Application of AI to Advance Structural Performance and Resiliency Quantification

Final Report for 2020 Biennial Research Project (1981/20784)

Principal Investigator

Dr Max Stephens, University of Auckland

Research Team:

Amin Ghasemi, University of Auckland

Prof Ken Elwood, Toka Tū Ake EQC/MBIE

Nick Horspool (GNS Science)

Date: March 2022

Executive Summary

Developing a robust structural assessment framework which can rapidly provide a quantitative seismic assessment of buildings and infrastructure at a regional scale enhances community resiliency by helping to identify critically vulnerable structures and lifelines. Central to the work developed in this project is the use of machine learning driven computational methods to efficiently cluster buildings into typologically similar groups and select representative indicator structures that can be used to estimate the response of a larger building set. Indicator buildings were used following the Christchurch earthquakes to limit the number of structural inspections of damaged buildings following aftershocks. These buildings were selected to include representative typologies in Christchurch and were reinspected following the aftershocks to determine if additional damage occurred. The intention here is to select indicator buildings before the occurrence of ground shaking to allow for the installation of structural instrumentation and development of detailed numerical models that can be used for scenario planning pre-event and can provide situational awareness post-event. Wellington was used as the region of study for this research due to the existence of a comprehensive database of buildings within the CBD. First, additional structural and site characteristics were added to the database including the estimated building periods and site periods. Next, three machine learning clustering methods were developed to cluster buildings in the database into typologically similar groups and select representative indicator buildings. Then, a range of numerical models were generated for the indicator buildings, and the response across all buildings was evaluated using the 2016 Kaikōura earthquake as a case study. Results from the case-study suggest that the indicator building approach can be effective in estimating building drifts and accelerations across a range of buildings. Ongoing work is focused on continued database development and refinement of the indicator building approach for regional seismic response modelling.

Technical Abstract

This research is focused on the development of a multi-level seismic impact framework for regional seismic response and damage assessment. The work presented here uses a detailed database of buildings in Wellington. First, structural and site characteristics that were necessary for development of response frameworks were added to the database including the estimated building periods, design accelerations and site periods. Next, a one-DOF framework that can be used to roughly identify structures where the seismic demands may have exceeded the design accelerations was presented and demonstrated using a small earthquake in Wellington. Then, machine learning driven computational methods to cluster buildings into typologically similar groups and select representative indicator structures were evaluated. Two prominent unsupervised machine learning clustering approaches are utilized to cluster the mixed categorical and numerical building database: namely k-prototype and k-mean. A novel autoencoder deep learning neural network is also designed and trained to convert the mixed data into a low-dimensional subspace called latent space and feed this into k-mean algorithm. The autoencoder method is demonstrated to be more effective at clustering buildings into useful typological clusters for seismic response analysis. Then, the concept of using indicator buildings for regional response modelling is introduced. Indicator buildings within each cluster were selected and modelled, with supplementary models generated by modifying the stiffness of the base indicator building models account for building flexibilities across each cluster (for a total four models per cluster). A case-study was undertaken using the 2016 Kaikōura earthquake, and the response of the models are utilized to estimate the seismic response of all buildings in the database. Results from the case-study suggest that the indicator building approach can be effective in estimating building drifts and accelerations across a range of buildings. Ongoing work is focused on developing methodologies to correct for building flexibility across each building cluster, developing more detailed structural models and instrumenting indicator buildings, and applying additional machine learning techniques for damage and response estimation.

Key Words
Regional Seismic Response Analysis, Representative Buildings Selection, Seismic Vulnerability Archetypes, Unsupervised Machine Learning Clustering, Deep Learning Autoencoder Neural Network, Latent Space

1. Introduction

Buildings and infrastructure systems in New Zealand are vulnerable to damage and disruption from seismic events. Developing a robust structural assessment framework which can rapidly provide a quantitative assessment of these systems at a regional scale enhances community resiliency by helping to identify critically vulnerable structures and lifelines. A number of research and commercial initiatives have been undertaken world-wide to develop regional impact tools for scenario planning and/or loss estimation following seismic events. These tools commonly utilize a fragility-based approach to estimate fatalities, damage and economic loss, based on a general prediction of the dominant behaviours of structures subjected to seismic hazards (e.g. [1]). However, using such tools it is difficult to pinpoint specific critically deficient structures for retrofit, evacuation or inspection.

There is an interest in developing a framework to provide building-specific nonlinear response and damage predictions following earthquakes. However, traditional numerical methods for evaluating the nonlinear seismic response of structures are computationally expensive, making it difficult to scale these methods to evaluate the performance of all structures within a community. This can make it impractical to perform detailed structural resilience analysis at community scale (where hundreds of numerical runs are required for each structure in the community), and difficult to provide rapid community-wide performance assessments immediately following seismic events.

To overcome these deficiencies, this project provides the building blocks for next generation tools for performing regional seismic resiliency and scenario assessments as well as rapid performance and damage estimations following earthquakes. A key concept of this study is the development of linked data sets that can be used to inform the seismic response and damage of a large number of buildings using multiple response and damage evaluation methodologies. The research worked in conjunction with aligned projects supported by QuakeCoRE and RNC2 to compile the information required to develop several regional seismic impact tools. Two tools were developed here: (1) a one-DOF framework that compares the design and imposed spectral accelerations, and (2) a more detailed machine learning driven computational method to efficiently cluster buildings into typologically similar groups and select representative indicator structures that can be used to estimate the response of structures at a regional scale. The following sections of this report provide a summary of the completed work and the report is organised as follows:

- Section 2 provides a summary of the Wellington Building Inventory (WBI) including the building-level structural and site characteristics that were added as part of this work. A one-DOF framework that can be used to roughly identify structures where the seismic demands may have exceeded the design accelerations is also presented and demonstrated for a small earthquake recorded in Wellington in 2020.
- Section 3 introduces machine learning driven computational methods to cluster buildings into typologically similar groups and select representative indicator structures for regional response modelling. Indicator buildings within each cluster were selected and modelled, with supplementary models generated by modifying the stiffness of the base indicator building models account for building flexibilities across each cluster, and the indicator building approach is demonstrated and evaluated using the 2016 Kaikōura earthquake as a case study.
- Section 4 evaluates the efficacy of multi-fidelity nonlinear structural modelling using recorded earthquake data to provide insight into the level of structural modelling required to

adequately capture nonlinear structural response to develop training data for deep learning surrogate models.

- Section 5 discusses recent advances in developing nonlinear structural surrogate models using deep learning model architectures. Two deep learning approaches are recommended for further evaluation and use in the proposed indicator building regional framework.
- Section 6 provides key findings and conclusions, and Section 7 provides a brief overview of ongoing and future work.

2. Wellington Building Inventory and One-DOF Rapid Assessment Framework

The building dataset used in the research is the Wellington Building Inventory (WBI) which was developed in-part through collaboration between QuakeCoRE and Wellington City Council [2]. Figure 1 shows a snapshot of Wellington CBD and the buildings included in the WBI as of 2021. The WBI consists of several different databases with buildings in the Wellington CBD, including Building Seismic Assessment, Hollowcore Floors, Targeted Damage Evaluation, 1935-1975 RC Buildings, CityScope, Earthquake Prone Buildings, and WCC Heritage, which have been combined in a master database and evaluated using a comprehensive site survey. Additionally, the database is enriched by a building drawing database which consists of building drawings and consents for a large number of buildings in the wider database. The readable access to the structural drawing of buildings facilitates the extraction of building structural properties and forms the foundation for further investigation in this research. This work focuses on medium to high rise concrete buildings located at the Wellington CBD. The most important building parameters in the WBI for the purpose of this work were the geographic location, number of storeys, above ground height, year of design and construction, lateral load bearing system, floor system, presence of vertical/horizontal irregularity, use category, and footprint area. However for the purpose of performing a response analysis, additional parameters were necessary to estimate the design forces and seismic demands, namely the site class, site period, building period, and strong motion station. These parameters were added to the WBI as part of this work and are discussed in more detail below.

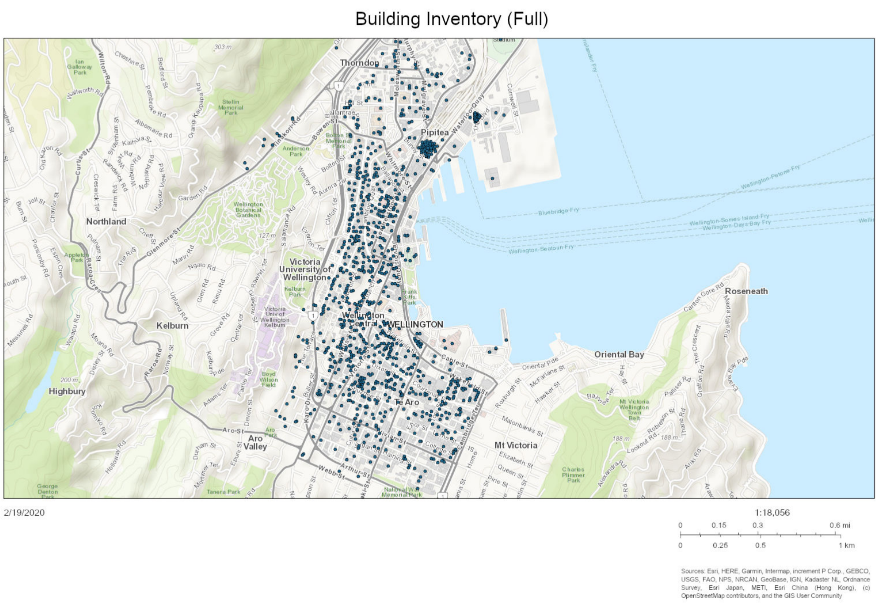


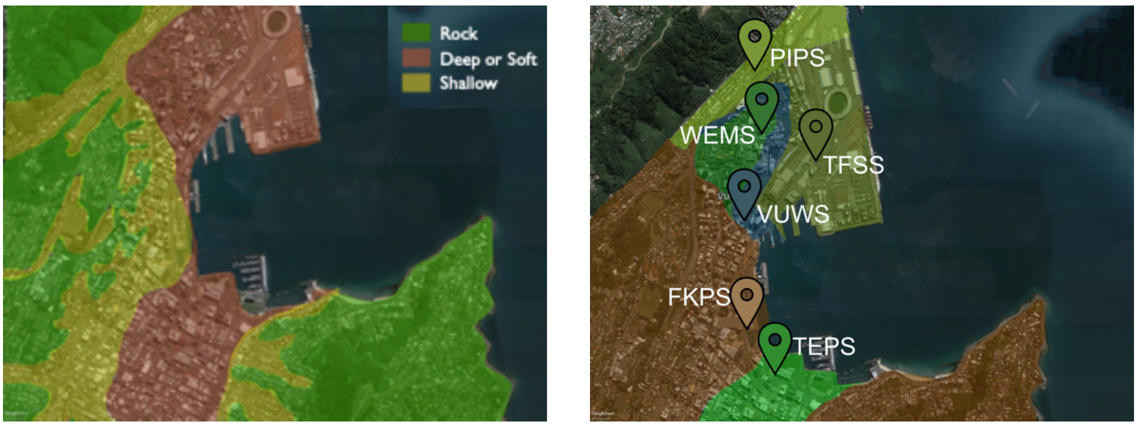
Figure 1. Wellington building inventory [2]

2.1 Estimating Building Design Accelerations

Several important parameters required to estimate the design spectral accelerations for each building were not included in the original WBI, namely the site subsoil class and structural period. This research utilised the available site subsoil class maps for central Wellington presented by Kaiser et al. [3] to identify the buildings' soil types based on the geographic location. Figure 2a shows the Wellington CBD including the boundaries of different site subsoil classes according to NZS 1170.5 [4], where soil type B is for Rock, C is for shallow soil, and D is for deep or soft soil, while Figure 3a shows the number of buildings in the database assigned to each site class. The building period is essential for the evaluation of earthquake design action. Consequently, this research used the analytical approach proposed in the Commentary of NZS1170.5 [5] to estimate the approximate period of the building based on the height. Adding site subsoil class and building period to the database, the design acceleration of each building is calculated in accordance with NZS 1170.5 [5] and added to the database for use in the framework.

