Selection of representative load profiles

Accompanying Appendix Six to

Understanding the value of residential solar PV and storage in New Zealand

29 January 2025

Prepared for the Energy Efficiency and Conservation Authority

Prepared by Dr Allan Miller

www.millercl.co.nz

Contents

Disclaimer

This appendix accompanies the report "Understanding the value of residential solar PV and storage in New Zealand". The information and results are supplied in good faith and reflect the expertise and experience of the author. The model used to derive the results is subject to assumptions and limitations referred to in the document and model specification. Any reliance on the model results is a matter for the recipient's own commercial judgement, taking into account the inputs and assumptions given. AMCL accepts no responsibility for any loss by any person acting or otherwise as a result of reliance on this document and the results.

1 Introduction

A key input to determining residential solar economics, such as rate of return and payback, is a consumer's load profile. This would ideally use individual consumers' load profiles. However, in a generalised study, this is impractical. Instead, a large number of load profiles were collected across the four cities of interest in the study. For simpler parts of the study, each load profile was run through a basic residential solar model to produce distributions of results. For more complex parts of the study, and to to show results by representative loads rather than over many load profiles, a set of loads was chosen to represent consumers.

This appendix deals with the collection of load profiles from the four cities for the study, and the selection of representative load profiles for most of the analysis. The process to achieve this is depicted in Figure 1.

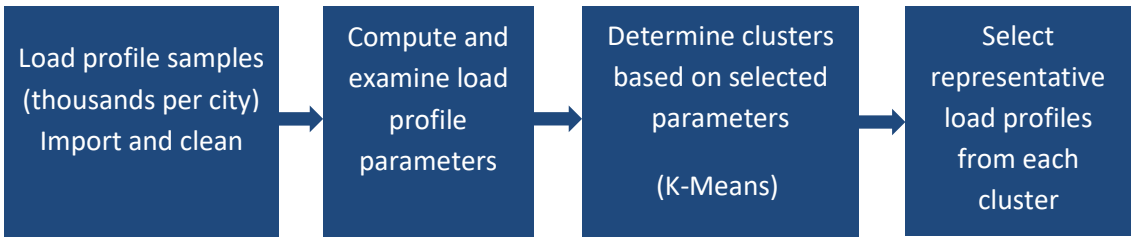


Figure 1: The process of obtaining representative load profiles for the study.

The appendix is divided into six sections which follow the steps in Figure 1. These are:

- Section 2 documents the number of load profiles collected, and measures taken to 'clean' them to make them suitable for the study.
- Section 3 presents parameters of the load profiles collected, ranging from time domain parameters to frequency domain parameters obtained from discrete Fourier transforms of the time-domain load profiles.
- Section 4 discusses methods of clustering the load profiles into representative sets, and the selection of a suitable subset of parameters determined above to group the load profiles into clusters.
- Section 5 discusses the clusters in more detail and/or summarises the clusters used throughout the study.
- Section 6 outlines the method used to select representative load from the clusters.
- Section 7 summarises the representative load profiles used throughout the study.

2 Load profiles collected and 'cleaning' of load profiles

A total of 49,385 load profiles were provided to EECA by Meridian Energy and Powershop. These were anonymised, where the ICP number was replaced with a generic ID from 1 to 49,385. All load profiles were collected from ICPs without solar, although no information about consumers at each ICP was available, other than city and postcode. Further, each ICP's retail price structure and whether they were on a night-rate only or controlled hot-water tariff were unknown.

The fact that the load profiles were from two retailers owned by one company introduces the possibility of bias in the load profile population used in the study as it is possible that consumers with certain load characteristics choose these two retailers. More likely is that these consumers in some way respond to the prices of the two retailers, which may shape the load profiles to some degree.

The number of load profiles collected per city considered in the study is given in Table 1. This table also indicates the steps taken to ensure whole load profiles that contained minimal invalid values, as well as load profiles that are most likely single phase residential consumers. To ensure this, any load profile with a maximum average load over a half-hour period greater than 15 kW was discarded (15 kW corresponds to a single-phase fuse rating of 65 Amps). Load profiles with annual consumption greater than 31,000 kWh per annum were also discarded. This maximum value was chosen to include very large residential consumers, with average load of 3.5 kW, but aimed to exclude businesses. Finally, load profiles with annual consumption below 900 kWh per annum were discarded to attempt to exclude commercial loads such as vending machines and streetside cabinets. This was based on an average load below 100 Watts, which is estimated to be below the average consumption of a domestic fridge.

