

RESEARCH ARTICLE

Non-Stationary Spatial Correlation in New Zealand Strong Ground-Motion Data

Yilin Chen*¹ | Brendon A. Bradley² | Jack W. Baker¹

¹Department of Civil & Environmental Engineering, Stanford University, California, USA

²Department of Civil & Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand

Correspondence

*Yilin Chen Email: yilinc2@stanford.edu

Summary

This paper presents new techniques for quantifying non-stationary spatial variations in strong ground-motion, using data from recent well-recorded earthquakes in New Zealand. The dataset is unique in that many recording stations are relatively densely spaced, and multiple strong ground motions have been recorded at the same stations. This allows calculation of site-specific and region-specific correlations in ground-motion amplitude for Wellington and Christchurch, and the results are compared to a model assuming stationary correlations at all locations. Strong non-stationarity in spatial correlation is observed in the Wellington and Christchurch regions. Heterogeneous geologic conditions appear to be associated with the non-stationary spatial correlation. Several factors influencing non-stationary spatial correlation were studied: (a) site-specific residuals indicate deviation of correlation from a stationary model; (b) most earthquakes have no systematic effect on spatial correlations and there is no indication of a trend in correlations with magnitude; (c) rupture complexity is related to the variation of spatial correlations in ground-motion residuals; (d) variation of site-specific correlations cannot be resolved by using $V_{S,30}$ terms in Ground-Motion Models. The non-stationary correlation approach provides an opportunity to incorporate the site-specific effects in future correlation models.

KEYWORDS:

Spatial Correlation, Non-stationary, Ground-Motion

Paper Citation: Chen, Y., Bradley, B. A., and Baker, J. W. (2021). “Nonstationary spatial correlation in New Zealand strong ground-motion data.” *Earthquake Engineering & Structural Dynamics*, 50(13), 3421-3440. <http://dx.doi.org/10.1002/eqe.3516>

1 | INTRODUCTION

When an earthquake causes ground-motion shaking in a region, the amplitude of shaking (measured, for example, using spectral acceleration at a given period) varies spatially. Some of that variation is predicable via Ground-Motion Models (GMMs), but there is significant remaining variation in ground-motion amplitudes not captured by those models. This remaining variation in ground-motion prediction “residuals” is significant, exhibiting spatial correlations at scales of tens of kilometers ^{e.g., 1,2}, and is important to consider when assessing risk to spatially distributed infrastructure or portfolios of properties ^{3,4}.

This spatial correlation is expected, since functional forms and parameters used in GMMs are relatively simple, so they only partially capture the spatial variation in ground-motion amplitudes. For example, ruptures of finite dimension have spatial variation in the slip amplitude, rupture velocity, and rise time⁵, which is typically not explained by GMMs⁶. Additionally, due to commonalities in crustal velocity structure and wave propagation paths, proximal stations tend to have common wave propagation amplification or de-amplification effects as waves propagate from the fault through the crust⁷. Finally, surficial geology impacts ground-motion amplitudes, while site parameters such as $V_{S,30}$ (average shear wave velocity over the top 30 m) in GMMs only partially capture these effects. This also introduces spatial correlation in ground-motion residuals⁸.

Current correlation models are developed by assuming stationarity (i.e., that the correlations in residuals at stations separated by the same distance are identical). This stationarity assumption has been made out of necessity, as there are few events recorded by a sufficient number and density of ground-motion instruments. A number of predictive models of spatial correlation have been developed based on this assumption e.g.,^{9,10,11,12,13,14}. These models were developed by measuring the correlation between the ground-motion residuals at different stations during a single event and predicting the correlation as an exponentially decaying function with separation distance. Some researchers find that site conditions can affect correlations, based on analyzing stationary correlation results from different soil conditions or regions e.g.,^{15,16}. Recent research have investigated the influence of source and path effects on non-stationary correlations^{17,18}. However, there has been no systematic study of site-specific correlation of ground-motion residuals.

This paper develops and applies new approaches for quantifying non-stationary spatial correlations in strong ground-motion observations, using densely recorded ground-motion data in New Zealand, where multiple strong ground motions have been recorded at the same stations. The strong ground-motion dataset used in this study is described in Section 2. The methodology for non-stationary spatial correlation calculation is presented in Section 3, and results are discussed in Section 4. Lastly, a sensitivity analysis for factors influencing non-stationary spatial correlations is presented in Section 5.

2 | DATA

We used strong ground-motion data from Van Houtte et al.¹⁹. The database contains ground-motion recordings from earthquakes that occurred between 1968 and 2016 in New Zealand. The recordings were individually processed to remove low-quality signals and any contaminating noise. Spectral accelerations for different damping ratios and periods were then computed from the processed recordings.

For this analysis, the dataset was filtered using several criteria. Only stations in the Christchurch and Wellington region are considered, as they contain the densest instrumentation in the country, and have repeated observations of strong shaking. For those stations, only recordings less than 100 km from the fault rupture are considered. Recordings were only considered if they came from crustal earthquakes with magnitude (M) of at least 3, intraslab subduction earthquakes with $M \geq 5$, or interface subduction earthquakes with $M \geq 6$. These data criteria follow the suggested parameter ranges from the GMMs used in this study, which are presented in Section 2.1. Figure 1 shows the locations of earthquake epicenters in the two regions, and Figure 2 shows the station locations. Additional earthquake information is listed in Table A1 and A2.

2.1 | Ground-motion models

Spatial correlations of ground motions are characterized based on residuals from GMMs. A typical GMM formulates the observed intensity measure as the sum of the prediction and residuals. In the case of spectral acceleration (SA), it is expressed as:

$$\ln SA_{i,j} = \mu_{\ln SA}(rup_i, site_j) + \delta B_i + \delta W_{i,j} \quad (1)$$

where $SA_{i,j}$ is the spectral acceleration at the period of interest at site j caused by rupture i ; $\mu_{\ln SA}(rup_i, site_j)$ is the predicted natural logarithmic mean of SA intensity, and $rup_i, site_j$ are the rupture and site parameters used in GMMs (e.g., magnitude, source-to-site distance, site condition); δB_i is the between-event residual for the i^{th} rupture; $\delta W_{i,j}$ is the within-event residual for site j from the rupture i . δB_i and $\delta W_{i,j}$ are normal random variables with zero means and standard deviations of τ and ϕ , respectively. Although the observed $SA_{i,j}$ is decomposed in this equation into three components, the correlation of $\delta W_{i,j}$ is the primary term of interest for spatial correlations, as the $\mu_{\ln SA}(rup_i, site_j)$ is already characterized by the prediction model and δB_i is constant for different sites.

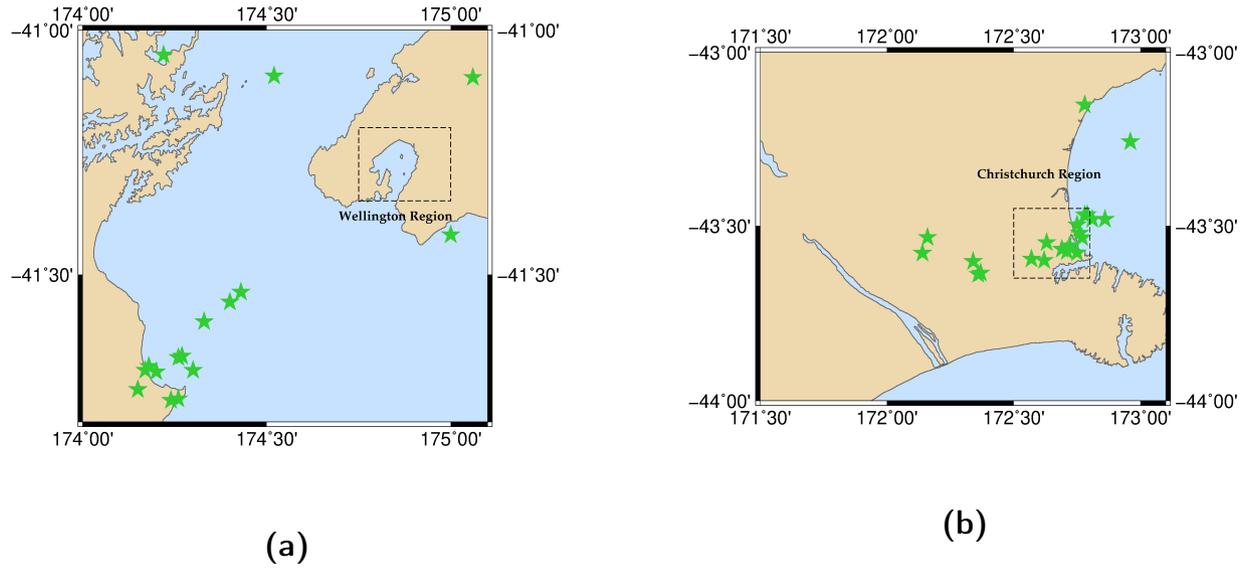


FIGURE 1 Locations of earthquake epicenters in the (a) Wellington region, and (b) Christchurch region. Dash lines show the study regions in Figure 2.