2.2 Strong Motion Stations

The other important parameter which was not available in the original WBI are the building site periods, which are used to link strong motion stations to individual buildings to estimate seismic demand. Since the Wellington CBD has very dense site period contours, the change in site periods of buildings over very short distances can be extreme. Hence, the geotechnical characteristics of buildings are more related to building site periods than relative geographic location to the nearest strong motion station. For this reason, the site period contour map of the Wellington CBD from Kaiser et al. [3] was modified (by adding additional contours) and used to link buildings to the appropriate strong motion station, where the nearest geographic station with the closest site period was linked to each building. Figure 2b shows the Wellington CBD including the boundaries of different strong motion stations, while Figure 3b shows the number of buildings in the database assigned to each strong motion station. From Figure 3b, it is clear additional strong motion stations should be added to central Wellington to provide a better estimate of individual building-level seismic demands following earthquakes. These findings prompted an aligned pilot project that has been funded through QuakeCoRE (agreement number 3723086) that was focused on evaluating the optimized locations for strong motion stations in urban environments with rapidly changing site conditions. That work has been completed, and additional work will be required before any recommendations can be made.



(a) (b)

Figure 2. (a) Site class and (b) strong motion station boundaries in the Wellington CBD [6]

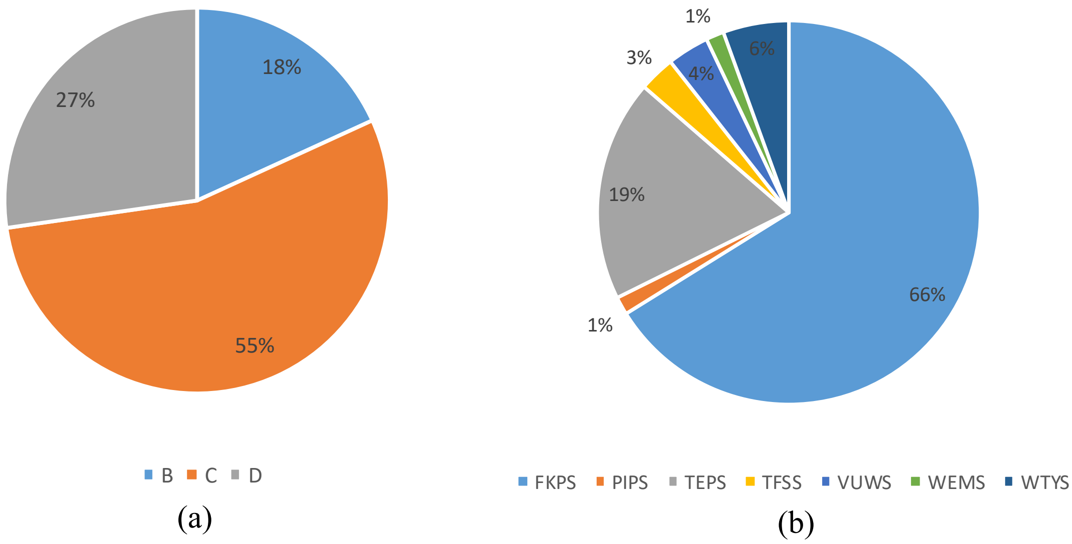
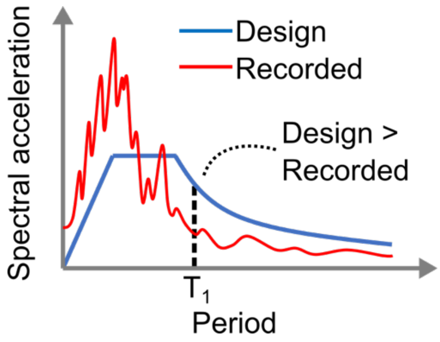


Figure 3. Distribution of (a) site subsoil class and (b) strong motion stations within the WBI [7]

2.3 One-DOF Rapid Assessment Framework

Using the building period, year of construction, site class, and strong motion station, a one-DOF rapid assessment framework was developed as illustrated in Figure 4a. Within this framework, the design accelerations for each structure (as estimated using historical data or based on site class and location) are compared to the spectral accelerations at the fundamental period of the structure recorded by the strong motion station that has been linked to each building. The initial version of the framework is designed as an offline webpage platform that can display the percentage of the measured spectral acceleration at the fundamental period of the structure to the design acceleration. A trial of the framework was created for the relatively small Foxton 5.8 M_w earthquake occurred at 24 May 2020, 19:53:33 UT. The buildings of central Wellington database are shaded according to a colour scale to represent the ratio of measured earthquake acceleration to the Ultimate Limit State (ULS) design acceleration as shown in Figure 4b. The map is designed interactively which enables the user to click on each building and view the building name and the corresponding demand ratio, $T_{1,design}/ T_{1,measured}$. This map also includes the locations of Wellington strong motion stations and allows the user to access information about each station including the name and site subsoil class . As shown in Figure 4b, the map displays a set of spectrum for each strong motion station, including design ULS and Serviceability Limit State (SLS) spectral acceleration spectrum for the corresponding soil type, as well as the spectra for both components of acceleration measured at the station for the earthquake.



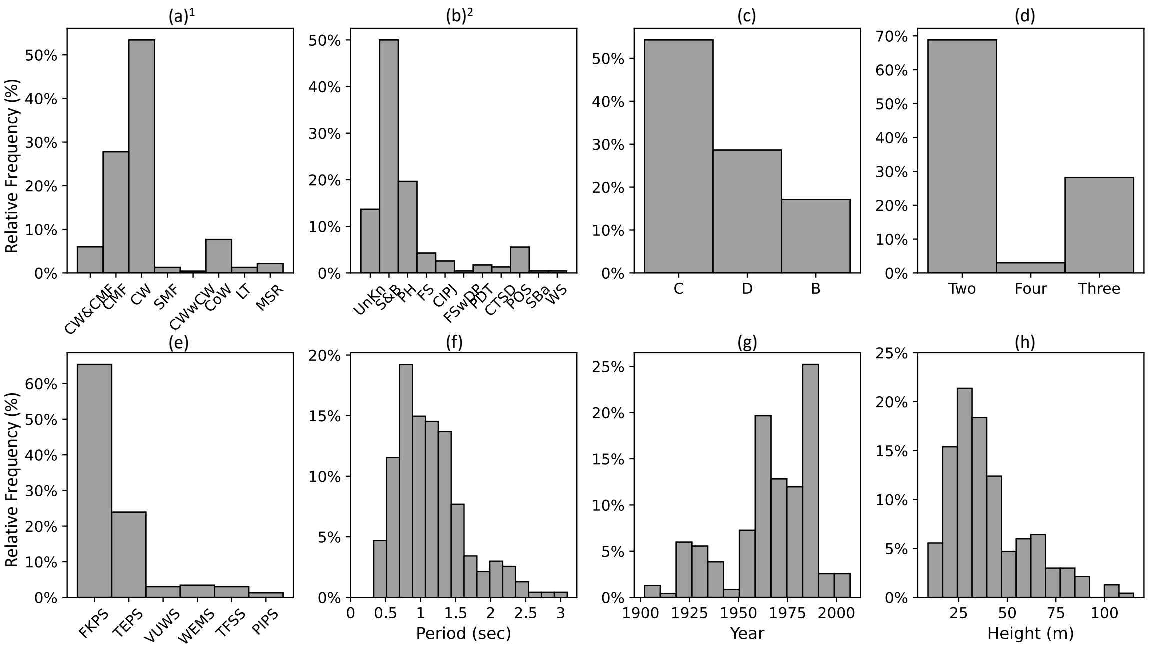
(a)

(b)

Figure 4. Distribution of (a) overview of one-DOF analysis approach and (b) snapshot of near real-time output following the 2020 Foxton Earthquake. Note the blue footprint indicates data is not available.

3. Machine Learning Driven Building Clustering for Regional Seismic Response and Damage Analysis

This section presents a machine learning driven framework to cluster buildings into typologically similar groups and select indicator buildings for regional seismic response analysis. The framework requires a robust database of buildings to provide high level structural and site information of buildings. Here, a reduced data set of 234 reinforced concrete buildings from the WBI were used to demonstrate the concept. First, key structural and site parameters that contribute to the seismic demand and structural response of each building were extracted from the database. Extracted parameters include three numerical and five categorical attributes of each building including year of construction, height, period, lateral load resisting system, floor system, site subsoil class, importance level, and strong motion station. The relative frequency of these attributes across the reduced building set of 234 buildings is illustrated in Figure 5. Next, two prominent unsupervised machine learning clustering approaches were utilized to cluster the mixed categorical and numerical building database: k-prototype on the mixed numerical and categorical database and k-mean on principal components numerical subspace adopted from Factor Analysis of Mixed Data (FAMD) (herein referred to as k-mean on FAMD). A novel autoencoder deep learning neural network was also designed and trained to convert the mixed data into a low-dimensional subspace called latent space and feed this into k-mean for clustering (herein referred to as k-mean on latent space). An overview of the neural network architecture developed here is shown in Figure 6, where IL, FS, LS, ST, and SMS represent importance level, floor system, lateral system, site subsoil class, and strong motion station respectively. Detailed information on the clustering techniques developed and implemented in the work can be found in Ghasemi and Stephens [8].



[1]Lateral systems:: CW&CMF: Concrete Walls and Concrete Moment Frame, CMF: Concrete Moment Frame, CW: Concrete Walls, SMF: Steel Moment Frame, CWwCW: 'Core Wall with Concrete Walls, CoW: Core Wall, LT: Light Timber, MSR: Masonry

[2]Floor systems:: UnKn: Unknown, S&B: Slab and Beams, PH: Precast Hollowcore, FS: Flat Slab, CIPJ: Cast In Place Joists, FSwDP: Flat Slab with Drop Panels, PDT: Precast Double T , CTSD: Concrete Topping on Steel Deck, POS: Precast Other Systems, Sba: Slab and Beam, WS: Waffle Slabs

Figure 5. Relative frequency of building attributes in the database. (a) Lateral System (LS), (b) Floor System (FS), (c) Soil Type (ST), (d) Importance Level (IL), (e) Strong Motion Station (SMS), (f) Period, (g) Year, and (h) Height.

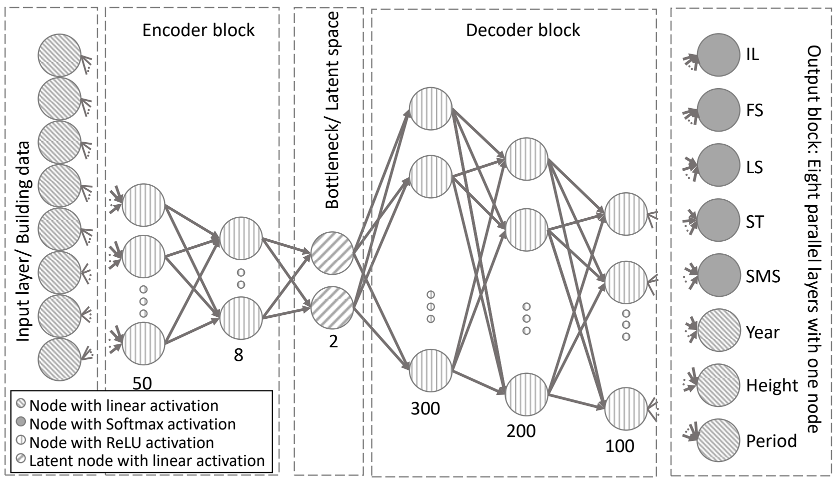


Figure 6. Architecture of designed autoencoder neural network [8]

3.1 Clustering Results – Data Science Methods

The results of three clustering methods were investigated and compared using common data-science methods including principle components (Figure 7) and t-Distributed Stochastic Neighbour embedding (Figure 8). A detailed description of these comparison methods and a more complete discussion regarding the data-science comparison of the building data is included in Ghasemi and Stephens [8]. What is important to note here are the distinct clusters developed using the novel autoencoder deep learning approach in both cases, which demonstrates that this method is most effective at clustering buildings based on the selected site and structural parameters.