In all cases each load profile began at the half-hour ending 00:30 (Trading Period 1) on 1 January 2023 and contained 17,520 half-hour values of energy (kWh). The values were contiguous half-hour periods, meaning a consistent time base could be constructed. This also allowed matching with solar generation, which was adjusted to begin at the same time for the typical meteorological year.

Table 1: Treatment of load profiles to produce the final sample set.

City	Number of load profiles obtained	Number of load profiles after removal of profiles with incomplete days or more than 200 invalid values	Number of load profiles after removal of those with maximum demand over 15 kW, annual consumption over 31,000 kWh, or annual consumption below 900 kWh
Auckland	15,481	14,888	14,573
Wellington	10,918	10,671	10,514
Christchurch	21,926	21,475	21,038
Queenstown	1,060	989	920
Total	**49,385**	**48,023**	**47,045**

3 Load profile parameters

Parameters were derived from each load profile as potential candidates for grouping the load profiles into representative groups (clusters), as discussed in Section 4. Broadly two types of parameters were assessed: time domain, outlined in Section 3.1 and frequency domain, outlined in Section 3.2. Section 4 discusses the use of these parameters to form clusters.

While only a subset of all parameters discussed in this section are used to form clusters, most of the parameters are presented as they give characteristics of load profiles and compare them between the four cities in the study, which may be of interest.

3.1 Time-domain parameters

The parameters given in Table 3 were calculated for each load profile as possible candidates for determining representative clusters. These parameters are summarised in the following figures by city. The definitions set out in Table 2 were part of establishing of these parameters.

Table 2: Definitions of time periods used to determine the parameters set out in Table 3.

Period	Definition	Notes
Winter	May to August inclusive except for seasonal ratios, where the seasons have the internationally recognised definition.	Set to align with traditional winter months and to include May, as demand can be higher in May than other non-winter months as days shorten and temperatures fall.
Morning peak	7am-11am	Set to capture morning hot-water heating and to align with peak periods of some retailers and most distributors.
Evening peak	5pm-9pm	Set to capture evening cooking and heating, and to align with peak periods of some retailers and most distributors.
Daytime	7am-11pm	Set to align night-time (that being the hours not included in the daytime definition) with most retailers' and distributors' night-time definition.

*Table 3: Time domain load profile parameters investigated. * Indicates parameters selected for clustering trials discussed in Section 4. ** Indicates parameters trialled and ultimately used to define the clusters used in the study.*

Period	Parameter	Reference and notes
1. Annual	a. Average annual load *	annual_ave_kw, Figure 2
	b. Average annual load factor *	load_factor, Figure 3
2. Weekly	a. Ratio of weekday average load to annual average load	awdl_to_ya Figure 4. Discarded due to little spread.
	b. Ratio of weekend day average load to annual average load	awkndl_to_ya, Figure 5. Discarded due to little spread.
	c. Average weekday load to average weekend load	awdl_to_awkndl (not shown, discarded as results were not consistent in some cities).
3a. Monthly	Ratio of average load in the maximum month to annual load	max_mth_ave_to_ave, Figure 6. Discarded, as seasons were selected to represent seasonality.
	Ratio of average load in the minimum month to annual load	min_mth_ave_to_ave, Figure 7. Discarded, as seasons were selected to represent seasonality.
	Months in which the average demand is higher than any other month	max_mth, Figure 8. Discarded, as seasons were selected to represent seasonality.
	Months in which the average demand is lower than any other month	min_mth, Figure 9. Discarded, as seasons were selected to represent seasonality.
	Ratio of the average load in the maximum to average load in the minimum month	max_to_min_mth (not shown, discarded as results were not consistent in some cities).
3b. Seasonal	Season in which the average demand is higher than any other season *	max_season, derived from Figure 8
	Season in which the average demand is lower than any other season *	min_season, derived from Figure 9
4. Intra day	a. Ratio of average weekday morning peak load to average weekday load over the year **	awdmpl_to_ya, Figure 10
	b. Ratio of average weekday evening peak load to average weekday load over the year **	awdepl_to_ya, Figure 11
	c. Average winter weekday morning peak load to average weekday load over the winter (May to August inclusive)	awwdmpl_to_ya, (not shown, discarded as 4a used instead).
	d. Average winter weekday evening peak load to average weekday load over the winter (May to August inclusive)	awwdepl_to_ya, (not shown, discarded as 4b used instead).
5. Day-night demand	Ratio of average daytime load to all day load**	day_ave_to_all_day_ave, Figure 12.