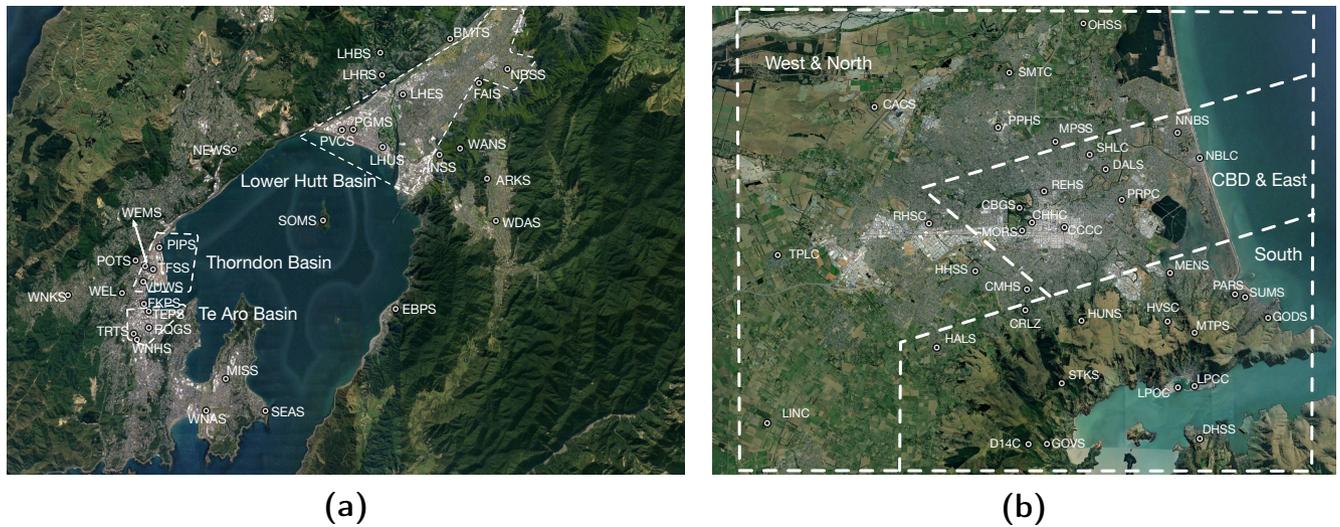


FIGURE 2 Locations of stations in the (a) Wellington region, and (b) Christchurch region.

In this study, the Abrahamson et al. GMM²⁰ is used for subduction zone events, and the Chiou and Youngs GMM²¹ for crustal events to calculate the residuals. Note that the residual calculations are dependent upon a reference GMM. However, the results should be similar across different GMMs because the variations in observations are large compared with the variations in predictions, so changing the reference GMM does not often impact residuals appreciably^{22,23}.

2.2 | Within-event residual calculation

For each earthquake and station considered in this study, we have an observed $\ln SA_{i,j}$, and predicted $\mu_{\ln SA}(rup_i, site_j)$, τ and ϕ terms from the reference GMM. Subtracting the predicted $\mu_{\ln SA}(rup_i, site_j)$ from the observed $\ln SA_{i,j}$ gives a total residual, $\epsilon_{i,j}^t$:

$$\varepsilon_{i,j}^t = \ln SA_{i,j} - \mu_{\ln SA}(rup_i, site_j) \quad (2)$$

Comparing Equations 1 and 2, we see that the total residual equals the sum of within-event and between-event residuals. Separating the between-event residuals from the total residuals requires solving a mixed-effect regression problem. In this study, the following method is used to estimate δB_i for each earthquake^{24,25,26}:

$$\delta \hat{B}_i = \frac{\sum_{j=1}^n \varepsilon_{i,j}^t}{\frac{\phi^2}{\tau^2} + n_i} \quad (3)$$

where $\delta \hat{B}_i$ denotes the estimated between-event residual, and n_i is the number of recordings of earthquake i . Then normalized within-event residuals $\delta \tilde{W}_{i,j}$ can be estimated by:

$$\delta \tilde{W}_{i,j} = \frac{\delta \hat{W}_{i,j}}{\phi} = \frac{\ln SA_{i,j} - \mu_{\ln SA}(rup_i, site_j) - \delta \hat{B}_i}{\phi} \quad (4)$$

We further estimate a site-specific residual at a given station as the mean over all observations of the normalized within-event residuals:

$$\delta S2S_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \delta \tilde{W}_{i,j} \quad (5)$$

where n_j is the number of earthquakes recorded at station j .

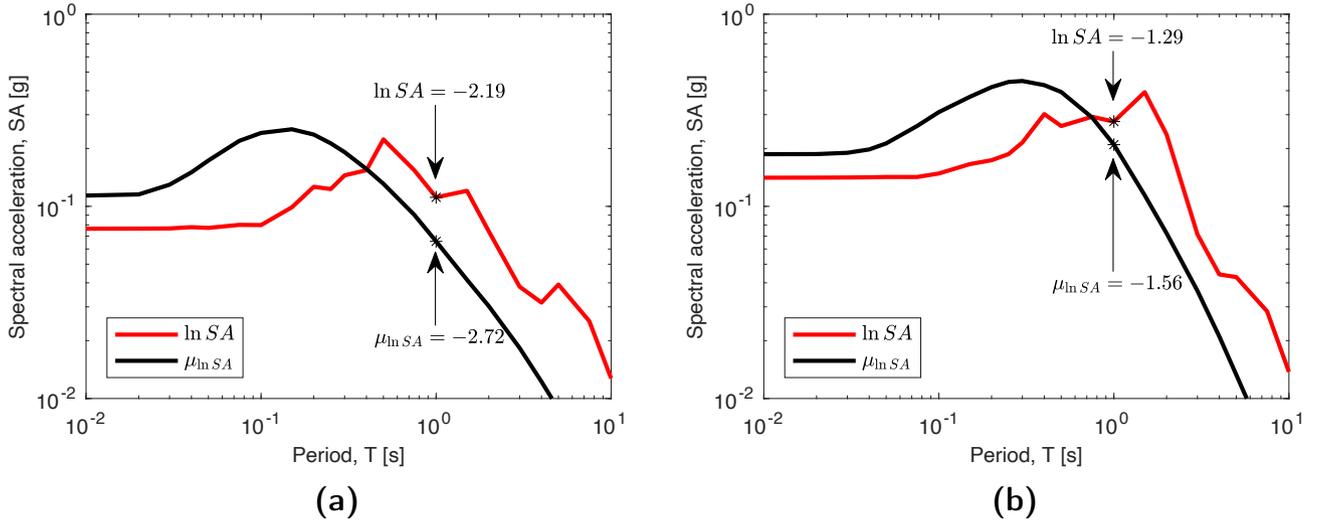


FIGURE 3 Example observed ground-motion response spectra and GMMs as used in this study: (a) Station POTS, and (b) Station WEMS recordings of the 13 November 2016 M 7.8 Kaikōura earthquake.

Figure 3 illustrates the data used to compute normalized $SA(1\text{ s})$ within-event residuals at stations POTS and WEMS. The normalized within-event residual for station POTS is calculated as:

$$\delta \tilde{W} = \frac{\delta \hat{W}}{\phi} = \frac{\ln SA - \mu_{\ln SA}(rup_i, site_j) - \delta \hat{B}}{\phi} = \frac{-2.19 + 2.72 - 0.19}{0.60} = 0.57 \quad (6)$$

where the estimated between-event residual for the Kaikōura earthquake from Equation 3 is $\hat{\delta}B = 0.19$ and the within-event standard deviation given by the GMM is $\phi = 0.60$. Similarly, the normalized within-event residual for station WEMS is calculated as $(-1.29 + 1.56 - 0.19)/0.56 = 0.14$. We use this method to calculate the normalized within-event residuals for all recordings. These residuals are used to calculate the correlations in the following analysis.

3 | METHODOLOGY

3.1 | Site-specific correlation

Traditionally, spatial correlations of ground motions are studied by calculating a semivariogram of residuals from a single event, assuming stationarity *e.g.*, 1,9,10,15. Here we assess this assumption explicitly by also determining the non-stationary correlations using a site-specific approach¹⁷. For every pair of stations (j, k) we select all earthquakes with suitable recordings at both stations. We assume the pair of within-event residuals could be represented by a bivariate normal distribution²⁷, and calculate the sample correlation coefficient for the normalized within-event residuals:

$$\hat{\rho}_{j,k} = \frac{\frac{1}{n} \sum_{i=1}^n \delta\tilde{W}_{i,j} \delta\tilde{W}_{i,k}}{\sqrt{\frac{1}{n} \sum_{i=1}^n \delta\tilde{W}_{i,j}^2} \sqrt{\frac{1}{n} \sum_{i=1}^n \delta\tilde{W}_{i,k}^2}} \quad (7)$$

where n is the number of earthquakes with pairs of recordings at the given stations. This gives a correlation estimate for every pair of stations. Notice that Equation 7 is an uncentered sample correlation coefficient, which corresponds to the maximum likelihood estimator for bivariate normal data conditioned on the means being zero²⁸. The uncentered approach is designed to capture any site-specific effect associated with station j and station k (because a centered approach will remove the mean of within-event residuals which could influence the resulting correlation.)

The standard deviation of the Equation 7 correlation coefficient estimate is²⁹:

$$std(\hat{\rho}) = \frac{1 - \rho^2}{\sqrt{n}} \quad (8)$$

where ρ is the real correlation coefficient. We examined station pairs with varying n , and determined that $n \geq 6$ pairs of recordings were needed to obtain stable estimates of site-specific correlations while maintaining a sufficient number of station pairs to draw scientific insights.

Figure 4 illustrates the site-specific correlation calculation for within-event residuals at stations WEMS and POTS. Each data point represents the residuals at the two stations from a single earthquake. For example, the point at (0.57, 0.14) represents the residuals of the Kaikōura earthquake calculated in Section 2.2. The estimated correlation is 0.90 according to Equation 7, and the standard deviation of the estimated correlation is 0.06 according to Equation 8.

