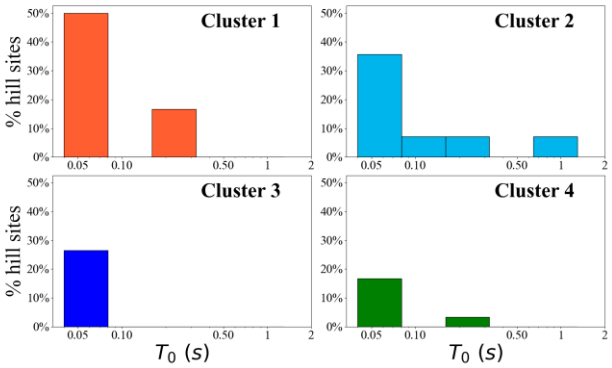


Figure 10: Distribution of T_0 among the hill sites as a percentage of the total sites within the selected clusters

Figure 11 shows the cumulative distribution function (CDF) of relative elevation within a 250m diameter circle, H_{250}, among the hill sites within the selected clusters. As evident in Figure 10, Cluster 1 primarily comprises hill sites with negative H_{250} values (~90%), indicating that most of the hill sites in this cluster have elevations lower than the average elevation within a 250m diameter circle. When examined on a site-by-site basis, these hill sites are typically located at the toe of a hill or on flat rock terrain. In contrast, Cluster 2 has all its hill sites with positive H_{250} values, indicating that every single hill site within this cluster has an elevation higher than the average elevation within a 250m diameter circle. When inspected on a site-by-site basis, these hill sites are typically located at the crest of a hill or near it. There is a significant amount of H_{250} variation among the hill sites within Clusters 3 and 4. Relative elevation parameters (like H_{250} and H_{1250}) are commonly adopted for terrain classification and hence were chosen for testing against residuals.

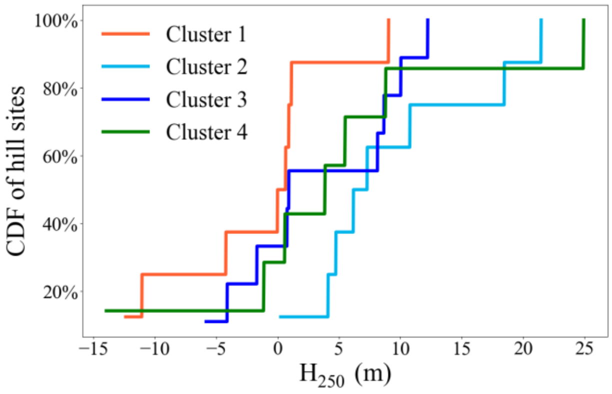


Figure 11: Cumulative distribution function (CDF) of H_{250} among the hill sites within the selected clusters

4 CONCLUSIONS

With the overarching goal of improving physics-based ground motion simulations, this study is aimed to identify and understand systematic site effects and their influencing attributes across New Zealand. Geomorphic categorization, proposed by Nweke et al. (2022), showed apparent differences between the four categories in the Wellington region, with residuals for valley sites illustrating a clear dependence with the inferred fundamental site period. While geomorphic classification aided in comprehending the general trends of systematic site residuals, it proved insufficient in explaining numerous other inaccuracies in ground-motion modelling. Therefore, hierarchical clustering is attempted on the site-to-site residuals across the entire country. These site-to-site residuals are examined against various site characterization parameters, in combination with geomorphic categorization. Hill sites within the selected clusters across the country exhibited diverse shapes of site-to-site residuals, primarily influenced by the relative elevations of the ground motion recording sites. However, the sole predictor variable for estimating shallow site response in physics-based ground motion simulations, V_{S30}, showed significant variability in each selected cluster, making it a poor differentiator for explaining the diverse shapes of site-to-site residuals and assessing sub-groups of hill sites. Relative elevation parameters, like H_{250} and H_{1250}, proved to be better differentiator among the hill sites of the country. An iterative process of utilizing site characterization and clustering is underway to reveal the physical mechanisms currently being overlooked in physics-based ground motion simulations.

REFERENCES

Atik, L. A., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., & Kuehn, N. (2010). The variability of ground-motion prediction models and its components. *Seismological Research Letters*, *81*(5), 794–801. (https://doi.org/10.1785/gssrl.81.5.794)

Bradley, B. A., Wotherspoon, L. M., Kaiser, A. E., Cox, B. R., & Jeong, S. (2018). Influence of site effects on Observed Ground Motions in the Wellington Region from the M_W 7.8 Kaikōura, New Zealand, Earthquake. *Bulletin of the Seismological Society of America*, 108(3B), 1722–1735. (https://doi.org/10.1785/0120170286)

Graves, R., & Pitarka, A. (2015). Refinements to the Graves and Pitarka (2010) broadband ground-motion simulation method. *Seismological Research Letters*, *86*(1), 75–80. (https://doi.org/10.1785/0220140101)

Lee, R. L., Bradley, B. A., Stafford, P. J., Graves, R. W., & Rodriguez-Marek, A. (2022). Hybrid broadband ground-motion simulation validation of small magnitude active shallow crustal earthquakes in New Zealand. *Earthquake Spectra*, *38*(4), 2548–2579. (https://doi.org/10.1177/87552930221109297)

Nweke, C. C., Stewart, J. P., Wang, P., & Brandenberg, S. J. (2022). Site response of sedimentary basins and other geomorphic provinces in Southern California. *Earthquake Spectra*, *38*(4), 2341–2370. (https://doi.org/10.1177/87552930221088609)

Thomson, E. M., Bradley, B. A., & Lee, R. L. (2020). Methodology and computational implementation of a New Zealand Velocity Model (NZVM2.0) for broadband ground motion simulation. *New Zealand Journal of Geology and Geophysics*, 63(1), 110–127. (https://doi.org/10.1080/00288306.2019.1636830)

Tiwari, A., de la Torre, C., Bradley, B., Lee, R., and Kuncar, F. (2023): Trends in systematic site residuals with geomorphic categories for New Zealand ground-motion instrument sites, *Proceedings of the 2023 New Zealand Society for Earthquake Engineering Annual Technical Conference, Auckland, New Zealand*. (https://repo.nzsee.org.nz/handle/nzsee/2561)

Wotherspoon LM, Kaiser AE, Manea EF, Stolte AC. 2022. National Seismic Hazard Model: Site Characterisation Database summary report. Lower Hutt (NZ): *GNS Science*. 54 p. (GNS Science report; 2022/28). (https://doi.org/10.21420/363X-CK83)