The following observations were made about the distributions of parameters in Figure 2 to Figure 12:

1. Average load, while showing a good spread in Figure 2 was removed as a parameter because clusters are selected based on average load (8,000 kWh pa and 12,000 kwh pa are chosen as the annual consumption and therefore effectively an average load for the study).
2. Annual load factor also shows a good spread, but the intra-day and day-night ratios in some way represent load factor, and there is a good correlation between annual average load and load factor, as discussed in Section 4.
3. Weekly ratios show little spread.
4. Monthly and seasonal ratios show promise but were discarded as seasonality is selected for when selecting representative load profiles, as discussed in alter sections.

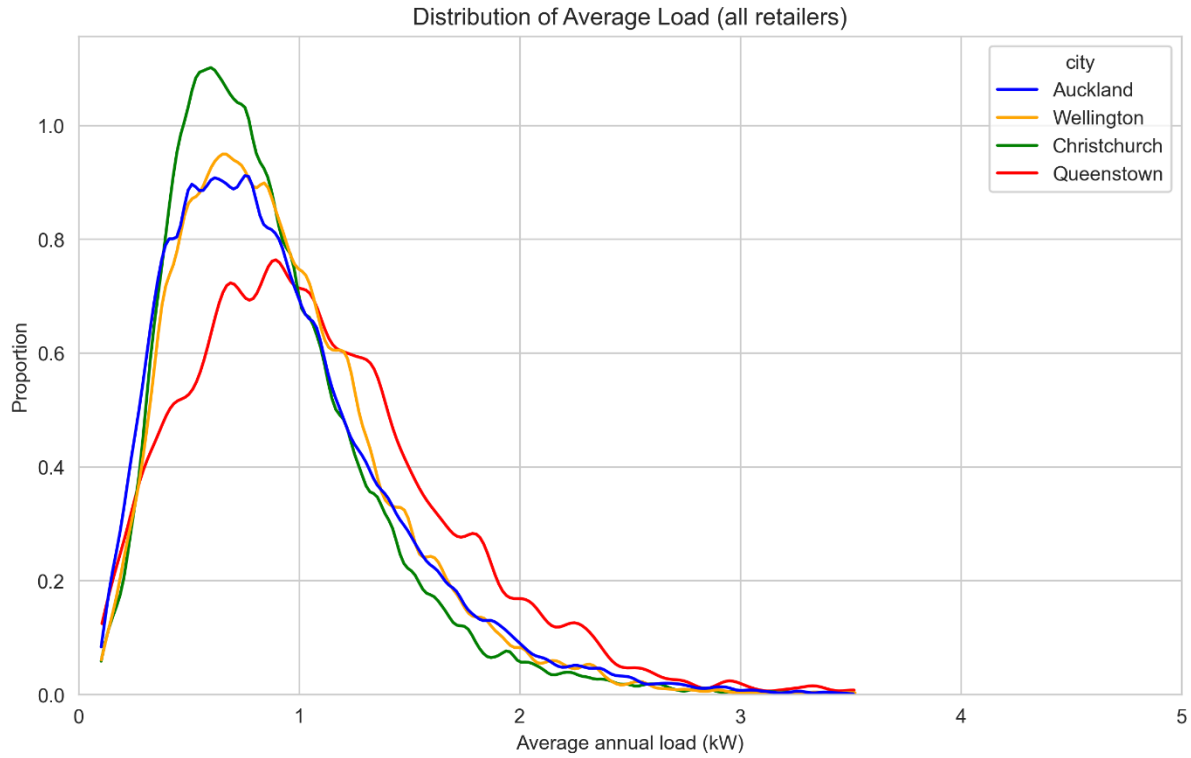


Figure 2: Characteristics of average annual load by city.