3.2 | Deviation of site-specific correlation

After calculating the site-specific correlations of all pairs of stations in the regions, we evaluate the deviation of these correlations relative to a stationary reference model. Fisher's Z-transformation is first applied³⁰:

$$Z(\hat{\rho}) = \frac{1}{2} \ln \left(\frac{1 + \hat{\rho}}{1 - \hat{\rho}} \right) \quad (9)$$

where $\hat{\rho}$ is the sample correlation coefficient from Equation 7. If the within-event residuals are bivariate normal, $Z(\hat{\rho})$ is approximately normally distributed with mean $1/2 \ln \left(\frac{1+\rho}{1-\rho} \right)$ and standard deviation $1/\sqrt{n-2}$. Then we can define

$$e = [Z(\hat{\rho}) - Z(\rho)] \sqrt{n-2} \quad (10)$$

as the measure of correlation deviation, where the real correlation coefficient ρ is predicted from a reference correlation model. Note that the $\sqrt{n-2}$ is slightly different then the original standard deviation of the Fisher's Z-transformation, because here the correlation is estimated using an uncentered approach. Under the above assumptions, e will follow the standard normal distribution.

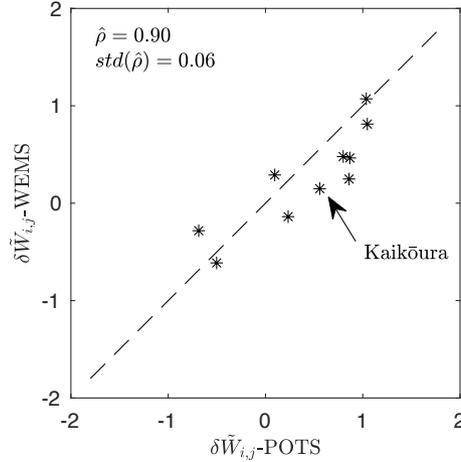


FIGURE 4 Scatter plot of $SA(1\text{ s})$ within-event residuals at stations POTS and WEMS ($n = 10$, separation distance = 0.46 km). A line with a slope of 1 is included for reference.

In the above example, the calculated $\hat{\rho}$ is 0.90 for stations POTS and WEMS, and the predicted correlation ρ is 0.95 from the reference model¹⁵. Therefore, the deviation of the site-specific correlation from the reference model is

$$e = [Z(0.90) - Z(0.95)] \times \sqrt{10 - 2} = -1.02 \quad (11)$$

which indicates that the estimated correlation between POTS and WEMS is lower than the stationary prediction from the reference model by 1.02 standard deviations.

3.3 | Influence of individual earthquakes on site-specific correlation estimates

Similarly, we can quantify the influence of a specific earthquake on the computed correlation by comparing the site-specific correlations with and without the specific earthquake data. Specifically, the influence of earthquake i on the correlation at a pair of stations is defined as

$$e_i = [Z(\hat{\rho}) - Z(\hat{\rho}_{-i})] (n - 2) \quad (12)$$

where $Z(\hat{\rho}_{-i})$ is the same transformed correlation, but with the correlation estimated after omitting data from earthquake i . The standard deviation of $Z(\hat{\rho}) - Z(\hat{\rho}_{-i})$ is estimated to be $n - 2$ using bootstrap method³¹, so division by this standard deviation makes the e_i terms have a standard normal distribution. A positive e_i score means that earthquake i increases the correlation at a given pair of stations (because its omission decreases the correlation), and vice versa.

Figure 5 shows an example of calculated e_i of earthquakes for stations TEPS and MISS. In this case, omitting the Kaikōura earthquake will change the site-specific correlation from 0.30 to 0.62, so the calculated e_i for the Kaikōura earthquake is:

$$e_i = [Z(0.30) - Z(0.62)] \times (8 - 2) = -2.49 \quad (13)$$

which indicates including the Kaikōura earthquake results in reducing the estimated correlation for this pair.

4 | NON-STATIONARY CORRELATION ANALYSIS

The Section 3.1 and 3.2 methods are next used to calculate the site-specific correlations and deviations, using 5%-damped SA at a period of 1 s. This SA period was chosen as it is potentially sensitive to the effects of deep geologic structures such as sedimentary basins, as shorter-period waves lose correlation more quickly and are more influenced by fine-scale crustal structure.

Figure 6 illustrates the site-specific correlations of within-event residuals for every pair of stations. The solid line shows the reference model from Jayaram and Baker¹⁵, and the dashed lines denote \pm one standard deviation intervals calculated from Equation 8 with $n = 6$. The estimated correlations are, on average, consistent with the reference model, but individual correlation

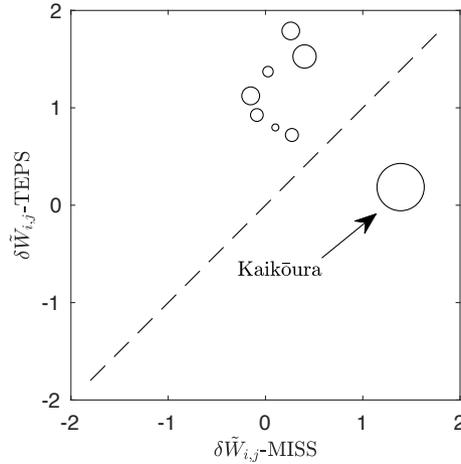


FIGURE 5 Scatter plot of within-event residuals of $SA(1\text{ s})$ at stations TEPS and MISS. The diameter of the circle corresponding to each data point is proportional to the e_i of that earthquake, as calculated by Equation 12.

estimates vary widely around that model. The color of each point shows the deviation of site-specific correlations from the reference model calculated using Equation 10. The sample mean (μ_e) and standard deviation (σ_e) of the deviations are annotated on the lower left of the figure.

In Figure 6, 400 of the 485 total station pairs are outside of the one standard deviation interval. If $\rho(h)$ was an appropriate model for all station pairs, we would expect only 1/3 (or ~ 160) pairs to be outside of this interval. We observe low-correlated pairs at short distances, and both high-correlated and negative-correlated pairs at large distances. There is significant variation relative to the predicted correlation from the reference model. This suggests that site-specific correlations are not well explained by their separation distances, h , alone. Similar results are also observed for other periods of interest. In the following sections, we select stations from different subregions defined in Figure 2b and 2a, to study the correlations for subregions with different characteristics.

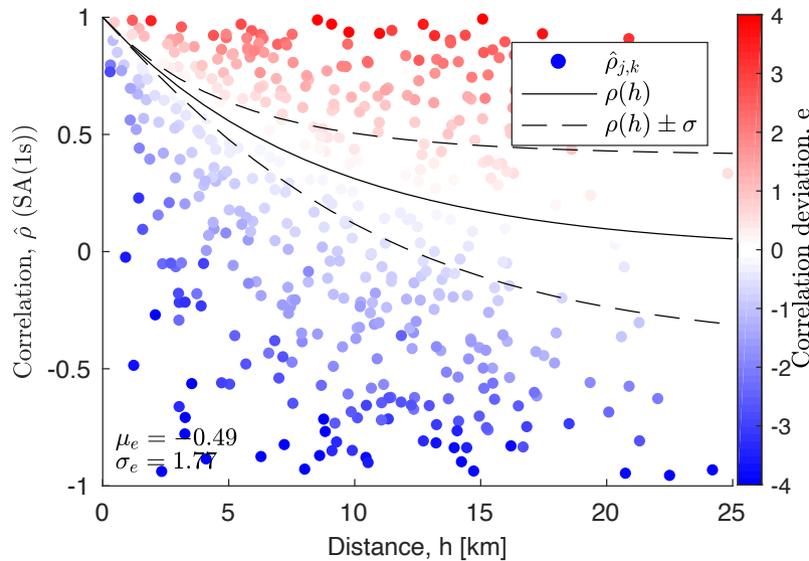


FIGURE 6 Site-specific correlations and deviations for $SA(1\text{ s})$ within-event residuals versus separation distance. Red points show pairs with a positive correlation deviations and blue points show pairs with a negative correlation deviation. μ_e and σ_e are the sample mean and standard deviation of the correlation deviations e .