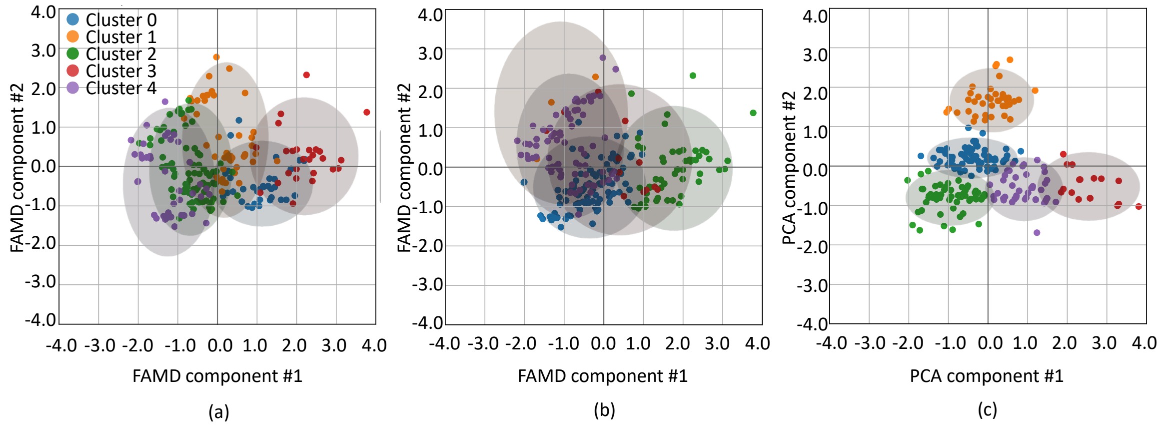


Figure 7. Scatter of buildings on principal component space (a) FAMD space for k-prototype, (b) FAMD space for k-mean on FAMD space, and (c) PCA space for k-mean on latent space clustering technique. Cluster labels identified by different colors [7].

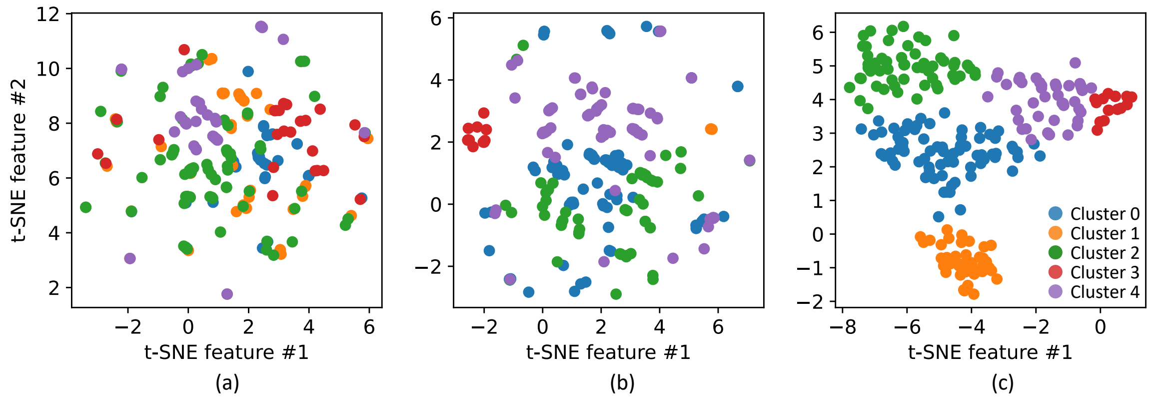


Figure 8. Scatter of buildings on the 2D t-SNE subspace adopted from (a) 34D FAMD space for k-prototype clustering, (b) 18D FMAD space for k-mean FAMD space clustering, and (c) latent space for k-mean clustering on latent space clustering. Cluster labels identified by colors [8].

3.2 Clustering Results – Engineering Comparison

The dominant structural and site characteristics in each cluster were evaluated to determine how the clusters could efficiently be used for regional seismic response assessment. Figure 9 shows the clustering results comparing the numerical data year, height, and period, where again the data points are coloured based on the cluster label. From Figure 9, it is clear that the k-prototype and k-means on latent space method results in distinct clusters with regards to year, height, and estimated periods, while the k-means on FAMD method results in significant overlap in buildings in different clusters. This suggests that the k-prototype and k-means on latent space methodologies are more effective when clustering buildings for the purpose of seismic response analysis, as structural response is highly dependent on height and period as well as the era of design, which can be identified based on the year.

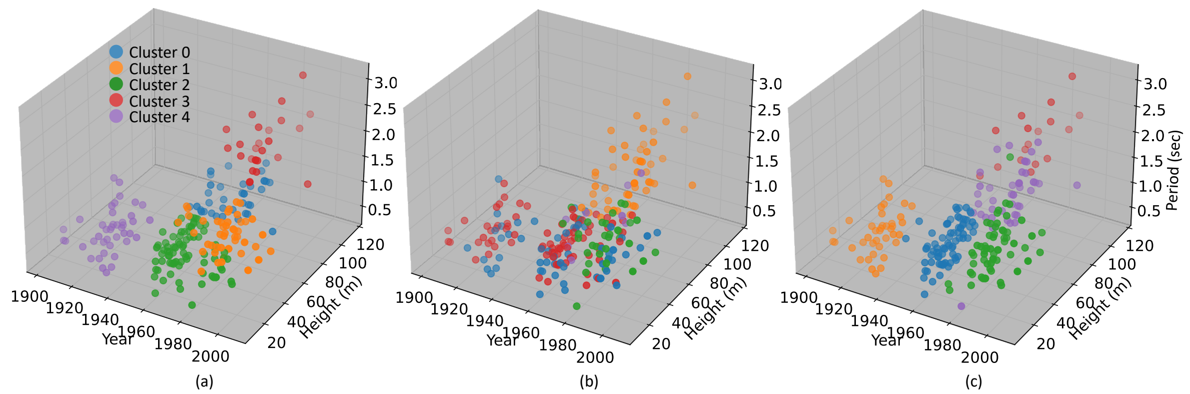


Figure 9. 3D Scatter of buildings on year, height, and period axes and colours in accordance with the clustering labels from (a) k-prototype, (b) k-mean on FAMD space, and (c) k-mean on latent space.

To further evaluate how the different methods clustered the numerical attributes, the relative frequency distribution of height and year within each cluster are illustrated in Figure 10. From Figure 10, it is clear that the clusters developed using k-prototype and k-means on FAMD include a larger range of building heights, and years than clusters developed using k-means on latent space. The difference is especially significant in the distribution of year where the bin widths in k-mean of latent space are mush shorter than the two other methods. The tighter bin widths for the numerical data in the clusters developed using k-means on latent space demonstrate that this method is more useful for seismic response analysis, where the seismic demands are dependent on the fundamental period (which is in-part based on height), and vulnerabilities are commonly linked to year of design.

The categorical attributes in the clustering results were also investigated using relative frequencies of each attribute in the clusters. Figure 11 shows the relative frequencies of LS and FS in the unpacked clusters of each method. In particular these attributes have been selected for investigation here since they most significantly influence seismic response and damage. From Figure 11 it is clear that in general a majority of buildings in each cluster use either CW and CMF lateral systems while the dominant floor systems are S&B and PH, which is consistent with the distribution of the input data. The k-means on latent space method resulted in a significant number of wall buildings (>40%) in four clusters (Cluster 0, 1, 2, and 4), with a relatively wide distribution of lateral systems in Cluster 3. The k-prototype method resulted in a significant number of wall buildings in three clusters (>40%) (Clusters 0, 2, and 4) and concrete moment frames (~50%) in two clusters (Clusters 1 and 3). The k-means on FAMD method resulted in a significant number of wall buildings (>40%) in Clusters 0, 1, and 3 and concrete moment frame (~60%) in Cluster 3. The distribution of lateral systems in Cluster 4 of this method is fairly distributed. The distribution of floor systems in each cluster reveals that k-mean on latent space has S&B as the significant FS (>50%)) in three clusters (Cluster 0, 1, and 2), precast PH (~40%)) in Cluster 2 and relatively distributed floor systems in cluster 4. For k-prototype, Clusters 0, 2, and 4 have S&B as the significant value (>50%) and Clusters 1 and 3 have PH as the significant floor system value (~50%). The k-mean on FAMD also resulted in three clusters (Cluster 0, 3, and 4) with significant number of S&B (>40%), Cluster 2 with significant number of PH (~60%), and Cluster 1 with nearly 40% of S&B and PH separately.

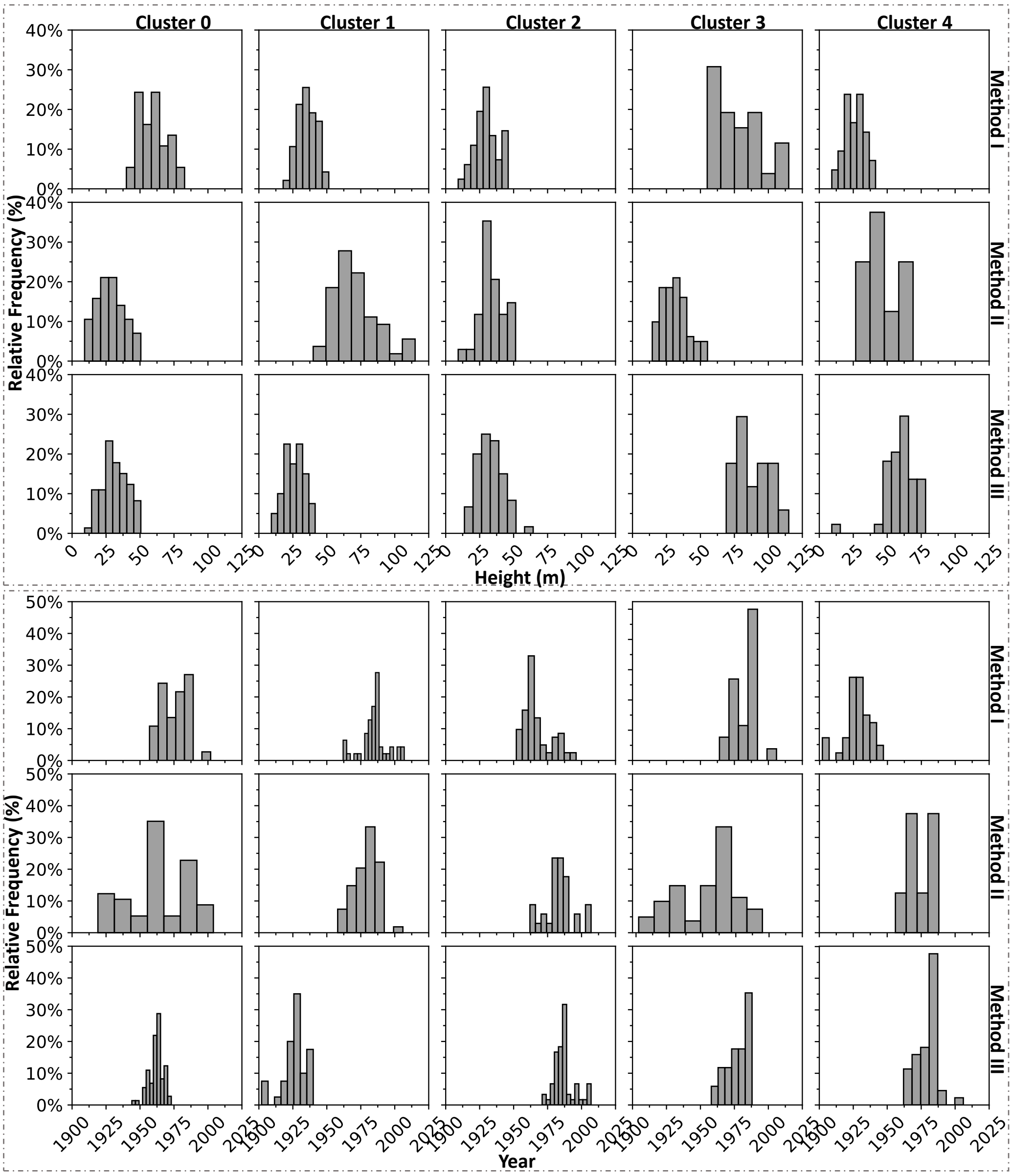


Figure 10. Relative frequency of height (top) and year (bottom) for buildings in each cluster of each adopted clustering technique. Columns represent cluster number and rows represent the clustering technique.