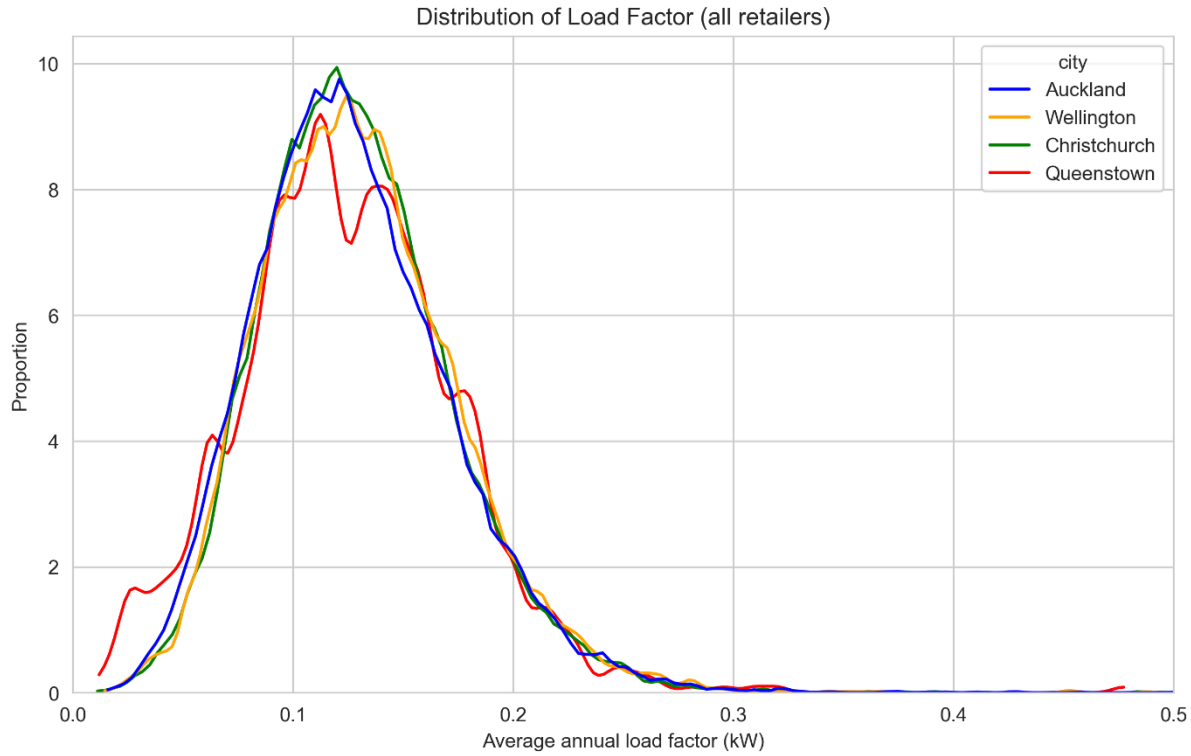


Figure 3: Characteristics of average annual load factor by city.

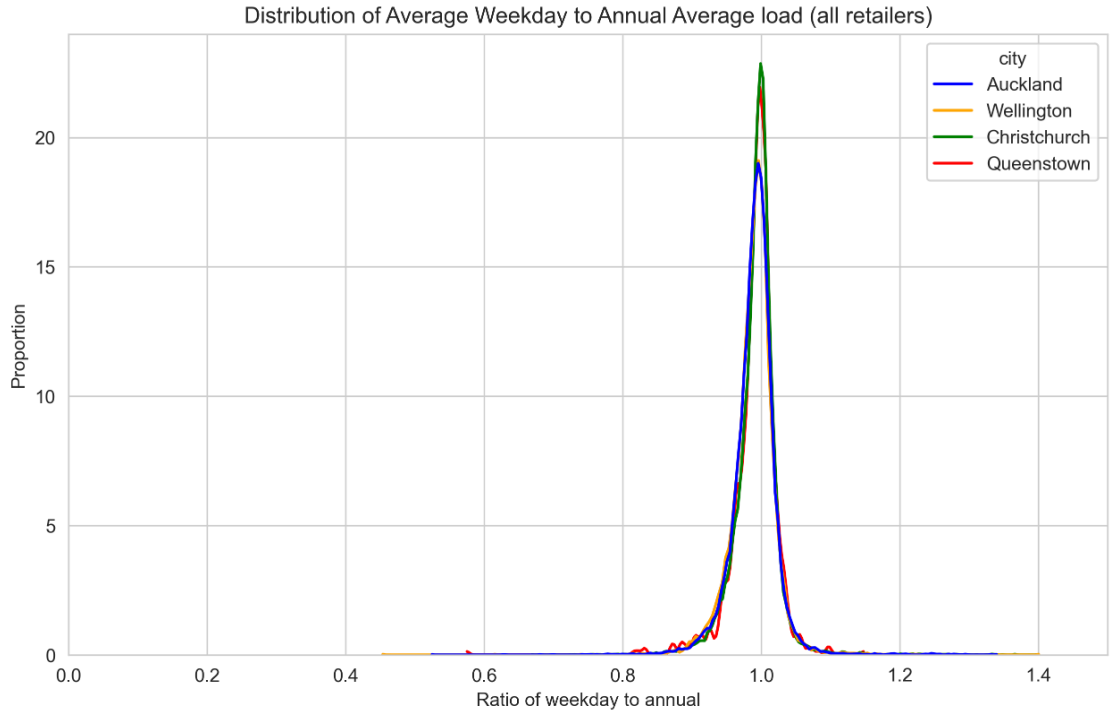


Figure 4: Characteristics of the ratio of weekday to annual average load by city. Discarded as a clustering parameter due to very little variance.