4.1 | Christchurch region

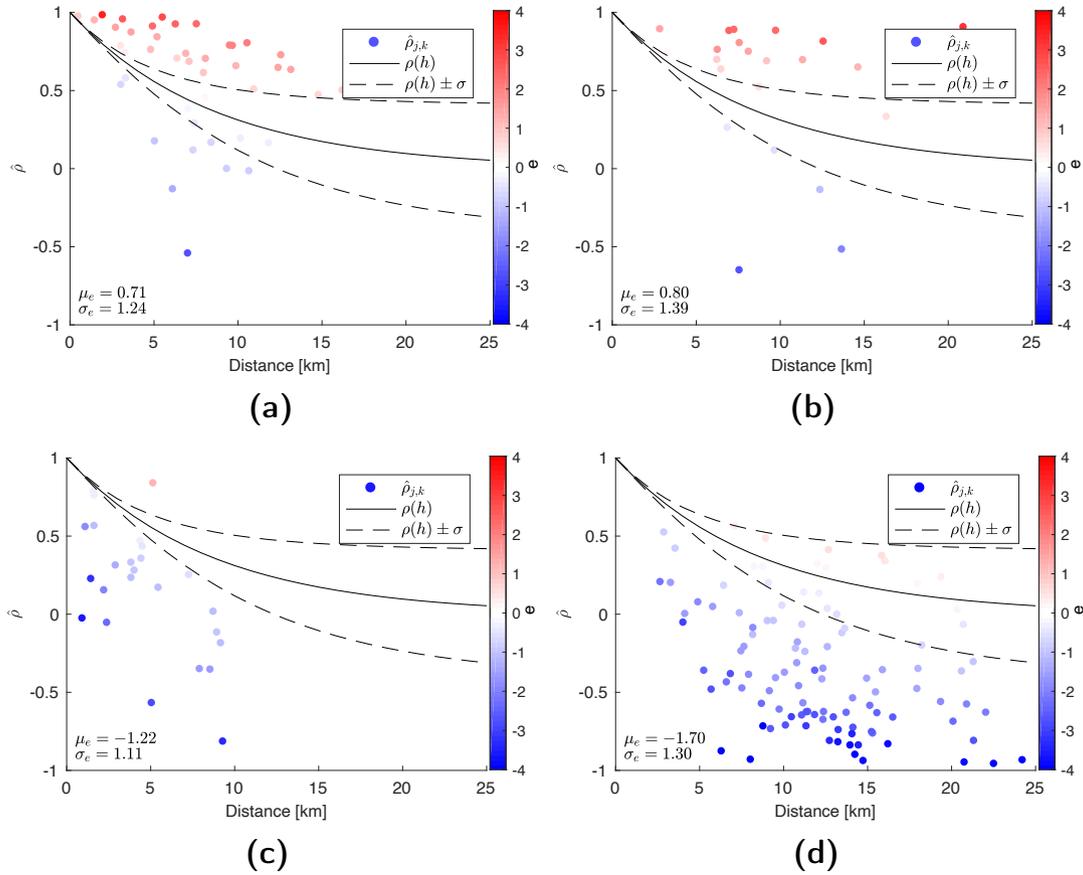


FIGURE 7 Site-specific correlations and deviations for $SA(1s)$ within-event residuals versus separation distance for Christchurch station pairs (a) in the south subregion, (b) in the north & west subregion, (c) in the CBD & east subregion, (d) with one in the south subregion and another outside. (See subregions in Figure 2b).

Figure 7 shows Christchurch station-pair correlations, grouped by their subregions. Figure 7a, 7b, and 7c show the correlations of station pairs in the south, west & north, and CBD & east subregions, respectively, and Figure 7d shows the correlation of the pairs with one station in the south subregion and another outside.

In Figure 7a, most pairs in the south subregion have a higher correlation than the reference model, and the difference is significant for most pairs with distance < 15 km. This is because the stations in this subregion are mainly on a rock condition, and are closer to the epicenters of the earthquakes in the region (Figure 1b). Therefore, the stations tend to have similar shaking intensity and similar GMM predictions^{32,33,34}, which leads to similar residuals and thus high correlation coefficients (i.e., positive μ_e). Similarly, for the stations in the west & north subregion, the stations are mostly located on dense gravel geological conditions and further from the epicenters of the earthquakes, which also results in similar residuals and high correlation coefficients (Figure 7b).

Conversely, for the stations in the CBD & east subregion in Figure 7c, the correlation of the station pairs shows faster decay than the reference model and thus a negative μ_e . The stations in this subregion are on a relatively soft soil condition, but soils vary from deep sedimentary deposits of interbedded gravels to fine-grain sediments³⁵. These stations tend to have appreciably different residuals for different earthquakes, which leads to a higher variation of residuals and thus a lower correlation.

Figure 7d shows the correlations for pairs of stations with one in the south subregion and another outside of the south subregion. These pairs of stations are on different geological conditions. It is common that the GMMs overpredict shaking at rock

stations and underpredict at soil stations, or vice versa, in this dataset. Therefore, these pairs of stations usually have low or even negative correlations.

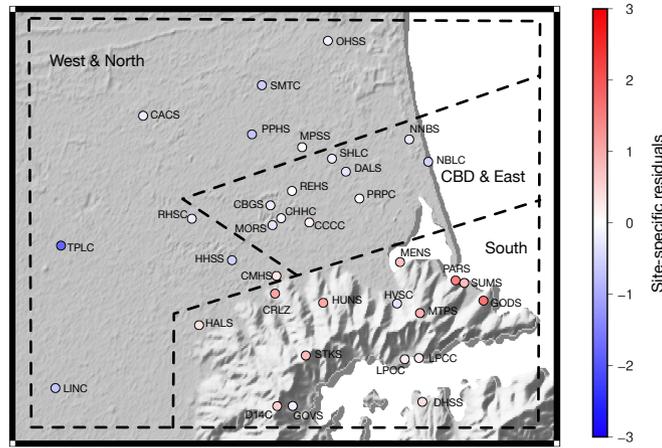


FIGURE 8 Site-specific $SA(1\text{ s})$ residuals in the Christchurch region.

Figure 8 shows the site-specific residuals from all earthquakes in Christchurch. For the south subregion, it can be seen that the GMMs underpredicted $SA(1\text{ s})$, which leads to an overall positive residual in the subregion. These systematic underpredictions caused higher correlations within the region. Similarly, the west & north subregion tends to have systematic negative residuals, leading to higher correlations. This opposite sign of residuals causes the low correlations for stations in the south subregion with others in Figure 7d.

4.2 | Wellington region

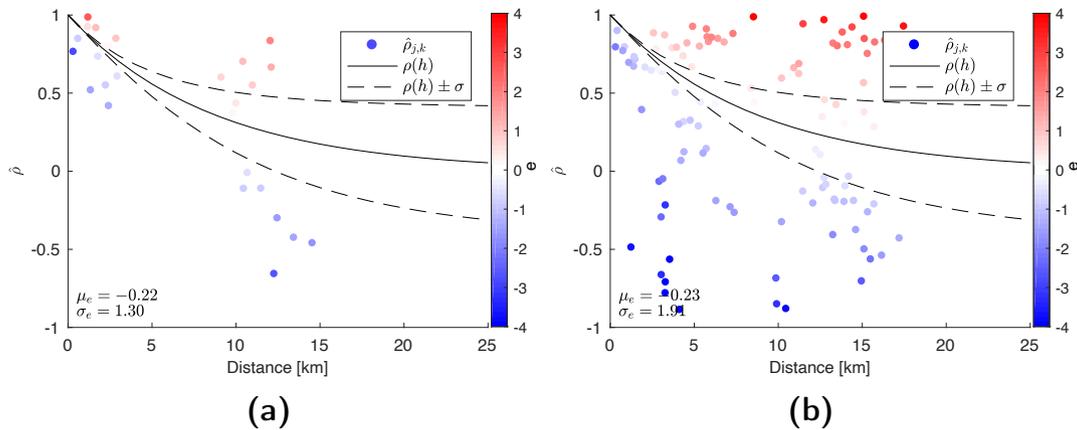


FIGURE 9 Site-specific correlations and deviations for $SA(1\text{ s})$ within-event residuals versus separation distance for Wellington station pairs (a) in the basin subregions (Thorndon, Te Aro, and Lower Hutt Basin), (b) with one in the basin subregions and another outside. (See subregions in Figure 2a).

Similarly, we study the correlations of the subregions in Wellington. Figure 9a shows the correlations for station pairs in the basin subregions. These stations have a very similar geological condition (soft soil), and are spatially dense compared with

the distance to earthquake epicenters. Their correlation shows good alignment to the reference model. However, in Figure 9b, the correlation for the basin stations with outside basin stations vary significantly. These stations are located at heterogeneous geological conditions, and are relatively scattered, which leads to high variation of $SA(1\text{ s})$ for different earthquakes and thus high variation of correlations. Figure 10 shows the site-specific residuals from all earthquakes in the Wellington region. The stations in the basin subregions show similar residuals, and thus have a higher correlation. However, for other stations, the site-specific residuals vary spatially, which leads to a high variation of their correlations. In addition, there are a few pairs with a station at the edge of a sedimentary basin and a station on a rock outcrop. The influence of basin and basin-edge-generated surface waves leads to significant variation over small distances³⁶, which causes the abnormally lower-correlated pairs at close distances observed in Figure 9b. For example, station BMTS, FAIS, and INSS are on a rock outcrop, but they have appreciably different residuals than the stations in the basin. The results for other periods are listed in Appendix B.

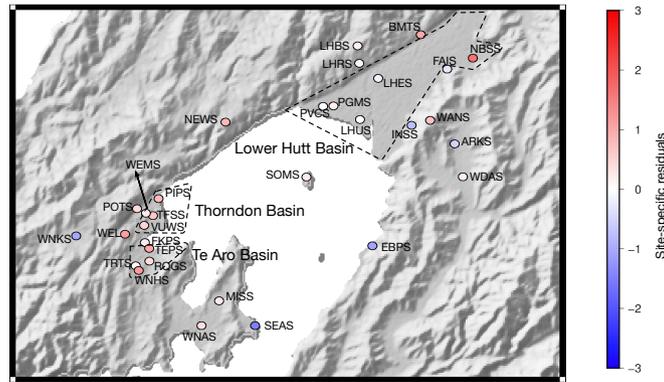


FIGURE 10 Site-specific $SA(1\text{ s})$ residuals in the Wellington region.

5 | SENSITIVITY ANALYSIS

5.1 | Effect of site-specific residuals

In Section 4, we saw that the site-specific residuals influenced the correlation of a pair of stations. In Figure 11, we plot the correlation deviation versus the product of site-specific residuals at a given pair of stations. As expected, a positive product (resulting when the site-specific residuals have the same sign) is usually associated with a positive correlation deviation, and vice versa. The relationship is close to a logistic function, as indicated by the red line showing a logistic function fitted to the data.

Since e follows a standard normal distribution, we can examine what values of $\delta S_2 S_i * \delta S_2 S_j$ cause an extreme e (i.e., strong non-stationarity). According to the fitted logistic function, $\delta S_2 S_i * \delta S_2 S_j > 0.47$ will result in $|e| > 1.96$. These values are indicated with dashed lines in Figure 11.