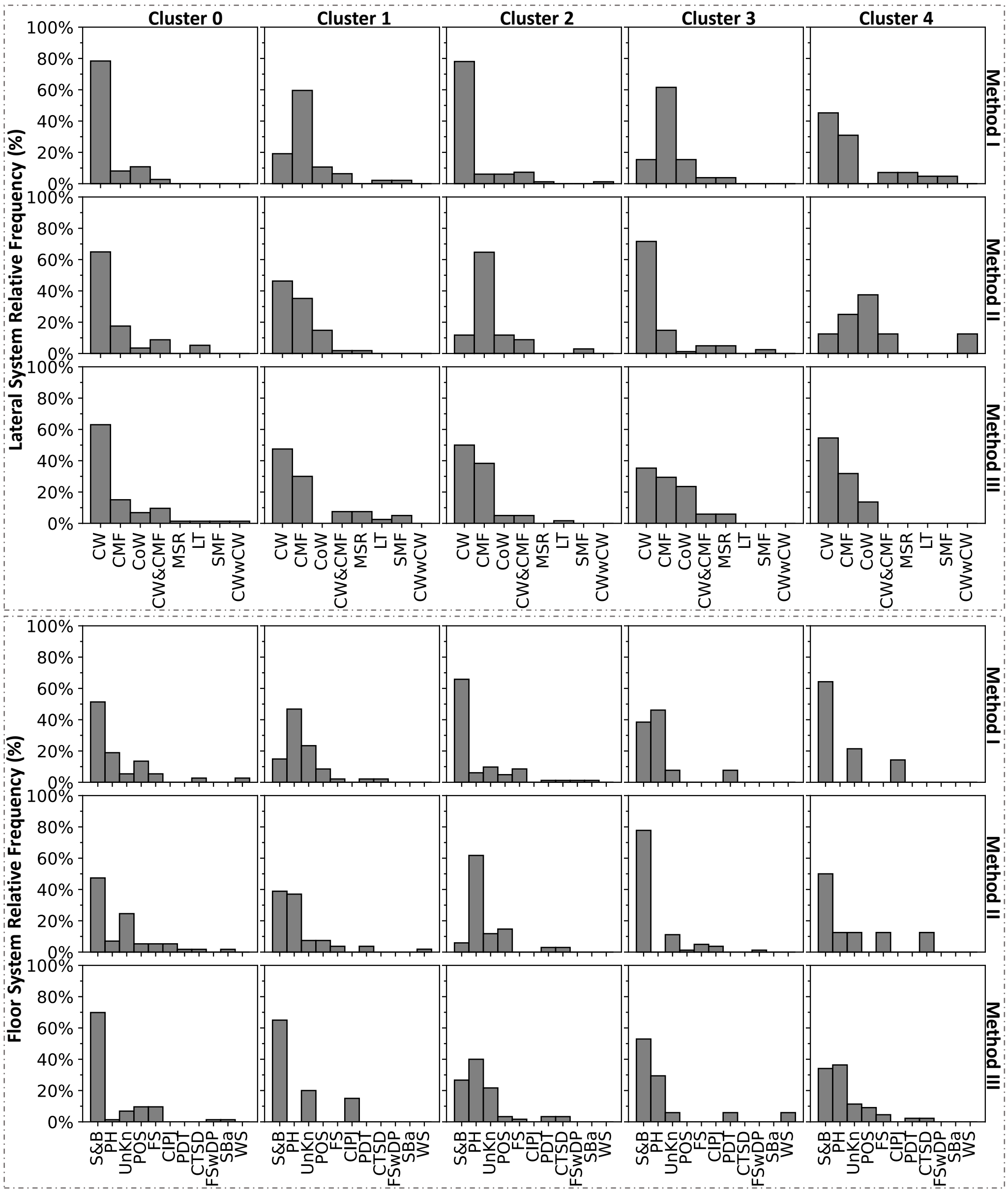


Figure 11. Distribution of lateral system (top) and floor system (bottom) for buildings in each cluster of each adopted clustering technique. Columns represent cluster number and rows represent the clustering technique.

3.3 Selection of Representative 'Indicator' Buildings

Using each of the clustering methods, buildings can be sorted by their distance to the corresponding cluster centroid which leads to a prioritized list of typologically representative buildings for each cluster. Indicator buildings were used following the Christchurch earthquakes to limit the number of structural inspections of damaged buildings following aftershocks. These buildings were selected to include representative typologies in Christchurch and were reinspected following aftershocks to determine if additional damage occurred. The intention here is to select indicator buildings before the occurrence of ground shaking to allow for the installation of structural instrumentation and development of detailed numerical models that can be used for scenario planning pre-event and can provide situational awareness post-event. Table 1 lists the attributes of the selected representative buildings as selected using each clustering technique. To evaluate how well the selected representative buildings in each cluster reflect the dominant structural and site characteristics, the distribution of numerical and categorical attributes across the entire data set (shown in Figure 5) and within each cluster (shown in Figure 10 and Figure 11) are used. Note that the most frequent bin(s) in numerical attributes, *i.e*. year and height, are selected as the dominant ranges of numerical attributes and the values with relative frequency greater than 20% in categorical attributes, *i.e*., lateral systems and floor systems, are referred to as the dominant categorical values of each cluster. The dominant range of buildings year and height and the dominant values of LS and FS in each cluster are tabulated in Table 1.

Table 1. Characteristics of selected representative buildings and dominant ranges/values in each cluster from each adopted clustering technique. Numbers in parenthesizes indicate the percentage of corresponding value in the cluster adopted from Figure 11.

		Representative buildings				Dominant cluster ranges/values			
Clustering technique	Cluster	Year	Height (m)	LS*	FS*	Year	Height (m)	LS*	FS*
Method I: K-prototype	0	1971	58.7	CW	S&B	[1982-1988]	[47-52] [60-65]	CW (78%)	S&B (51%)
	1	1984	36.8	CW&CMF	PH	[1985-1987]	[28-36]	CMF (59%)	PH (45%)
	2	1969	28.5	CW	S&B	[1960-1964]	[24-29]	CW (79%)	S&B (67%)
	3	1979	78.2	CMF	S&B	[1984-1990]	[26-37]	CMF (61%)	PH (45%) S&B (38%)
	4	1926	24.6	CW	S&B	[1923-1935]	[20-32]	CW (45%) CMF (30%)	S&B (70%)
Method II: K-mean on FAMD space	0	1952	27.6	CW	S&B	[1959-1965]	[20-32]	CW (65%)	S&B (47%)
	1	1958	69	CoW	WS	[1980-1985]	[60-68]	CW (46%) CMF (35%)	S&B (38%) PH (37%)
	2	1996	41.4	CMF	CTSD	[1979-1986]	[27-33]	CMF (65%)	PH (62%)
	3	1956	25.7	CW	FSwDP	[1954-1960]	[30-35]	CW (70%)	S&B (78%)
	4	1976	68	CW	CTSD	[1963-1971] [1980-1988]	[37-43]	CoW (38%) CMF (25%)	S&B (50%)
Method III: K-mean on latent space	0	1963	32.2	CW&CMF	S&B	[1962-1965]	[25-30]	CW (61%)	S$B (70%)
	1	1926	23	CW	S&B	[1926-1930]	[19-24] [29-34]	CW (44%) CMF (30%)	S&B (65%)
	2	1986	34	CMF	PH	[1985-1988]	[28-34]	CW (50%) CMF (39%)	PH (40%)
	3	1972	88.8	CoW	PH	[1982-1987]	[76-86]	CW (34%) CMF (30%) CoW (25%)	S&B (52%) PH (30%)
	4	1981	59.8	CW	S&B	[1981-1987]	[59-65]	CW (53%) CMF (30%)	PH (36%) S&B (34%)

* LS and FS values are explained in the legend of Figure 5.

With regards to numerical attributes, the representative buildings selected using k-means on latent space were more reflective of the characteristics within the clusters than the buildings selected using k-prototype and k-means on FAMD. For k-prototype and k-means on FAMD, only one representative building out the five clusters had a year of construction within the dominant range within the cluster, while for k-means on latent space three buildings had a year within the dominant range. Similar trends can be observed for building height, where only one representative building had a height within the dominant range of the cluster for k-prototype and k-means on FAMD, while three buildings were within the dominant height range using k-means on latent space. Further, it should be noted that when using k-prototype and k-means on FAMD the representative buildings with numerical attributes outside the dominant ranges tended to be well outside the dominant range (*e.g*. the height of the representative building for Cluster 3 was 78.2m while the dominant height range in the cluster was 26-32m). In contrast for the representative buildings selected using k-means on latent space with attributes outside the dominant range tended to be relatively close to that range (*e.g* the year of representative building for Cluster 3 was 1972 while the corresponding dominant year range in the cluster was 1982-87). These comparisons highlight the advantage of k-mean on latent space in that it results in narrower ranges of numerical attributes across the clusters while also selecting indicator buildings that better represent the buildings within each cluster for seismic response and vulnerability assessment.

The other important attributes in seismic vulnerability of buildings such as LS and FS were also better reflected in the representative buildings of k-mean on latent space. According to Table 1, k-means on FAMD was not effective in selecting representative buildings within the dominant LS and FS across the clusters, *e.g*., FS of cluster 1 representee building was waffle slab which had a very negligible relative frequency and FS of cluster 0 representative building was concrete wall which had the lowest relative frequency in that cluster. It can be concluded that k-prototype and k-mean on latent space were fairly effective at selecting truly indicator buildings in terms of the categorical attributes. Moreover, the better performance of k-mean on latent space in selecting more effective representative buildings in terms of numerical attributes was also concluded previously. Consecutively, the representative buildings of k-mean on latent space were selected for further investigation on regional vulnerability archetypes which is discussed shortly. Note that in cases where there are multiple attributes with larger than 20% representation in a single cluster, it may be necessary to select multiple representative buildings for response and damage analysis, *e.g*. for cluster 1 where concrete walls and moment frames had more than 30% relative frequency.

3.4 Application of 'Indicator' Buildings for Regional Response Evaluation

The key parameters influencing the seismic response and vulnerability of each cluster resulting from k-mean on latent space clustering was investigated to evaluate how effectively representative buildings can be used to estimate the seismic behavior of all clusters in terms of response (*e.g*. drift and accelerations) and damage (based on previously defined vulnerabilities). Here, the attributes most likely to affect the seismic response were taken as the year, period, and LS. From Table 1, it is clear the year and periods of the representative buildings effectively represent the years and periods within each cluster more broadly. However within clusters 1, 2, and 4, there are multiple LS with greater than 20% representation within the clusters, which would make it difficult to select a single representative building to estimate the seismic response and damage across the cluster (*e.g*. Cluster 1 with CW and CMF which consist 44% and 30% of buildings LS respectively). In these cases, it may be necessary to select additional representative buildings to ensure all dominant lateral systems are captured by the representative building response. This will be the topic of study in later work. However, additional buildings were selected from the prioritized list to include representative buildings with the dominant characteristics in each cluster. These additional buildings have been summarized in Table 2 and were selected to accurately represent the cluster in terms of all important attributes.

To address the effectiveness with which the representative buildings would be expected to capture the damage and failure modes across all buildings in each cluster, vulnerabilities defined in the New Zealand RC buildings in Section C5 of the New Zealand Seismic Assessment Guidelines [9] were used. According to C5, five major categories for seismic vulnerability of reinforced concrete buildings in New Zealand are categorized as: (1) buildings prior to 1970s, (2) non-ductile columns from 1982 to 1995, (3) precast floors after 1980, (4) shear walls in 1970s and 80s, and (5) buildings with plan/vertical irregularity. Note that the importance of building floor system in seismic vulnerability of buildings is emphasized by these identified archetypes which was also considered here as one of the effective attributes. Table 2 links the dominant attributes of the clusters developed using the k-means on latent space approach to the C5 vulnerabilities. From this table, it is clear the selected indicator buildings have effective attributes, *i.e.*, year, LS, and FS, which matches the applicable seismic vulnerability of the corresponding cluster. Moreover, the C5 vulnerabilities have been distributed well across each cluster, with each cluster containing one or two distinct vulnerabilities (with the exception of plan or vertical irregularities, which were not incorporated into development of the clusters here). It could be claimed that the proposed clustering technique in this paper is able to identify similar seismic response and vulnerability typologies across the entire building database and self-select representative buildings as indicator of each typology.