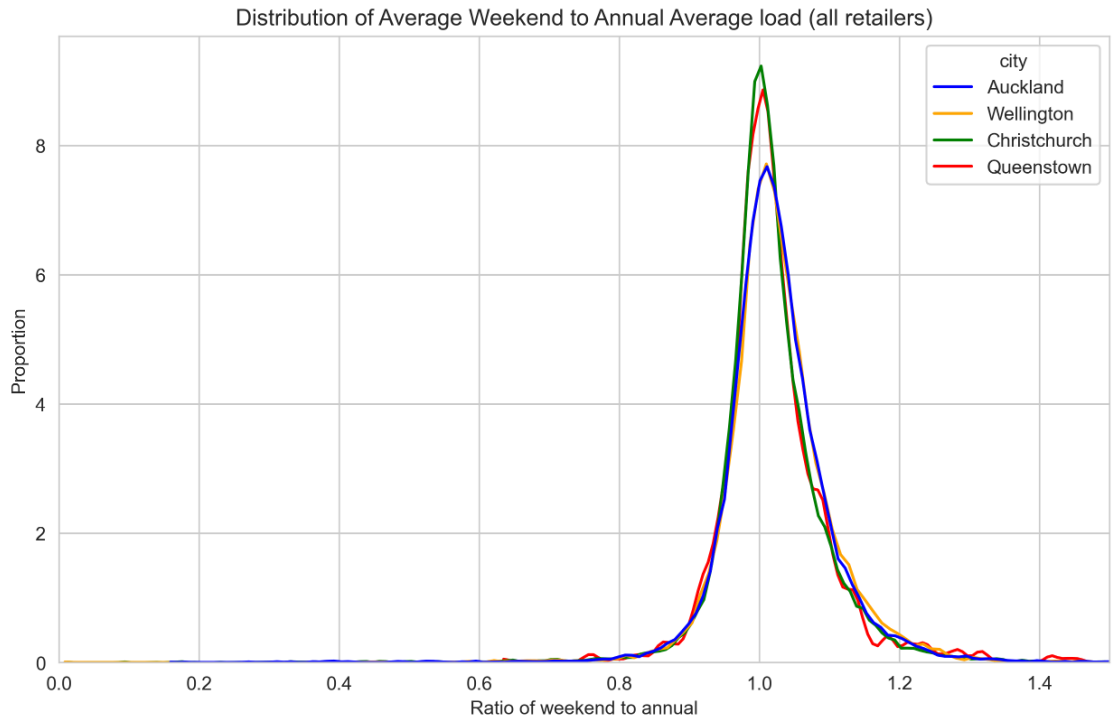


Figure 5: Characteristics of the ratio of weekend day to annual average load by city. Discarded as a clustering parameter due to very little variance.