We next plot the correlation deviations versus the site-specific residuals at a pair of stations. In Figure 12, the coordinates of each data point indicate the site-specific residuals at a pair of stations, and the color indicates its correlation deviation. The dashed lines show the contours of $|\delta S_2 S_i * \delta S_2 S_j| = 0.47$. It is clear that if a pair of stations has a product of site-specific residuals close to zero (region ①), it tends to have a e also close to zero (i.e., stationary correlation). On the other hand, if a pair of stations has a large site-specific residual product (region ②), they tend to have a positive deviation and thus a higher correlation than the reference model. Conversely, an opposite sign of site-specific residuals is associated with the station pairs with negative deviations (region ③). These non-zero site-specific residuals can be caused by the systemic under/overprediction of GMMs, which are observed to be associated with similar geological conditions or similar locations relative to earthquake epicenters in Section 4.

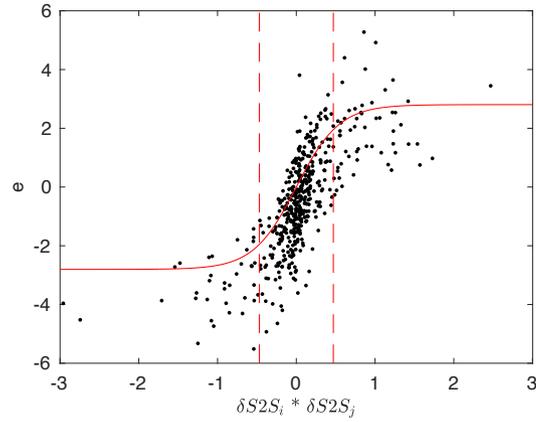


FIGURE 11 Correlation deviation versus the product of site-specific residual pairs. The solid red line shows the fitted logistic curve of the data. The dashed lines show the range of $\delta S2S_i * \delta S2S_j$ that has a e between -1.96 and 1.96 according to the fitted logistic function.

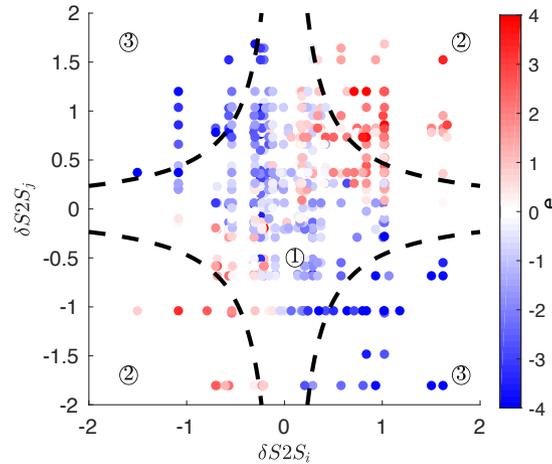


FIGURE 12 Scatter plot of site-specific residuals for all station pairs, colored by the pair's correlation deviation. The dashed lines show the contours of $|\delta S2S_i * \delta S2S_j| = 0.47$.

5.2 | Effect of individual events

We use Equation 12 to calculate the influence of earthquakes (e_i) on all pairs of stations in Christchurch and Wellington. These e_i values are plotted versus earthquake magnitude. In Figure 13, the mean e_i for all earthquakes are around zero, which indicates that no earthquake produces systematically higher or lower correlations, and that the correlations caused by these earthquakes are consistent with each other. However, the Kaikōura earthquake produces some extreme e_i values and has the largest standard deviation ($\hat{\sigma} = 1.54$). This suggests that the Kaikōura earthquake has large-amplitude within-event residuals, and thus adding or removing the earthquake's data will substantially influence a number of site-specific correlations. This is because the Kaikōura earthquake involved complex rupture of many faults with highly variable slip (southern faults had smaller slip) and significant directivity effects (propagation from south to north)^{37,36}. However, GMM predictions use simple parameters such as closest distance to the fault (R_{rup}) that do not capture these effects, which causes regional errors in predictions and thus results in high variations of prediction residuals. On the other hand, the standard deviation of e_i was 0.91 for the Darfield earthquake. It had a somewhat smaller rupture, and ground-motion amplitudes are more easily predicted as a function of R_{rup} , leading to fewer regional deviations in ground-motion amplitudes and thus no systematic correlation effects. This pattern suggests a relationship between rupture complexity and spatial correlations in ground-motion residuals.

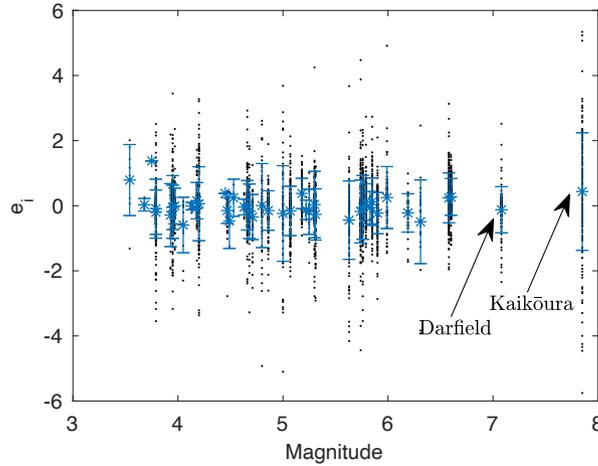


FIGURE 13 Computed e_i from Equation 12 for all earthquakes and station pairs at a period of 1 s, plotted versus earthquake magnitude. Points indicate e_i values for a particular earthquake and station pair. Stars indicate the mean of the e_i scores and bars show the \pm standard deviations, for each earthquake.

Figures 14 and 15 show the within-event residuals for the Kaikōura and Darfield earthquakes. There is noticeable regional clustering of within-event residuals from the Kaikōura earthquake. Due to the directivity effect, positive within-event residuals cluster in the Wellington region (average $\delta\tilde{W}_{i,j} = 0.82$) and negative within-event residuals cluster in the Christchurch region (average $\delta\tilde{W}_{i,j} = -1.72$). Therefore, recordings from this event tend to imply high spatial correlations and thus can result in a high e_i standard deviation, as observed above. The Darfield earthquake (Figure 15), on the other hand, has no apparent regional clustering of residuals. Therefore, removing this event from the correlation coefficient calculation has less influence on the estimated site-specific correlations, resulting in e_i values that are close to zero on average and have a lower standard deviation.

5.3 | Effect of GMM $V_{S,30}$ term

To investigate the role of $V_{S,30}$ in spatial correlations, we use a constant $V_{S,30}$ in GMM predictions (rather than the measured values) to estimate residuals. We then examine whether the spatial correlations between soil and rock stations show more significant non-stationary behavior since GMMs already predict part of the site effects with the $V_{S,30}$ term. We use the mean $V_{S,30}$ of the station dataset (471 m/s) for all stations to calculate the normalized within-event residuals at a period of 1 s, and then calculate $\hat{\rho}$ and e using these residuals.

Figure 16a shows the correlations for rock-soil station pairs using a constant $V_{S,30}$. Most of the station pairs have negative correlation. This is expected, because a constant $V_{S,30}$ introduces opposite biases for rock and soil SA prediction, which results in a drop of correlations. Figure 16b shows the correlations for rock-soil station pairs using best-estimate $V_{S,30}$. By incorporating the $V_{S,30}$ term, we observe that correlations are no longer consistently lower than the reference model, but the variation of correlations are still significant around the reference model. This indicates that including the $V_{S,30}$ term in the GMMs helps modify the systematically lower correlations for rock-soil station pairs. However, the variations remain significant and cannot be eliminated using a $V_{S,30}$ term in GMMs. Further refinement of site-specific predictions using site response modeling should further reduce correlations and nonstationarities³⁸, but it is clear that $V_{S,30}$ alone is not sufficient to capture these effects.

5.4 | Effect of SA period

The above results are specific to the analysis of $SA(1\text{ s})$ values. To evaluate the effect of the considered SA period, the same calculations are repeated for spectral acceleration data at other periods. Appendix B provides the results for $SA(0.01\text{ s})$ and $SA(0.25\text{ s})$. The results at other periods are similar to those of $SA(1\text{ s})$, showing consistency in which subregion correlations deviate from the reference model. The effects are slightly weaker than the $SA(1\text{ s})$ results, which was expected as the shorter-period amplitudes will be less sensitive to deep geologic structures that may be causing some of the correlation. It was not possible to analyze spectral accelerations at periods longer than 1 s, because there were not a sufficient number of ground motion