Table 2. Dominant cluster and representative buildings attributes of Method III and corresponding C5 vulnerability category.

	Dominant range and values				Representative buildings				
Cluster	Year	Height (m)	LS	FS	Year	Height (m)	LS	FS	C5 vulnerability category
0	[1962-1965]	[25-30]	CW (61%)	S&B (70%)	1963	32	CW&CMF	S&B	Prior to 1070s
1	[1926-1930]	[19-24] [29-34]	CW (44%) CMF (30%)	S&B (65%)	1926	23	CW	S&B	Prior to 1970s
					1927	28	CMF	S&B	Prior to 1970s
2	[1985-1988]	[28-34]	CW (50%) CMF (39%)	PH (40%)	1986	34	CMF	PH	Precast floors after 1980
					1986	32	CW	PH	Precast floors after 1980 Non-ductile columns 1982-85
3	[1982-1987]	[76-86]	CW (34%) CMF (30%) CoW (25%)	S&B (52%) PH (30%)	1972	88	CoW	PH	Precast floors after 1980
					1984	92	CMF	S&B	Non-ductile columns 1982-85
4	[1981-1987]	[59-65]	CW (53%) CMF (30%)	PH (36%) S&B (34%)	1981	56	CW	S&B	Shear walls in 1970s and 80s
					1985	64	CMF	PH	Precast floors after 1980

3.5 'Indicator' Building Models

The five top-ranked buildings from each cluster identified in the previous section were selected as candidate indicator buildings for evaluation in a case-study. From the top five ranked buildings in each cluster, the selection of an indicator building was primarily based on the floor plans, structural systems, and building period. Detailed CAD drawings were developed for Ground Floor and Typical

Floors layouts for each building to help in selection of the indicator buildings for detailed modelling. The selected indicator building in each cluster was modelled in CSI ETABS using linear modelling procedures with concrete stiffness modification recommendations in ASCE-41. A demonstration of the indicator building selection procedure for Cluster 0 is shown in Figure 12, where Figure 12a shows the ground floor layouts for the top five buildings, Figure 12b shows the selected indicator building, and Figure 12c shows a rendered sketch of the building model in CSI ETABS. A more detailed overview of this procedure can be found in [10].

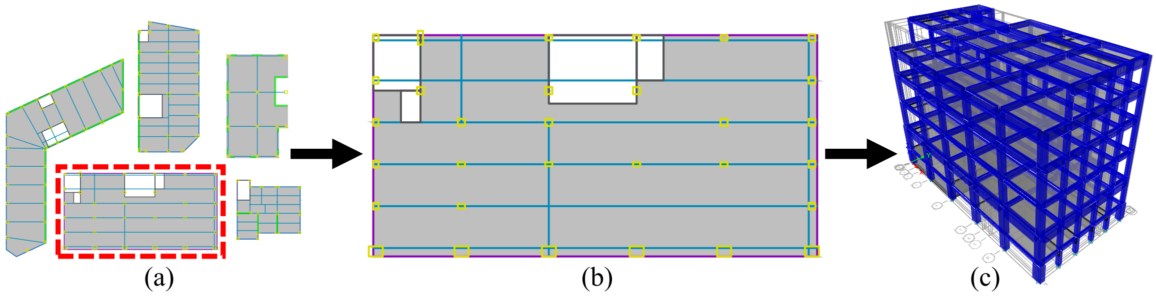


Figure 12. Cluster 0 (a): top five ranked buildings floor plans, (b): the selected indicator building floor plan, and (c): the rendered sketch of ETABS model for indicator building.

To provide additional models to reflect the period ranges in the buildings across each cluster (period ranges for each cluster shown in Figure 13), supplementary models were generated by modifying the stiffness of components in the base indicator building models. This was done using statistical analysis on cluster period distributions and adjusting the stiffness modification coefficients in the base indicator building models. Estimated building periods for all buildings in each cluster were calculated using approximate methods from NZS 1170.5 Commentary [4]. A total of three supplementary models were developed for each cluster based on the mean and standard deviation of building periods in the clusters, resulting in a total of four models for each cluster. The stiffness modification coefficients on the primary structural elements in the models were modified to target building periods equal to the mean period of all buildings within the cluster as well as the mean plus one standard deviation and the mean minus one standard deviation. These models were created to generate response estimates for a range of periods in each cluster for more accurate response estimation of all building in each cluster. Figure 14 shows the fundamental periods of all models developed for each cluster, where ULS represents the base indicator building model.

3.6 Instrumentation of Indicator Buildings

In aligned work funded by QuakeCoRE (agreement number ADMIN-2021-SF-15) and conducted in collaboration with Canterbury Seismic Instruments, 15 buildings in Wellington were temporarily instrumented (including the selected indicator buildings) to determine the actual fundamental periods for model validation and comparison to periods calculated using the simplified approach in 1170.5. The buildings were instrumented using a single accelerograph with a 120dB resolution and a 10ug RMS noise floor, and ambient vibration measurements were taken on a windy day to capture the excited state of the structures. The measured and estimated periods as calculated using 1170.5 are summarised in Table 3. The measured periods were unexpectedly low, suggesting the measurement time was too short or additional instruments were required – this is the focus of ongoing work as instrumenting the selected indicator buildings is essential to the development of the framework presented here. Due to the errors in the measured building periods, they have not been used for model validation in the following section.

Table 3. Estimated and Measured Periods for 15 Buildings in Wellington

Building Name	1170.5 Period - X	1170.5 Period - Y	Measured Period - X	Measured Period - Y	% Difference X	% Difference Y
Plumbers Building	0.74	0.92	0.49	0.51	-34	-45
The Wakefield	0.93	1.16	0.40	0.32	-57	-72
Hope Gibbons Building	0.63	0.79	0.26	0.74	-60	-7
2-8 Maginnity Street	0.53	0.66	0.57	0.58	9	-12
Colenso House	0.88	1.10	0.78	0.91	-11	-18
EMC Building	0.84	1.05	1.33	0.55	59	-48
Tower Building	1.14	1.42	0.59	0.51	-49	-65
Council Administration Building	1.09	1.37	0.68	0.98	-38	-29
Grant Thornton Building	1.85	2.31	1.75	1.84	-5	-21
49 Boulcott Street	1.88	2.35	1.44	0.89	-23	-62
Todd Building	1.12	1.41	1.22	1.74	8	24
17 Whitmore Street	0.70	0.88	0.41	0.58	-42	-34
38 Waring Taylor Street	1.15	1.44	0.49	0.59	-58	-59
St Helens Apartments	0.58	0.72	0.28	0.68	-52	-6
VUW School of Architecture and Design	0.31	0.39	0.32	0.25	0	-36
Endeavour Apartments	0.59	0.74	0.25	0.28	-58	-63

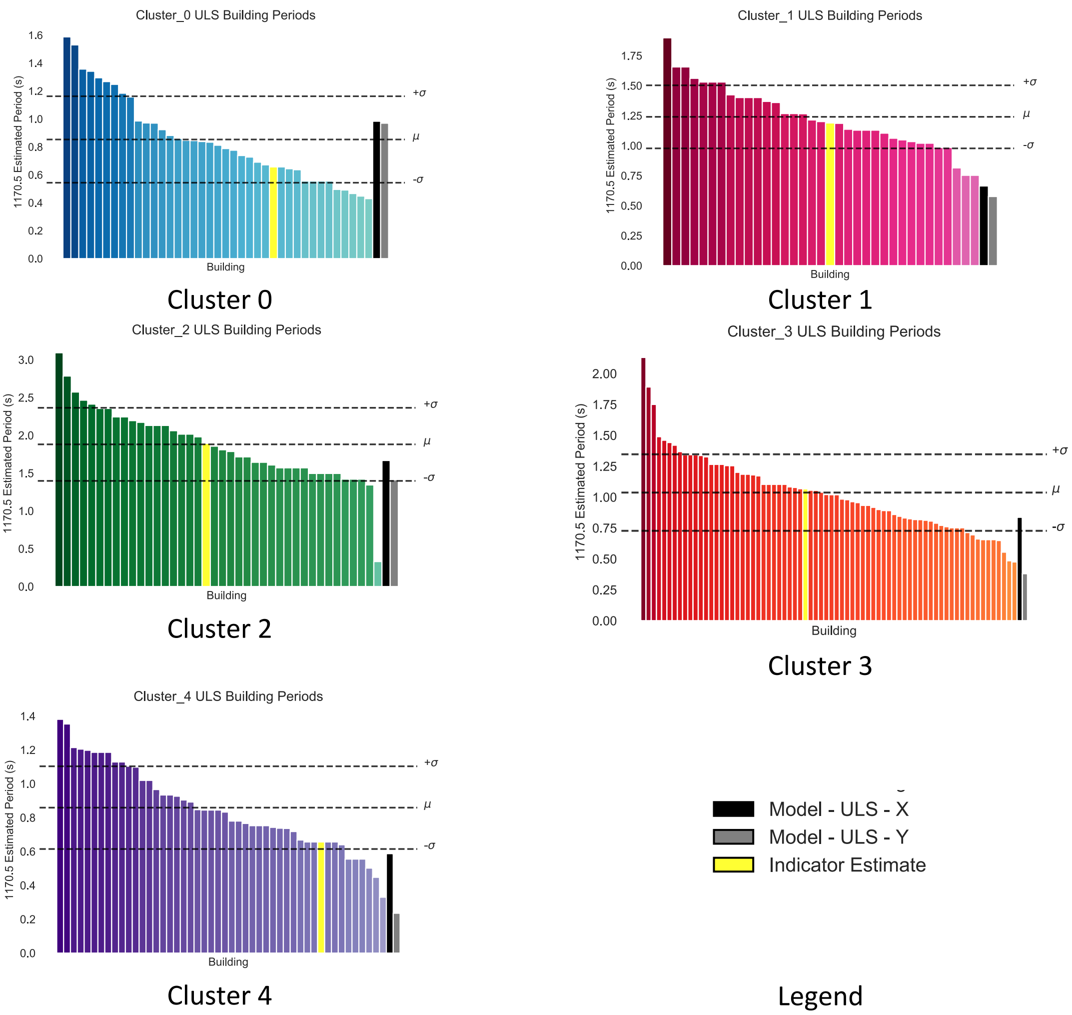


Figure 13. Building period distribution in all clusters

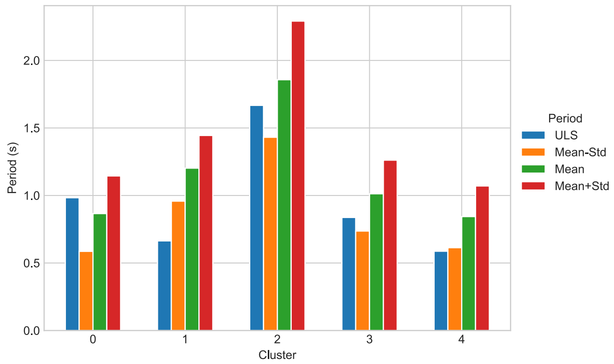
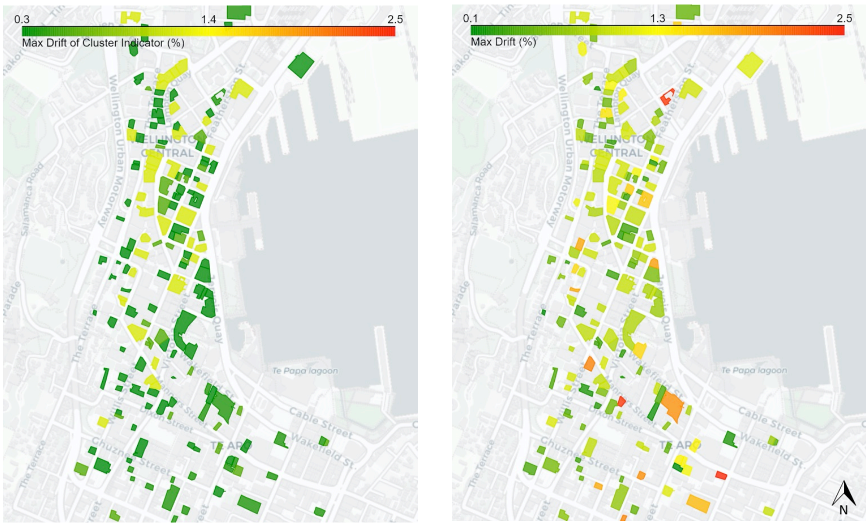