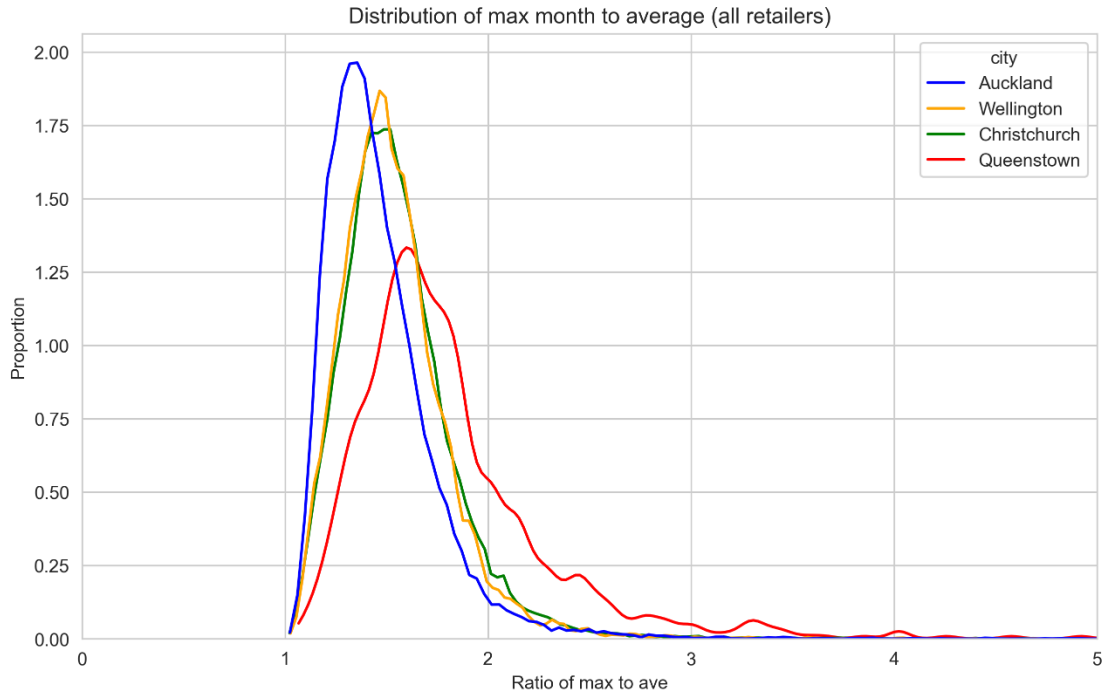


Figure 6: Characteristics of the ratio of the average load in the month of maximum consumption to the annual average load by city.

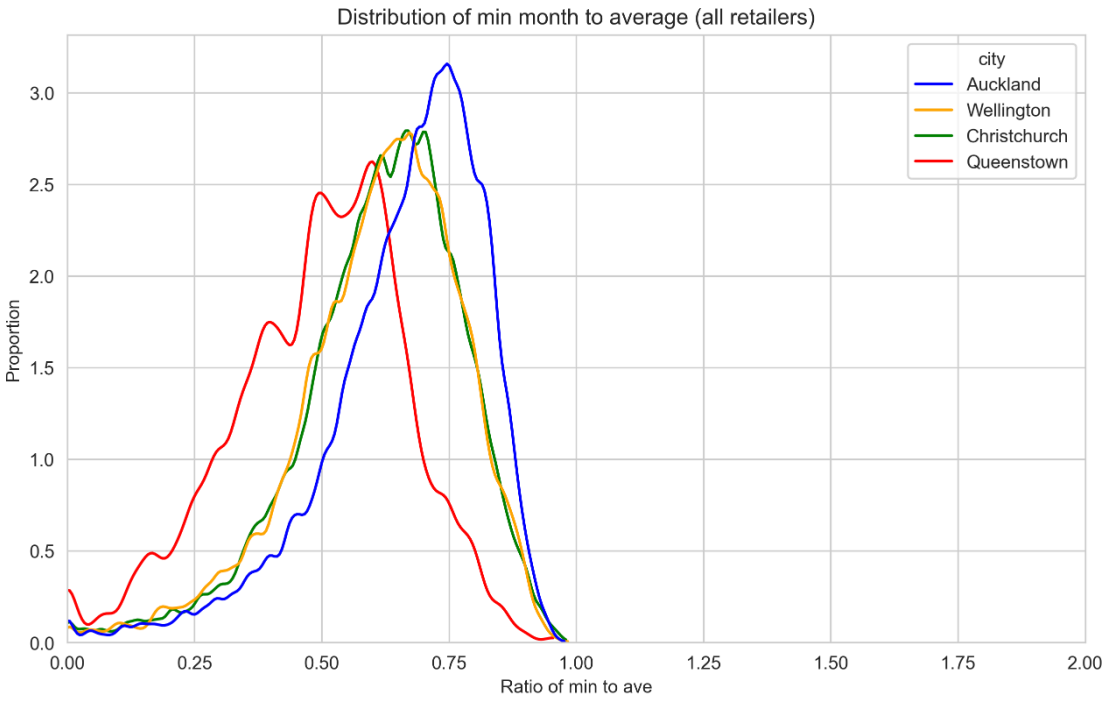


Figure 7: Characteristics of the ratio of the average load in the month of minimum consumption to the annual average load by city.