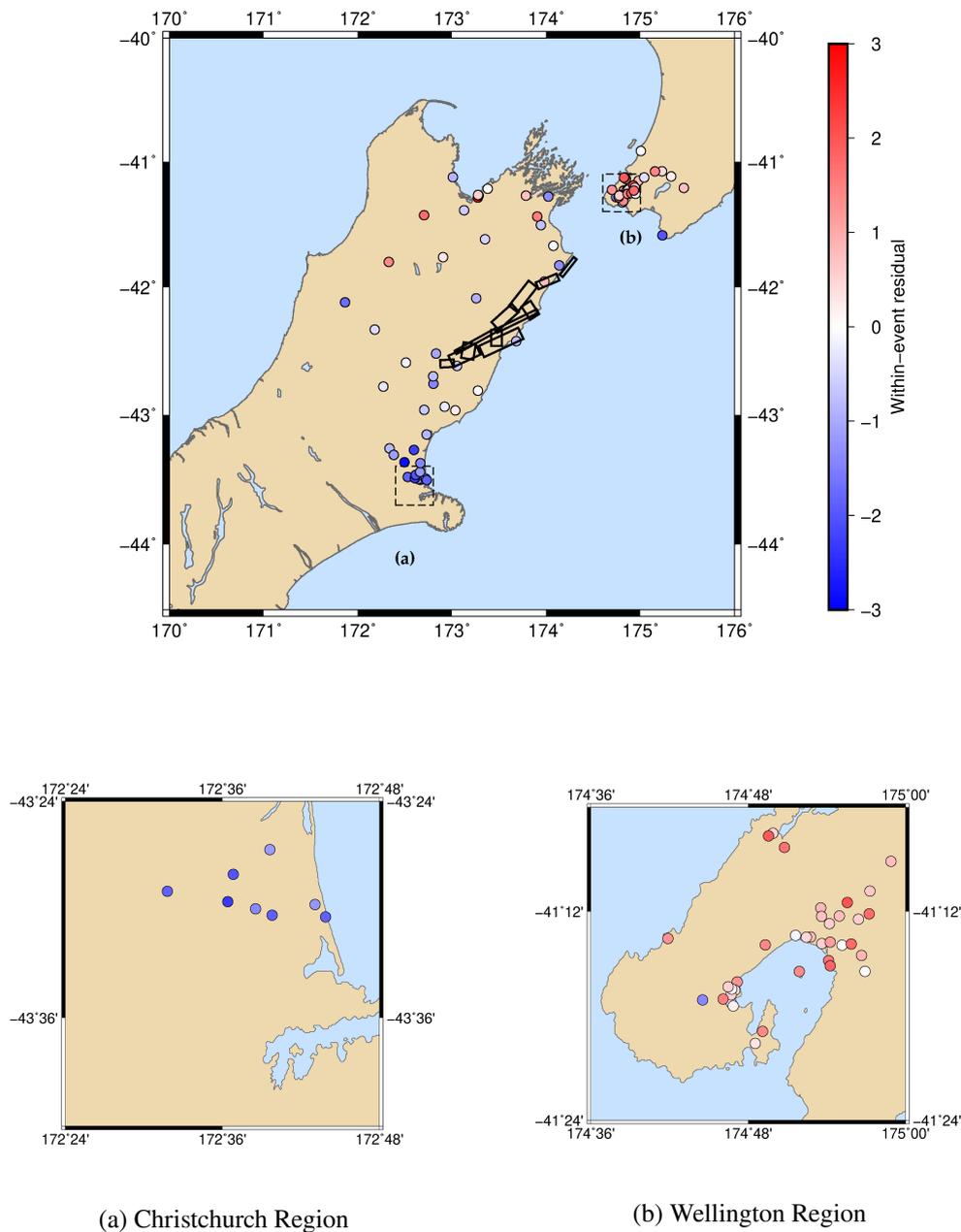


FIGURE 14 $SA(1\text{ s})$ within-event residuals from the 13 November 2016 M 7.8 Kaikōura earthquake. Black lines show the surface projection of the rupture.

recordings with usable signal at those longer periods. In total, these results suggest that subregions may consistently produce non-stationary correlations for a range of spectral acceleration periods, but that longer-period spectral accelerations may have stronger effects. This limited set of results is not definitive, but is consistent with effects seen in numerically simulated ground motions¹⁷.

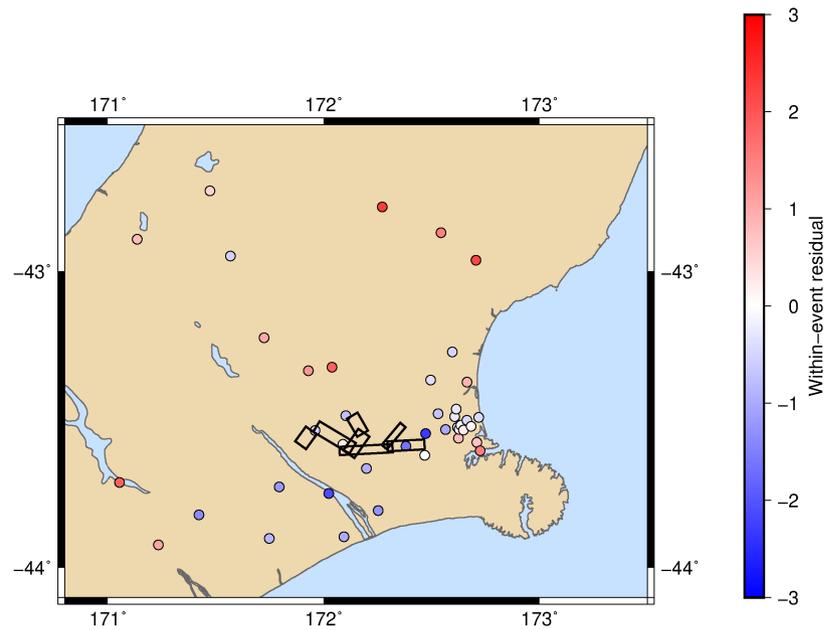


FIGURE 15 $SA(1\text{ s})$ within-event residuals from the 4 September 2010 $M\ 7.1$ Darfield earthquake. Black lines show the surface projection of the rupture.

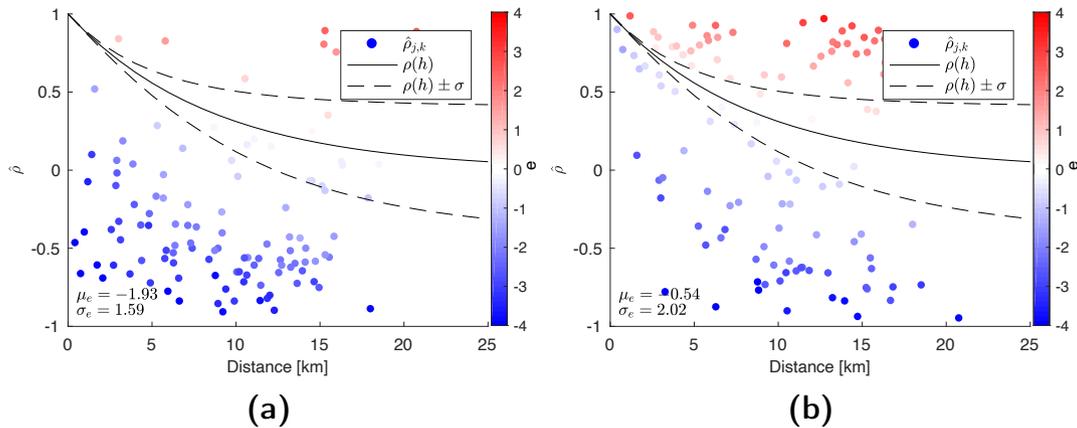


FIGURE 16 Site-specific correlations and deviations for $SA(1\text{ s})$ within-event residuals versus separation distance for rock-soil station pairs (a) using a constant $V_{S,30}$, (b) using best-estimate $V_{S,30}$.

6 | CONCLUSIONS

This paper introduced a methodology to measure and evaluate non-stationary correlations in ground motions. We calculated site-specific correlation coefficients, and used Fisher's Z-transformation to quantify the deviation of these correlations from a reference model. We applied this methodology to New Zealand ground-motion data, to understand how correlations in ground motions vary among station pairs.

On average, the New Zealand ground motions have spatial correlations comparable to a reference model that was previously calibrated based on global data sets. However, there is appreciable variability in the pair-wise correlations between stations for a given separation distance. Strong non-stationarity in correlations is observed in the Christchurch and Wellington regions, with some subregions of those areas exhibiting higher correlations than the reference model, and station pairs from differing subregions having lower correlations.

The site-specific residuals at considered stations are strongly related to the correlation deviation. Station pairs with the same sign of site-specific residuals tend to have higher correlation than the reference model and vice versa. In general, stations with similar geological conditions and similar relative location to earthquake epicenters tend to have similar site-specific residuals and higher correlations. Station pairs on homogeneous conditions tend to have good alignment with a stationary model, but heterogeneous conditions tend to be associated with higher variation of spatial correlations.

Most earthquakes had no systematic effect on correlations, and there was no apparent trend in this effect with the magnitude or other rupture properties. Large variation of residual values, on the other hand, may cause a large variation in spatial correlations. For example, the Kaikōura earthquake has the largest standard deviation of e_i at a period of 1 s due to the clustering of residuals caused by the complexity of the rupture.

Variation of site-specific correlations cannot be resolved by using $V_{S,30}$ terms in GMMs. We show this by computing residuals and spatial correlations with best-estimate $V_{S,30}$ values, and alternatively by using a constant $V_{S,30}$ value for all stations in the region, and find that the variation of site-specific correlations remains significant. This suggests that the variation in GMM prediction caused by $V_{S,30}$ is not as significant as the variation in site-specific ground motions. More comprehensive site response modeling (e.g., by using site-specific site response estimates along with a GMM) should lead to a further decrease in such variations.

ACKNOWLEDGMENTS

We thank Chris Van Houtte for help with the ground motion dataset, and Robin Lee and Viktor Polak for help with data used in Figures 8, 10, 14 and 15. This work was supported by the U.S. Geological Survey (USGS) via External Research Program award G20AP00019. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the USGS. This work was partially supported by QuakeCoRE: The New Zealand centre for Earthquake Resilience. This is QuakeCoRE publication number 0641.

References

1. Boore DM, Gibbs JF, Joyner WB, Tinsley JC, Ponti DJ. Estimated Ground Motion From the 1994 Northridge, California, Earthquake at the Site of the Interstate 10 and La Cienega Boulevard Bridge Collapse, West Los Angeles, California. *Bulletin of the Seismological Society of America* 2003; 93(6): 2737–2751.
2. Schiappapietra E, Douglas J. Modelling the Spatial Correlation of Earthquake Ground Motion: Insights from the Literature, Data from the 2016–2017 Central Italy Earthquake Sequence and Ground-Motion Simulations. *Earth-Science Reviews* 2020: 103139. doi: 10.1016/j.earscirev.2020.103139
3. Wesson RL, Perkins DM. Spatial Correlation of Probabilistic Earthquake Ground Motion and Loss. *Bulletin of the Seismological Society of America* 2001; 91(6): 1498–1515.
4. Adachi T, Ellingwood B. Impact of Infrastructure Interdependency and Spatial Correlation of Seismic Intensities on Performance Assessment of a Water Distribution System. In: Proceedings of the 10th International Conference on Applications of Statistics and Probability in Civil Engineering; 2007; Tokyo, Japan.
5. Graves R, Pitarka A. Refinements to the Graves and Pitarka (2010) broadband ground-motion simulation method. *Seismological Research Letters* 2015; 86(1): 75–80.
6. Park J, Bazzurro P, Baker JW. Modeling Spatial Correlation of Ground Motion Intensity Measures for Regional Seismic Hazard and Portfolio Loss Estimation. In: ; 2007; Tokyo, Japan: 8p.