Figure 14. Period distribution of models developed for each cluster

3.7 Regional Response Analysis using 'Indicator' Buildings - A Case Study

A case study was conducted using the 2016 Kaikōura Earthquake, which was an Mw 7.8 earthquake that had an epicenter approximately 200 km from Wellington [11]. The nearest location of fault rupture occurred approximately 50 km South of Wellington and caused significant structural damage to many buildings [11]. Time series acceleration data for this earthquake was obtained using the linked several strong motion stations (SMSs) in Wellington [12]. A linear time history analysis was conducted for the four models in each cluster, where the time history data used for each building was assigned based on site period rather than geographic distance to the SMS as previously stated. Using the maximum drift and acceleration of four different models in each cluster, two methods were utilized to estimate the response parameters of all buildings within the cluster. In Method I, the maximum drift from the indicator building in each cluster was applied to all buildings in that cluster. In Method II, a linear regression model was developed based on the response parameters and periods of the numerical results from the four models in each cluster. The developed model was used to estimate the maximum drift and maximum acceleration of all buildings in each cluster according to building periods (e.g. the response was corrected based on period). Figure 15 shows the estimated drifts for Method I (a), and Method II (b) for the reduced building set in the WBI for the Kaikōura earthquake case study. This comparison demonstrates that Method I tends to result in an underprediction of drift, while a preliminary analysis of Method II indicates that these drift estimates are consistent with response and damage observed during the earthquake. However additional work is necessary to validate this approach.



(a) (b)

Figure 15. Drift maps using two methods; (a) Method One: drifts driven from indicator buildings (b) Method Two: drifts driven from linear regression response model.

The results of the case study demonstrated several important considerations regarding the influence of structural and demand parameters when using indicator buildings to represent the response of a wider cluster of buildings. In particular, the range of periods across each cluster and the shape of the response spectrum of the input motion can have a large impact on the estimated response and must be accounted for. This is clearly demonstrated by comparing the responses of Clusters 2 and 0 for the Kaikōura case study. Figure 16 shows the estimated storey drifts, storey accelerations, and response spectra of the input motions used in each cluster. From this figure, it is clear that period correction is not necessary when the input response spectrum is relatively constant across the period range of the cluster (Cluster 0) but is required when the input response spectrum varies significantly across the period range of the cluster (Cluster 2). Ongoing work is focused on developing robust methods to ensure spectral characteristics of the input motion are accounted for when using indicator buildings to evaluate the response of a wider range of structures.

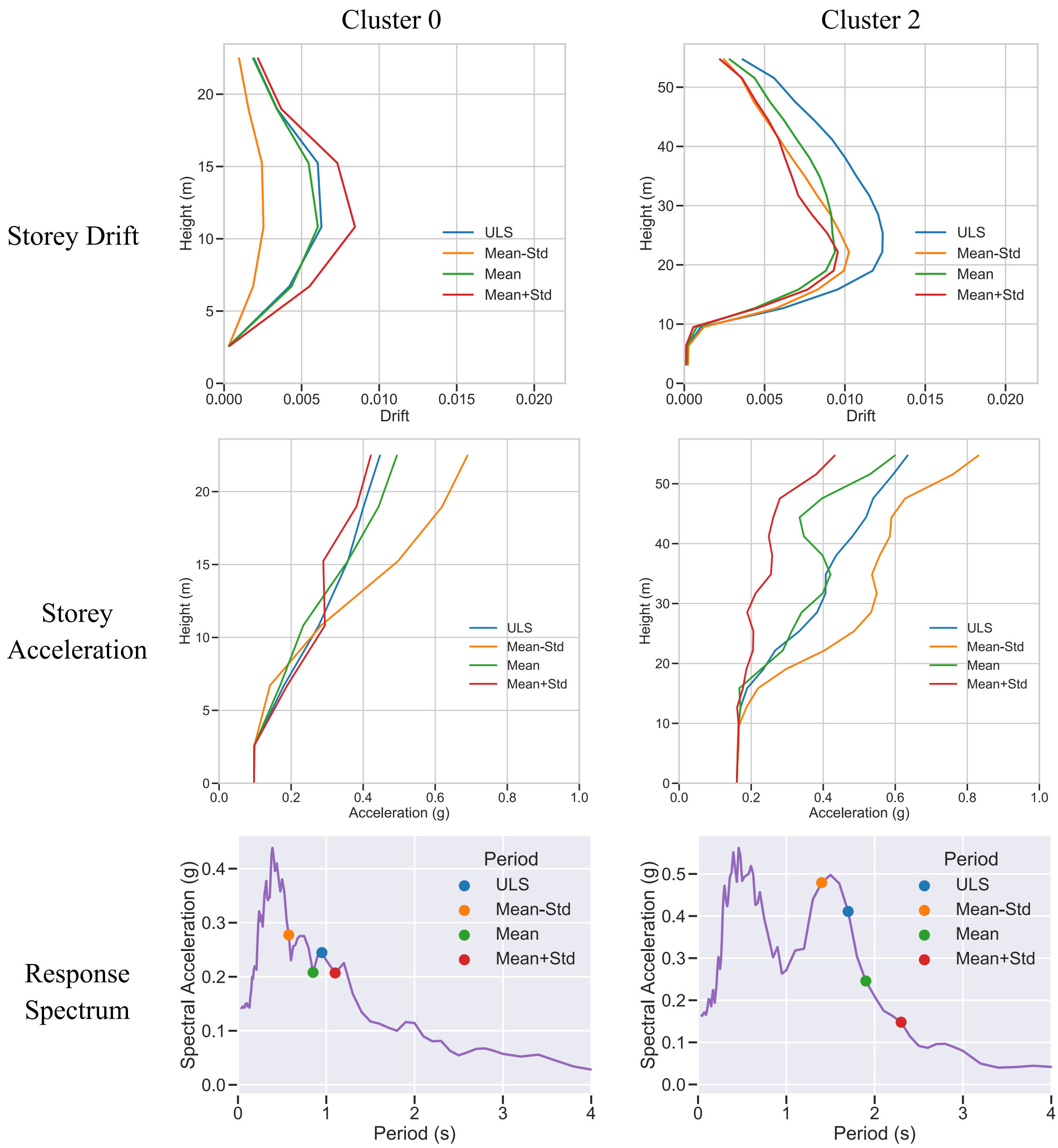


Figure 16. Comparison of Cluster 0 and Cluster 2 modelling results for case study

4. Evaluation of Various Nonlinear Modelling Approaches for Regional Response Analysis

This section presents a study that evaluates the effectiveness of a range of nonlinear numerical analysis methods in capturing the nonlinear response of structures. In the context of this study, this task was undertaken to evaluate the complexity of model required to generate training data for machine learning surrogate models, however the results are useful even in a context where traditional analysis methods are to be used for a large number of structures. Here, six common nonlinear analysis methods for concrete buildings were used to model the response of three instrumented buildings for which damaging earthquake records were available. The analysis methods are summarised in Table 4 and range from phenomenological (highly simplified) cases to more physical macro modelling approaches. The focus of this study was concrete moment frame buildings, and included a full-scale four-storey building tested on the E-defence shake table in Japan [13], a seven-storey building located in Van Nuys CA [14], and a 6-storey building located in Wellington, NZ [15]. An overview of these buildings and the instrumentation locations are shown in Figure 17.

Table 4. Analysis methods

Phenomenological	• Level 1: Equivalent SDOF model
	• Level 2: Simplified MDOF model
	• Level 3: 2D frame model with moment-rotation hinges
	• Level 4: 2D frame model with zero length fibre hinges
	• Level 5: 2D frame model with distributed plasticity fibre elements
	• Level 6: 3D model with moment rotation hinges
	• Level 7: 3D model with zero length fibre element hinges
Physical	• Level 8: 3D model with distributed plasticity fibre elements

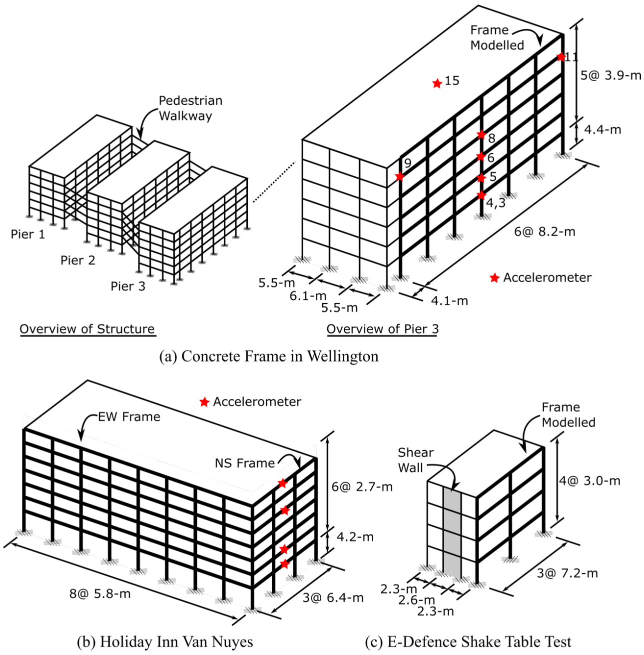


Figure 17. Instrumented study buildings

4.1 Overview of Modelling Approaches

This section provides a brief description of the analysis methods summarised in Table 4. All models were developed in the opensource structural analysis software OpenSeesPy. Detailed descriptions of the models can be found in Scaria and Stephens [16].

Equivalent Single Degree of Freedom Method

The equivalent single degree of freedom (SDOF) modelling approach simplifies a structural frame into a point mass, spring and damper system. This formulation seeks to simplify a frame building into an equivalent SDOF system and analyse the response-history of the structure for the applied ground motion. This method extends the nonlinear static capacity spectrum method proposed by Freeman [17] to enable nonlinear dynamic analyses using an equivalent conversion procedure such as that proposed by Kuramoto et al. [18]. Effectively, this method utilises the nonlinear monotonic pushover characteristic of a Multi-Degree of Freedom (MDOF) system to define the cyclic force-deformation response of the equivalent mass corresponding to the first mode of vibration of the structure.

Simplified Multi Degree of Freedom Method

Commonly known as the ‘fishbone’ model, the equivalent multi-degree of freedom system is a simplified two-dimensional representation of a building frame whereby a single column element represents all the participating columns of a storey, and two half-span beams extending either side of the column at every storey is a condensed representation of the beams at each floor level. Sliding supports are assigned at the ends of the half-beams Khaloo & Khosravi [19]. The two major simplifying assumption in this analysis are equal joint rotations at beam-column connections at any given floor, and negligible axial elongation of columns.

2D and 3D Modelling Using a Lumped Moment-Rotation Hinge Approach

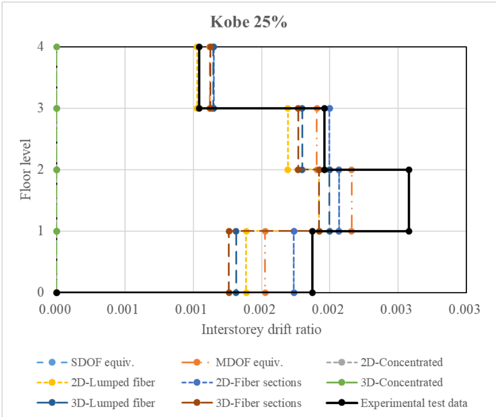
In a single-component model of concentrated plasticity, two inelastic springs are assigned at the ends of frame elements [20]. The overall flexibility of an element using this model is the sum of the flexibilities of the inelastic springs and the elastic beam element. On the other hand, a two-component model has a nonlinear element in parallel with an elastic element, from which the individual stiffness contributions are directly additive. The inelastic element develops plastic hinge only once the yield moment is exceeded at the hinge location [21]. Here, moment rotation hinges were defined based on beam and column cross sections using hinge definitions in ASCE 41.