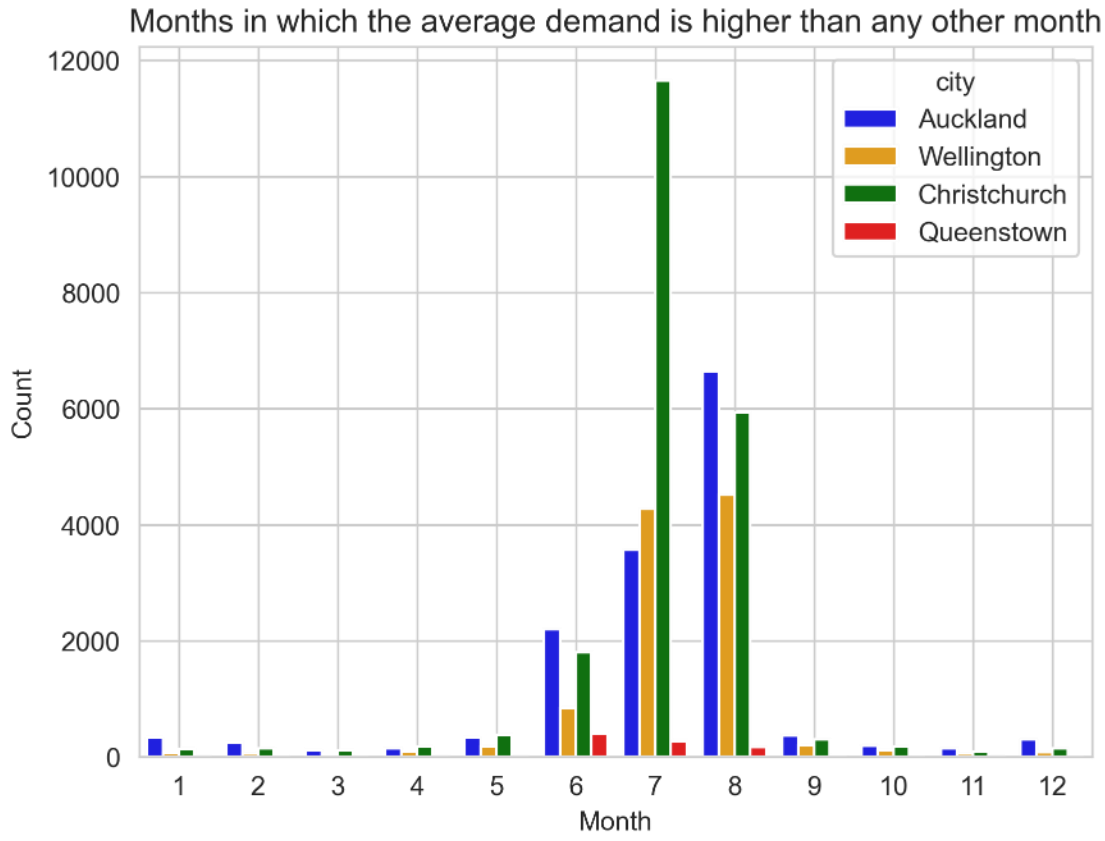


Figure 8: Months in which maximum average load occurred.