7. Bradley BA. Ground motions observed in the Darfield and Christchurch earthquakes and the importance of local site response effects. *New Zealand journal of geology and geophysics* 2012; 55(3): 279–286.
8. Sokolov V, Wenzel F, Wen KL, Jean WY. On the influence of site conditions and earthquake magnitude on ground-motion within-earthquake correlation: Analysis of PGA data from TSMIP (Taiwan) network. *Bulletin of Earthquake Engineering* 2012; 10(5): 1401–1429. doi: 10.1007/s10518-012-9368-5
9. Wang M, Takada T. Macrospatial Correlation Model of Seismic Ground Motions. *Earthquake Spectra* 2005; 21(4): 1137–1156.
10. Goda K, Hong HP. Spatial Correlation of Peak Ground Motions and Response Spectra. *Bulletin of the Seismological Society of America* 2008; 98(1): 354–365. doi: 10.1785/0120070078
11. Foulser-Piggott R, Stafford PJ. A predictive model for Arias intensity at multiple sites and consideration of spatial correlations. *Earthquake Engineering & Structural Dynamics* 2012; 41(3): 431–451.
12. Goda K. Interevent Variability of Spatial Correlation of Peak Ground Motions and Response Spectra. *Bulletin of the Seismological Society of America* 2011; 101(5): 2522–2531. doi: 10.1785/0120110092
13. Loth C, Baker JW. A spatial cross-correlation model for ground motion spectral accelerations at multiple periods. *Earthquake Engineering & Structural Dynamics* 2013; 42(3): 397–417. doi: DOI: 10.1002/eqe.2212
14. Heresi P, Miranda E. Uncertainty in intraevent spatial correlation of elastic pseudo-acceleration spectral ordinates. *Bulletin of Earthquake Engineering* 2019; 17(3): 1099–1115.
15. Jayaram N, Baker JW. Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering & Structural Dynamics* 2009; 38(15): 1687–1708.
16. Du W, Wang G. Intra-Event Spatial Correlations for Cumulative Absolute Velocity, Arias Intensity, and Spectral Accelerations Based on Regional Site Conditions. *Bulletin of the Seismological Society of America* 2013; 103(2A): 1117–1129. doi: 10.1785/0120120185
17. Chen Y, Baker JW. Spatial Correlations in CyberShake Physics-Based Ground-Motion Simulations. *Bulletin of the Seismological Society of America* 2019; 109(6): 2447–2458. doi: 10.1785/0120190065
18. Kuehn NM, Abrahamson NA. Spatial correlations of ground motion for non-ergodic seismic hazard analysis. *Earthquake Engineering & Structural Dynamics* 2020; 49(1): 4–23.
19. Van Houtte C, Bannister S, Holden C, Bourguignon S, McVerry G. The New Zealand strong motion database. *Bulletin of the New Zealand Society for Earthquake Engineering* 2017; 50(1): 1–20.
20. Abrahamson N, Gregor N, Addo K. BC Hydro ground motion prediction equations for subduction earthquakes. *Earthquake Spectra* 2016; 32(1): 23–44.
21. Chiou BSJ, Youngs RR. Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra* 2014; 30(3): 1117–1153.
22. Abrahamson N, Atkinson G, Boore D, et al. Comparisons of the NGA ground-motion relations. *Earthquake Spectra* 2008; 24(1): 45–66. doi: 10.1193/1.2924363
23. Douglas J, Akkar S, Ameri G, et al. Comparisons among the five ground-motion models developed using RESORCE for the prediction of response spectral accelerations due to earthquakes in Europe and the Middle East. *Bulletin of Earthquake Engineering* 2014; 12(1): 341–358. doi: 10.1007/s10518-013-9522-8
24. Abrahamson NA, Youngs R. A stable algorithm for regression analyses using the random effects model. *Bulletin of the Seismological Society of America* 1992; 82(1): 505–510.
25. Jayaram N, Baker JW. Considering spatial correlation in mixed-effects regression and the impact on ground-motion models. *Bulletin of the Seismological Society of America* 2010; 100(6): 3295–3303. doi: 10.1785/0120090366

26. Stafford PJ. Evaluation of structural performance in the immediate aftermath of an earthquake: a case study of the 2011 Christchurch earthquake. *International Journal of Forensic Engineering* 2012; 1(1): 58–77.
27. Jayaram N, Baker JW. Statistical tests of the joint distribution of spectral acceleration values. *Bulletin of the Seismological Society of America* 2008; 98(5): 2231–2243.
28. Fosdick BK, Raftery AE. Estimating the correlation in bivariate normal data with known variances and small sample sizes. *The American Statistician* 2012; 66(1): 34–41.
29. Bowley A. The standard deviation of the correlation coefficient. *Journal of the American Statistical Association* 1928; 23(161): 31–34.
30. Fisher RA. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* 1915; 10(4): 507–521.
31. Efron B, Tibshirani RJ. *An introduction to the bootstrap*. CRC press . 1994.
32. Bradley BA. Site-specific and spatially-distributed ground-motion intensity estimation in the 2010–2011 Canterbury earthquakes. *Soil Dynamics and Earthquake Engineering* 2014; 61: 83–91.
33. Bradley BA, Cubrinovski M. Near-source strong ground motions observed in the 22 February 2011 Christchurch earthquake. *Seismological Research Letters* 2011; 82(6): 853–865.
34. Bradley BA. Strong ground motion characteristics observed in the 13 June 2011 Mw6. 0 Christchurch, New Zealand earthquake. *Soil Dynamics and Earthquake Engineering* 2016; 91: 23–38.
35. Wotherspoon LM, Orense RP, Bradley BA, Cox BR, Wood CM, Green RA. Soil profile characterisation of Christchurch central business district strong motion stations. *Bulletin of the New Zealand Society for Earthquake Engineering* 2015; 48(3): 146–156.
36. Bradley BA, Wotherspoon LM, Kaiser AE, Cox BR, Jeong S. Influence of Site Effects on Observed Ground Motions in the Wellington Region from the Mw 7.8 Kaikōura, New Zealand, Earthquake. *Bulletin of the Seismological Society of America* 2018; 108(3B): 1722–1735.
37. Bradley BA, Razafindrakoto HN, Polak V. Ground-motion observations from the 14 November 2016 M w 7.8 Kaikoura, New Zealand, earthquake and insights from broadband simulations. *Seismological Research Letters* 2017; 88(3): 740–756.
38. Stafford PJ, Zurek BD, Ntinalexis M, Bommer JJ. Extensions to the Groningen ground-motion model for seismic risk calculations: component-to-component variability and spatial correlation. *Bulletin of Earthquake Engineering* 2019; 17(8): 4417–4439.

How to cite this article: Chen Y., Bradley B.A., and Baker J.W. (2021), Non-stationary Spatial Correlation in New Zealand Strong Ground-Motion Data, *Earthquake Engng Struct Dyn*, 50(13), 3421-3440.

APPENDIX

A EARTHQUAKES USED IN THE STUDY

TABLE A1 Summary of considered earthquakes in the Christchurch Region. CuspID is an event index value from the original database¹⁹. The number of usable stations are for $SA(1\text{ s})$ data.

Id	CuspID	Event Date	Magnitude	Type	# of usable stations
166	3711648	2012-05-25	5.1	Crustal	26
75	2016p118944	2016-02-14	5.8	Crustal	24
162	3631359	2011-12-23	5.8	Crustal	23
167	3734186	2012-07-06	4.7	Crustal	22
24	2012p764736	2012-10-09	4.0	Crustal	22
163	3631380	2011-12-23	5.8	Crustal	22
23	2012p713691	2012-09-21	4.2	Crustal	21
160	3528810	2011-06-13	5.3	Crustal	21
54	2014p237547	2014-03-29	3.8	Crustal	19
65	2014p965622	2014-12-24	4.0	Crustal	18
159	3468635	2011-02-22	5.6	Crustal	17
153	3366146	2010-09-03	7.1	Crustal	16
161	3528839	2011-06-13	6.0	Crustal	16
155	3391440	2010-10-18	4.8	Crustal	15
64	2014p933966	2014-12-12	4.2	Crustal	15
27	2012p801609	2012-10-23	4.0	Crustal	14
157	3437105	2010-12-25	4.7	Crustal	13
158	3468575	2011-02-21	6.2	Crustal	10
33	2013p368016	2013-05-17	3.9	Crustal	8
76	2016p858000	2016-11-13	7.8	Crustal	8
81	2122842	2003-09-29	4.6	Crustal	7
53	2014p211339	2014-03-20	3.5	Crustal	7
60	2014p686520	2014-09-12	4.1	Crustal	5
107	2472534	2005-10-13	4.5	Crustal	3
69	2015p080815	2015-01-30	3.8	Crustal	3

TABLE A2 Summary of considered earthquakes in the Wellington Region. CuspID is an event index value from the original database¹⁹. The number of usable stations are for $SA(1\text{ s})$ data.