2D and 3D Modelling Using a Lumped and Distributed Fibre Based Approach

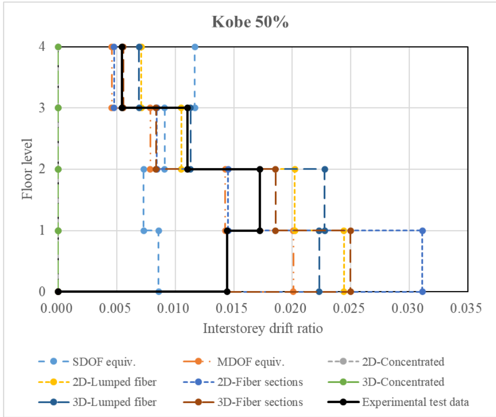
When using a fibre based modelling approach, both finite-hinge-zone model and distributed flexibility models were evaluated. In the finite-hinge-zone model, the length of the plastic hinge is predetermined, and the flexibility of the member is assumed constant between the end hinges. In contrast, for a variable flexibility model, the relation between member end-moments and rotations are functions of the assumed curvature distribution along the length of the member [22]. Lumped fibre-hinges can be assigned at member ends, with a linear beam-column element between the hinges to develop a finite-hinge-zone model. On the other hand, variable flexibility distributed plasticity models can be developed using nonlinear beam-column elements based on a force or displacement-based formulation, which are assigned fibre sections at discrete Gaussian integration points. The total response of the component is then analysed by integrating sectional deformations over the length of the element.

4.2 Comparison of Recorded and Numerical Data

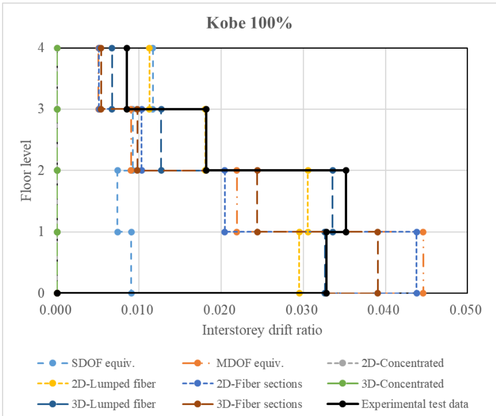
The response of the study structures recorded during damaging earthquakes was compared to numerical results generated using each of the modelling approaches described above. Only the E-Defence results are presented here for brevity, however the results from all buildings followed similar trends and can be found in Scaria and Stephens [16]. The E-Defence structure was subjected to motions recorded during the 1995 Kobe earthquake scaled to varying intensities [13]. Figure 18 shows the maximum recorded and numerical drifts for input motions scaled 25%, 50%, and 100% of the earthquake motion, while Table 5 summarises the maximum storey drifts and percent errors for the 100% scaled motion. InTable 5, the models are ranked by their representative error margin to identify their reliability in accurately predicting drift, which is a common metric to determine expected structural damage. From these results, it is clear significant care must be taken when developing structural models for the purpose of developing training data for surrogate models as relatively significant errors were observed even for the most detailed fibre-based models. Additional work is required on this topic to better quantify the uncertainties in modelling assumptions, constitutive models, etc. associated with these errors to provide more confidence in the baseline data used for the purpose of surrogate model training. Further, structural instrumentation should be used to constrain and validate any models that are to be used for developing training data.



(a)



(b)



(c)

Figure 18. Maximum storey drifts recorded for the E-Defence building for the Kobe earthquake scaled to (a) 25%, (b) 50%, and (c) 100%

Table 5. Ranking of multifidelity models based on Kobe 100% ground motion input

Parametric level	Description	Maximum interstorey drift ratio				% Error				Average (% Error)	Ranking
		L1	L2	L3	L4	L1	L2	L3	L4		
1	SDOF equiv.	0.01300	0.01367	0.01000	0.00333	60	61	45	61	56.9	**6**
2	MDOF equiv.	0.01484	0.02313	0.02367	0.01079	55	34	30	27	36.5	**4**
3	2D-Concentrated	0.01973	0.02140	0.01301	0.00513	40	39	28	40	36.8	**5**
4	2D-Fiber sections	0.02123	0.02734	0.01325	0.00700	35	22	27	18	25.7	**2**
5	3D-Concentrated	0.01920	0.02085	0.01404	0.00618	41	41	23	27	33.1	**3**
6	3D-Fiber sections	0.02052	0.02875	0.01478	0.00835	37	18	19	2	19.1	**1**
Observed	**Experimental data**	**0.03282**	**0.03521**	**0.01819**	**0.00851**					-	-

5. Machine Learning Surrogate Models for Structural Response Modelling

One of the original concepts of this work was the development of machine learning surrogate models to replace conventional finite element analyses to simulate the nonlinear response history of buildings at a community scale. Through compilation of the data required to complete this task, several practicality issues emerged, the largest being the ability to rapidly generate finite element models and training data for a large number of buildings. Although tools exist to generate simplified numerical models which can give a good approximation of structural response, the errors observed in these methods are too large to justify their use for training surrogate models (as illustrated in Section 4 of this report). However, the 'indicator' building approach demonstrated in Section 3 of this report has shown significant promise in approximating city-scale structural response. Therefore it is recommended that detailed models of the indicator buildings, coupled with instrumentation and recorded data from these buildings, could be used to generate training data to develop machine learning surrogate models for those buildings. Those models could then be used to rapidly evaluate regional structural response both for scenario planning and near real-time information following events. The following provides a brief review of promising research focused on the development machine learning surrogate models for estimating structural response and damage. In follow-on research to this work, these methods will be evaluated using data generated from detailed models and instrumentation data from the selected indicator buildings.

This topic is focused on developing machine learning models that can emulate a structural system and predict the seismic response. Work in this area can be segregated into two groups based on the data used to generate the models and the purpose of the models: (1) machine learning methods to identify component failure modes, capacities, and constitutive behaviours (e.g. [23], [24], [25], [26]) and (2) machine learning methods to that use structural parameters and ground motion inputs to predict global structural response. As this work is focused on predicting the seismic response of buildings, methods focused on the later are most relevant here. Deep learning approaches (and in particular artificial neural nets, ANNs) have been shown to be effective in capturing linear and slightly nonlinear structural response of buildings subjected to earthquakes (e.g. [27], [28], [29], [30], [31]). However more recent work completed in the last few years has shown promise in capturing the response of buildings for large earthquake demands and high levels of inelasticity. Zhang et al. [32] introduced a deep learning approach known as a long short-term memory (LSTM) network to predict the seismic response of a non-linear hysteretic system, an instrumented 6-storey concrete moment frame building, and synthetic data from a 3-storey steel moment frame building. Results demonstrated that the LSTM network could accurately capture the nonlinear response of the building. In follow-on work, Zhang et al. [33] developed a modified deep learning approach known as a physics guided convolutional network (PhyCNN) that utilises the laws of physics to provide constraints to the network outputs, alleviate overfitting, and reduce the need for large training data sets. This model architecture was shown to accurately capture the nonlinear response of the 6-storey concrete moment frame building used in previous work. Both of these methods have shown very promising results in accurately capturing nonlinear structure response. The model architectures are freely available on GitHub and will be evaluated for the selected indicator buildings when detailed numerical and experimental data is available for those buildings. This is a topic of ongoing work.

6. Conclusions and Key Findings

This project developed the building blocks for a next generation tool for use in performing regional resiliency and scenario assessments as well as rapid performance and damage estimates following seismic events in Wellington. Key to this work was compiling and linking data sets that can be used to estimate the seismic response of a large number of buildings. These datasets include relevant structural, site, and demand information. The outputs of this project include (1) an enhanced Wellington building inventory with additional information required for seismic response estimates, (2) a rapid framework to compare measured ground motions and design level accelerations in buildings across the Wellington region, (3) a novel machine learning driven clustering methodology for clustering buildings into typologically similar groups and selecting representative indicator buildings, (4) a demonstration of the use of indicator for regional seismic response evaluation, and (5) a study on the accuracy of different nonlinear modelling techniques, and (6) a pathway and recommendations towards developing deep learning surrogate models of selected indicator buildings. Key findings from the work are broadly broken into these categories following a brief comment on priorities for future instrumentation plans in Wellington.

Free-field Strong Motion and Building Instrumentation in Wellington

- Due to the varying site conditions in Wellington, free field strong motion stations need to be strategically placed to supply accurate demand information for seismic response estimation of structures. The current placement of the Geonet strong motion stations in Wellington are not adequate for this purpose, as 66% of buildings within central Wellington are currently tied to a single strong motion station based on site period. The locations of the existing strong motion stations should be revisited, or preferably additional strong motion stations should be added to central Wellington. The locations of any additional strong motion stations in Wellington should be selected based on site period to maximise accurate seismic demand coverage of the building stock.
- Building instrumentation is essential to validate and update numerical models for seismic response and damage estimation. Buildings identified as indicator buildings using the methodology described in Section 3 should be targeted for permanent instrumentation for this purpose. Aligned work funded through QuakeCoRE (agreement number ADMIN-2021-SF-15) which attempted to temporarily instrument many of the identified indicator buildings in collaboration with Canterbury Seismic Instruments demonstrated that this will be a very difficult task, as most building owners are reluctant to agree to instrumentation. To overcome this obstacle, researchers and policy makers need to develop a strategy to encourage wider building instrumentation in Wellington.

Wellington Building Inventory

- The Wellington Building Inventory is a useful tool that combines information from disparate building databases to improve seismic resiliency by informing strategic retrofit prioritization through the identification of critical structural deficiencies which can lead to building failures, and quantifying the downstream economic and social impacts of these failures. The curation and continued development of this database is ongoing, and needs continued support from researchers and stakeholders.

Rapid One-DOF Framework

- The Rapid One-DOF framework requires limited structural information and can be used to quickly identify buildings where the earthquake demands exceeded the design demands following seismic events.
- Depending on the spectral shape of the applied ground motion, the estimated demand in this framework can be very sensitive to the building period. The fundamental period estimates for the buildings in this framework have been calculated using the approximate approach outlined in NZS 1170.5.

Building Clustering and Indicator Buildings

- Machine learning driven clustering of buildings into typologically similar groups for indicator building selection and seismic response and damage assessment shows promising results.
- The k-means on latent space clustering approach is more effective at clustering numerical attributes than the k-prototype or k-means on FAMD methods. The latent space approach resulted in narrow year, height, and period bands within each cluster, while wide distributions of these attributes were observed when using the other methods.
- The k-means on latent space method is more effective at clustering categorical data, however all approaches had difficulty due to unbalanced data distribution for the categorical label in the Wellington dataset which has also resulted in more than one dominant attribute in each cluster.
- When selecting representative buildings, the k-means on latent space method selected buildings that were both more representative of the dominant typologies in the clusters and provided a relatively good distribution of attributes across the entire dataset.
- In cases where there are multiple dominant attributes within a cluster that can influence response and/or damage, multiple indicator buildings may be necessary within a single cluster. This is the focus of ongoing work.
- Common vulnerabilities in concrete buildings in New Zealand were well distributed across the clusters developed using the k-means on latent space method, with a maximum of two vulnerabilities across any given cluster.
- The indicator building approach demonstrated promising results in estimating the seismic response of buildings within typologically similar groups when using the Kaikōura earthquake as a case-study.
- When using indicator buildings, care must be taken to correct the estimated building responses based on the period ranges of buildings within each group to account for differences in spectral shape of the input motion. This is the focus of ongoing work.
- It would be beneficial if the selected cluster typologies and indicator buildings were evaluated by experienced practicing engineers in the region of interest (Wellington in this case). Several attempts were made to schedule a workshop for this purpose as part of this work, however the work shop was postponed and eventually cancelled due to COVID-19 restrictions.