Id	CuspID	Event Date	Magnitude	Type	# of usable stations
37	2013p543824	2013-07-21	6.6	Crustal	26
76	2016p858000	2016-11-13	7.8	Crustal	25
45	2013p614135	2013-08-16	5.9	Crustal	24
43	2013p613797	2013-08-16	6.6	Crustal	24
36	2013p542711	2013-07-20	5.7	Crustal	23
169	3765940	2012-09-01	3.8	Crustal	23
41	2013p563639	2013-07-28	4.9	Crustal	21
44	2013p613947	2013-08-16	5.2	Crustal	20
95	2354133	2005-01-20	5.3	Slab	19
35	2013p538215	2013-07-19	4.5	Crustal	16
72	2015p302557	2015-04-22	4.7	Crustal	16
156	3413873	2010-11-28	5.0	Slab	15
49	2014p022834	2014-01-09	3.7	Crustal	12
6	1502698	2000-03-29	5.2	Slab	10
50	2014p051675	2014-01-20	6.3	Slab	9
40	2013p561823	2013-07-27	4.0	Crustal	6
4	1348340	1999-01-03	5.2	Slab	4
109	2480685	2005-11-01	4.5	Crustal	4
177	722185	1995-03-22	5.8	Slab	3
38	2013p551065	2013-07-23	4.2	Crustal	2
110	2481737	2005-11-03	4.5	Crustal	2

B OTHER PERIOD RESULTS

B.1 Christchurch SA(0.25 s)

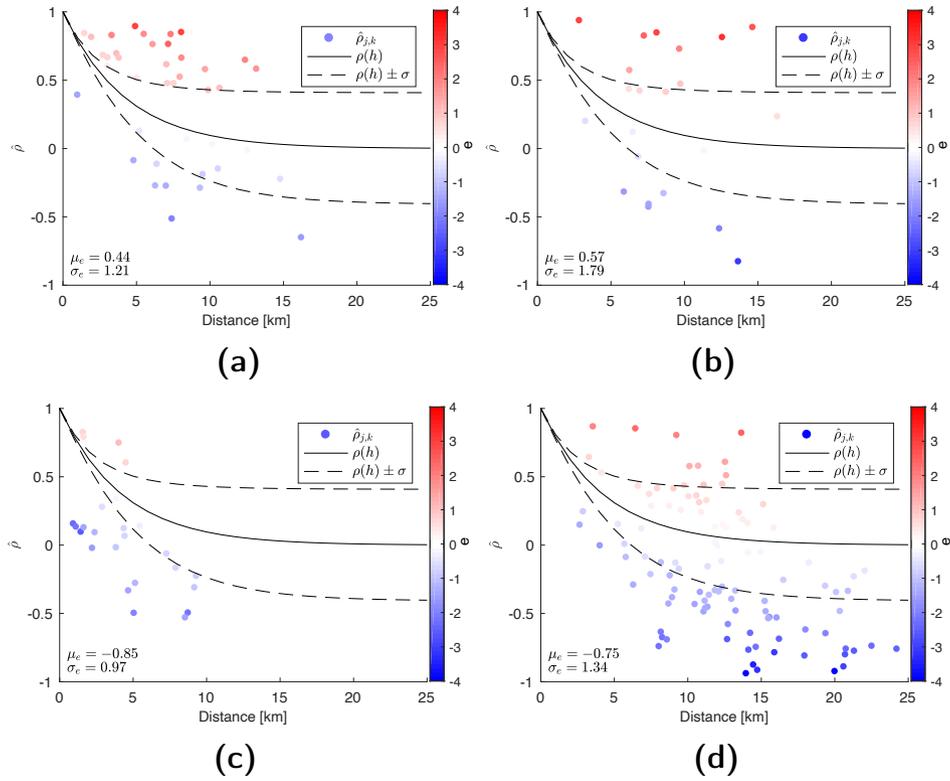


FIGURE B1 Site-specific correlations and deviations for $SA(0.25\text{ s})$ within-event residuals versus separation distance for Christchurch station pairs (a) in the south subregion, (b) in the north & west subregion, (c) in the CBD & east subregion, (d) with one in the south subregion and another outside. (See subregions in Figure 2b).

B.2 Wellington SA(0.25 s)

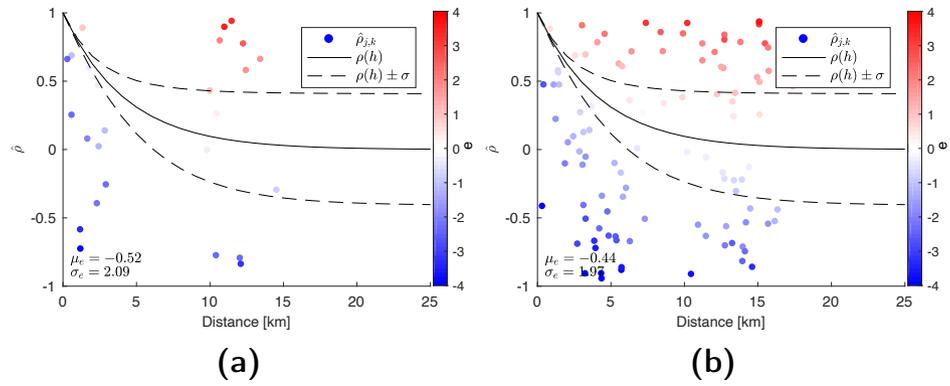


FIGURE B2 Site-specific correlations and deviations for $SA(0.25\text{ s})$ within-event residuals versus separation distance for Wellington station pairs (a) in the basin subregions (Thorndon, Te Aro, and Lower Hutt Basin), (b) with one in the basin subregions and another outside. (See subregions in Figure 2a).

B.3 Christchurch SA(0.01 s)

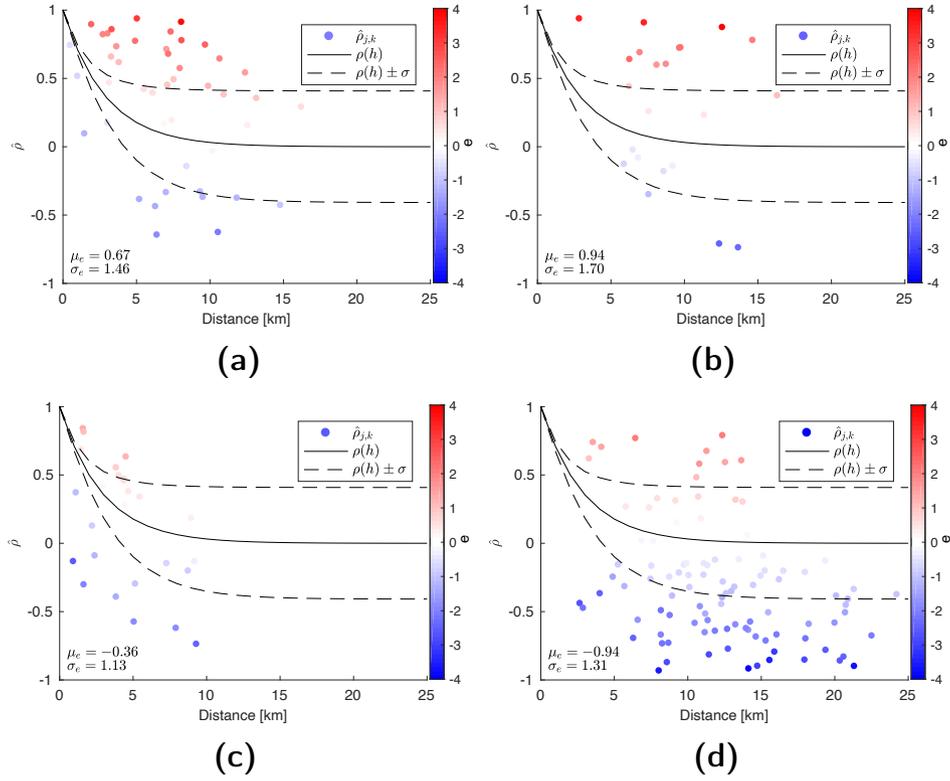


FIGURE B3 Site-specific correlations and deviations for $SA(0.01\text{ s})$ within-event residuals versus separation distance for Christchurch station pairs (a) in the south subregion, (b) in the north & west subregion, (c) in the CBD & east subregion, (d) with one in the south subregion and another outside. (See subregions in Figure 2b).

B.4 Wellington SA(0.01 s)

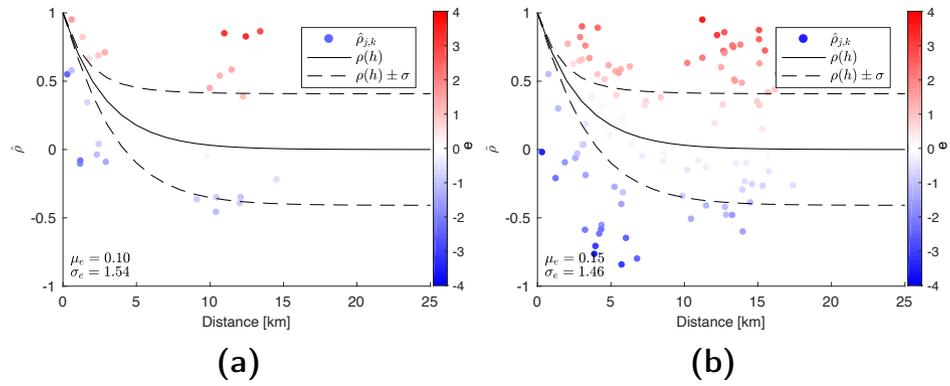


FIGURE B4 Site-specific correlations and deviations for $SA(0.01\text{ s})$ within-event residuals versus separation distance for Wellington station pairs (a) in the basin subregions (Thorndon, Te Aro, and Lower Hutt Basin), (b) with one in the basin subregions and another outside. (See subregions in Figure 2a).