Development of Training Data for Machine Learning Surrogate Models and Deep Learning for Structural Response Modelling

- Significant care must be taken when developing building response training data for machine learning surrogate models using traditional finite element models. Even detailed fibre-based

macro-models demonstrated relatively significant errors in estimated drifts and accelerations in this study. It is therefore recommended that building instrumentation data be used in conjunction with traditional finite element approaches when developing training data for surrogate models. Indicator buildings selected according to the clustering techniques developed here should be targeted for building instrumentation and detailed modeling for the purpose of training surrogate models.
- Several promising deep learning approaches for nonlinear response modelling of structures have been proposed in the last few years including LSTM and PhyCNN networks ([32], [33]). These methodologies should be evaluated for machine learning surrogate modelling of the selected indicator buildings.

7. Future Work

- Continued development of the Wellington Building Inventory as additional structural, site, and demand information is made available. Detailed drawings from an additional 200 buildings were recently obtained from WCC. Data from these drawings will be extracted to expand the number of buildings for which detailed structural information is available.
- Obtaining more granular seismic demand information in the Wellington CBD through partnerships with commercial and government entities that are providing seismic instrumentation. Discussions on this topic are ongoing.
- New PhD student Alex Kirby (funded by a University of Auckland Doctoral Scholarship) has been selected to continue the development of the regional response framework discussed here. His work will continue the evaluation of machine learning driven techniques for building clustering and indicator building selection, and will also focus on developing accurate data for surrogate model training through detailed modelling and building instrumentation.

8. Impact

This project developed the building blocks for a next generation tool for use in performing regional resiliency and scenario assessments as well as rapid performance estimations following seismic events in New Zealand. While pieces of the framework were developed and demonstrated for Wellington, the work-flow itself is fundamental in nature, and is therefore scalable for use in other communities and with other infrastructure systems (including bridges, dams and energy/power infrastructure). As the innovative engineering approach developed here is further refined, it will have significant impact in New Zealand in terms of planning resilient cities and infrastructure. Central to the work developed here is the use of machine learning driven computational methods to efficiently cluster buildings into typologically similar groups and select representative indicator structures that can be used to estimate the response of structures at a regional scale. Ongoing work is focused on detailed modelling and instrumentation of the selected indicator buildings to generate a seismic resiliency and evaluation quantification tool that offers granularity regarding structural response and damage at the individual structure level across a region.

9. Acknowledgements

- PhD Students – Amin Ghasemi
- QuakeCoRE Flagship 3 – Development of baseline Wellington Building Inventory

- QuakeCoRE Agreement Number 3723086 – Data Compilation Towards Optimisation of Strong Motion Sensor Networks
- QuakeCoRE Agreement Number ADMIN-2021-SF-15 – Instrumentation of Indicator Buildings in Wellington
- Resilience to Natures Challenges 2 (RNC 2) – funding for Amin Ghasemi
- Hamish Avery and Canterbury Seismic Instruments

10. References

[1] FEMA, *HAZUS Earthquake Model User Manual*. Washington D.C.: Federal Emergency Management Agency, 2011.

[2] A. Puranam *et al.*, “A detailed inventory of medium to high-rise buildings in Wellington’s central business district,” *Bull. N. Z. Soc. Earthq. Eng.*, vol. 52, no. 4, Art. no. 4, Dec. 2019, doi: 10.5459/bnzsee.52.4.172-192.

[3] A. Kaiser, C. V. Houtte, N. Perrin, L. Wotherspoon, and G. McVerry, “Site characterisation of GeoNet stations for the New Zealand Strong Motion Database,” *Bull. N. Z. Soc. Earthq. Eng.*, vol. 50, no. 1, Art. no. 1, Mar. 2017, doi: 10.5459/bnzsee.50.1.39-49.

[4] “NZS 1170.5:2004 :: Standards New Zealand.” https://www.standards.govt.nz/shop/nzs-1170-52004/ (accessed Mar. 07, 2022).

[5] “NZS 1170.5 Supp 1:2004 A1 :: Standards New Zealand.” https://www.standards.govt.nz/shop/nzs-1170-5-supp-12004-a1/ (accessed Mar. 07, 2022).

[6] A. Ghasemi, M. Stephens, and K. Elwood, “Wellington Building Inventory: Rapid Earthquake Response Framework,” presented at the NZSEE Conference, Christchurch, New Zealand, Apr. 2021. Accessed: Mar. 07, 2022. [Online]. Available: https://repo.nzsee.org.nz/xmlui/handle/nzsee/2356

[7] A. Ghasemi, “Wellington Building Inventory: Rapid Earthquake Response Framework,” presented at the 2021 NZSEE Conference, Christchurch, New Zealand, Apr. 15, 2021.

[8] Ghasemi, A. and Stephens, M.T., “Building Clustering for Regional Seismic Response and Damage Analysis,” *Earthq. Spectra*, In Review.

[9] MBIE, “The Seismic Assessment of Existing Buildings- Technical Guidelines for Engineering Assessments. Section C5: Concrete Buildings,” Wellington, New Zealand: Ministry of Business, Innovation, and Employment, 2017.

[10] Petreski, D., Cashmore, J., Ghasemi, A., and Stephens, M.T., “Development of a Regional Seismic Respnse Model Using Building Clustering,” presented at the 12th National Conference on Earthquake Engineering, Salt Lake City, Utah, Jul. 2022.

[11] B. A. Bradley, L. M. Wotherspoon, and A. E. Kaiser, “Ground motion and site effect observations in the wellington region from the 2016 Mw7.8 Kaikōura, New Zealand earthquake,” *Bull. N. Z. Soc. Earthq. Eng.*, vol. 50, no. 2, Art. no. 2, Jun. 2017, doi: 10.5459/bnzsee.50.2.94-105.

[12] C. V. Houtte, S. Bannister, C. Holden, S. Bourguignon, and G. McVerry, “The New Zealand Strong Motion Database,” *Bull. N. Z. Soc. Earthq. Eng.*, vol. 50, no. 1, Art. no. 1, Mar. 2017, doi: 10.5459/bnzsee.50.1.1-20.

[13] Wallace, J., “U.S. Instrumentation and Data Processing for the 4-story RC and Post-Tensioned E-Defense Buildng Tests,” NEES Report NEES-2011-1005. [Online]. Available: https://datacenterhub.org/resources/14266

[14] Y. R. Li and J. O. Jirsa, “Nonlinear Analyses of an Instrumented Structure Damaged in the 1994 Northridge Earthquake,” *Earthq. Spectra*, vol. 14, no. 2, pp. 265–283, May 1998, doi: 10.1193/1.1585999.

[15] R. Chandramohan, L. M. Wotherspoon, B. Bradley, M. Nayyerloo, S. R. Uma, and M. T. Stephens, “Response of instrumented buildings under the 2016 Kaikoura earthquake.,” 2017, Accessed: Mar. 08, 2022. [Online]. Available: https://ir.canterbury.ac.nz/handle/10092/13311

[16] Scaria, A. and Stephens, M.T., “A study on the efficacy of multi-fidelity macro-modelling of RC buildings for nonlinear seismic simulation.,” *Bull. N. Z. Soc. Earthq. Eng.*, In preparation.
[17] Freeman, S. A., “The capacity spectrum method,” presented at the 11th European Conference on Earthquake Engineering.
[18] Kuramoto, H. and Teshigawara, M., “Predicting the response of buildings using equivalent single degree of freedom system,” presented at the 12th World Conference in Earthquake Engineering.
[19] Khaloo, A. R. and Khosravi, H., “Modified Fishbone Model: A Simplified MDOF Model for Simulation of Seismic Responses of Moment Resisting Frames,” *Soil Dyn. Earthq. Eng.*, vol. 55, pp. 195–210, 2013.
[20] Giberson, M.F., “Two Nonlinear Beams with Definitions of Ductility,” *J. Struct. Div.*, vol. 95, pp. 137–157, 1969.
[21] Clough, R., Benuska, K.L., and Wilson, E.L., “Inelastic Earthquake Response of Tall Buildings,” 1965, pp. 68–89.
[22] Kunnath, S. K., “Modeling of Reinforced Concrete Structures for Nonlinear Seismic Simulation,” *J. Struct. Integr. Maint.*, vol. 3, no. 3, pp. 137–149, doi: https://doi.org/10.1080/24705314.2018.1492669.
[23] Mitra, N, Mitra, S., and Lowes, L.N., “Probabilistic model for failure initiation of reinforced concrete interior beam-column connections subjected to seismic loading,” *Eng. Struct.*, vol. 33, no. 1, pp. 154–162, 2011.
[24] Jeon, J-S, Shafieezadeh, A., and DesRoches, R, “Statistical models for shear strength of RC beam column joints using machine-learning techniques,” *Earthq. Eng. Struct. Dyn.*, vol. 43, pp. 2075–2095, 2014.
[25] Yaseen, Z.M., Afan, H.A., and Tran, M.T., “Beam-column joint shear prediction using hybridized deep learning neural network with genetic algorithm,” 2018, vol. 143, pp. 1–7.
[26] Naderpour, H. and Mirrashid, M., “Shear failure capacity prediction of concrete beam–column joints in terms of ANFIS and GMDH,” *Pract. Period. Struct. Des. Constr.*, vol. 24, no. 2, 2019.
[27] Xu, B., Wu, Z., Chen, G., and Yokoyama, K., “Direct identification of structural parameters from dynamic responses with neural networks,” *Eng. Appl. Artif. Intell.*, vol. 17, no. 8, pp. 931–943.
[28] Jeng, C.H. and Mo, Y.L, “Quick seismic response estimation of prestressed concrete bridges using artificial neural networks,” *J. Comput. Civ. Eng.*, vol. 18, no. 4, pp. 360–372.
[29] Joghataie, A. and Farrokh, M., “Dynamic analysis of nonlinear frames by Prandtl neural networks,” *J. Eng. Mech.*, vol. 134, no. 11, pp. 961–969, 2008.
[30] Tsompanakis, Y., Lagaros, N.D., Psarropoulos, P.N., and Georgopoulos, E.C., “Simulating the seismic response of embankments via artificial neural networks,” *Adv. Eng. Softw.*, vol. 40, no. 8, pp. 640–651, 2009.
[31] Akbas, B., Shen, J., and Sabol, T.A., “Estimation of seismic-induced demands on column splices with a neural network model,” *Appl. Soft Comput.*, vol. 11, no. 8, pp. 4820–4829, 2011.
[32] Zhang, R., Chen, Z., Chen, S., Zheng, J., Buyukozturk, O., and Sun, H., “Deep long short-term memory networks for nonlinear structural seismic response prediction,” *Comput. Struct.*, vol. 220, pp. 55–68, 2019.
[33] Zhang, R., Liu, Y., and Sun, H., “Physics-guided Convolutional Neural Network (PhyCNN) for Data-driven Seismic Response Modeling,” *Eng. Struct.*, vol. 215, 2020.

11. Outputs and Dissemination

Presentations:

- Ghasemi, A. “Wellington Building Inventory: Rapid Earthquake Response Framework,” presented at the 2021 NZSEE Conference, Christchurch, New Zealand, Apr. 15, 2021.
- Ghasemi, A. “Building Clustering using Unsupervised Machine Learning Methods,” presented at TechWeek 2021: Smart Resilient Cities Showcase, Wellington, New Zealand, May 28, 2021.

- Stephens, M.T., “Development of a Regional Seismic Response Model Using Building Clustering,” presented at the 12th National Conference on Earthquake Engineering, Salt Lake City, Utah, Jul. 2022.

Conference Papers:

- A. Ghasemi, M. Stephens, and K. Elwood, “Wellington Building Inventory: Rapid Earthquake Response Framework,” Proceedings of the 2021 NZSEE Conference, Christchurch, New Zealand, 2021.
- Petreski, D., Cashmore, J., Ghasemi, A., and Stephens, M.T., “Development of a Regional Seismic Response Model Using Building Clustering,” Proceedings of the 12th National Conference on Earthquake Engineering, Salt Lake City, Utah, Jul. 2022.

Journal Papers:

- Ghasemi, A. and Stephens, M.T., “Building Clustering for Regional Seismic Response and Damage Analysis,” Earthq. Spectra, In Review.
- Scaria, A. and Stephens, M.T., “A study on the efficacy of multi-fidelity macro-modelling of RC buildings for nonlinear seismic simulation.,” Bull. N. Z. Soc. Earthq. Eng., In preparation.