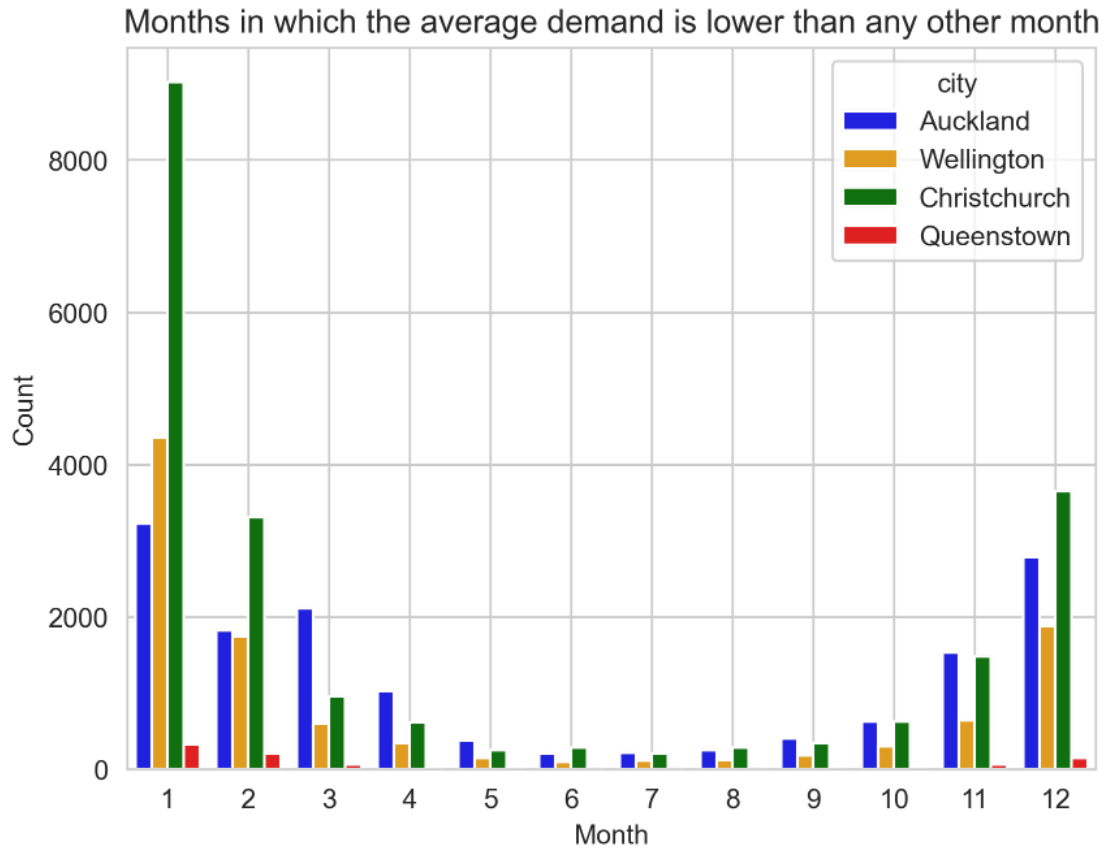


Figure 9: Months in which minimum average load occurred.

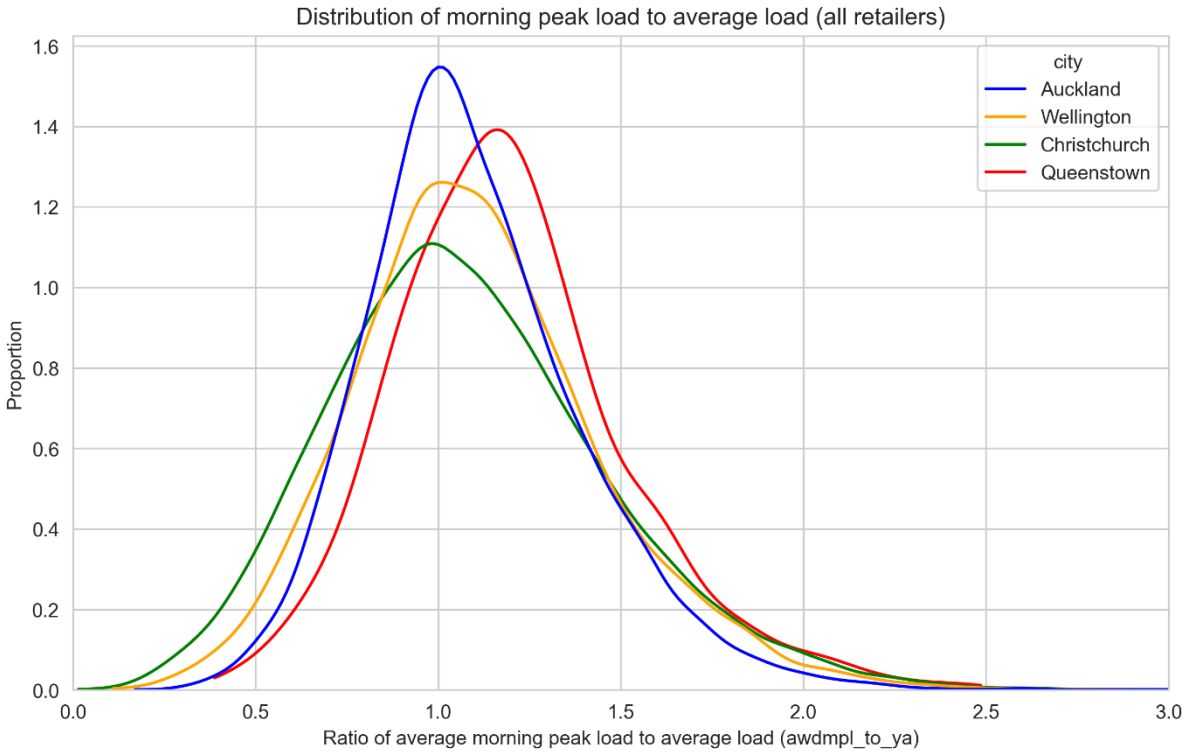


Figure 10: Ratio of average weekday morning peak load to average load over the year. The morning peak period is defined as 7am to 11am.

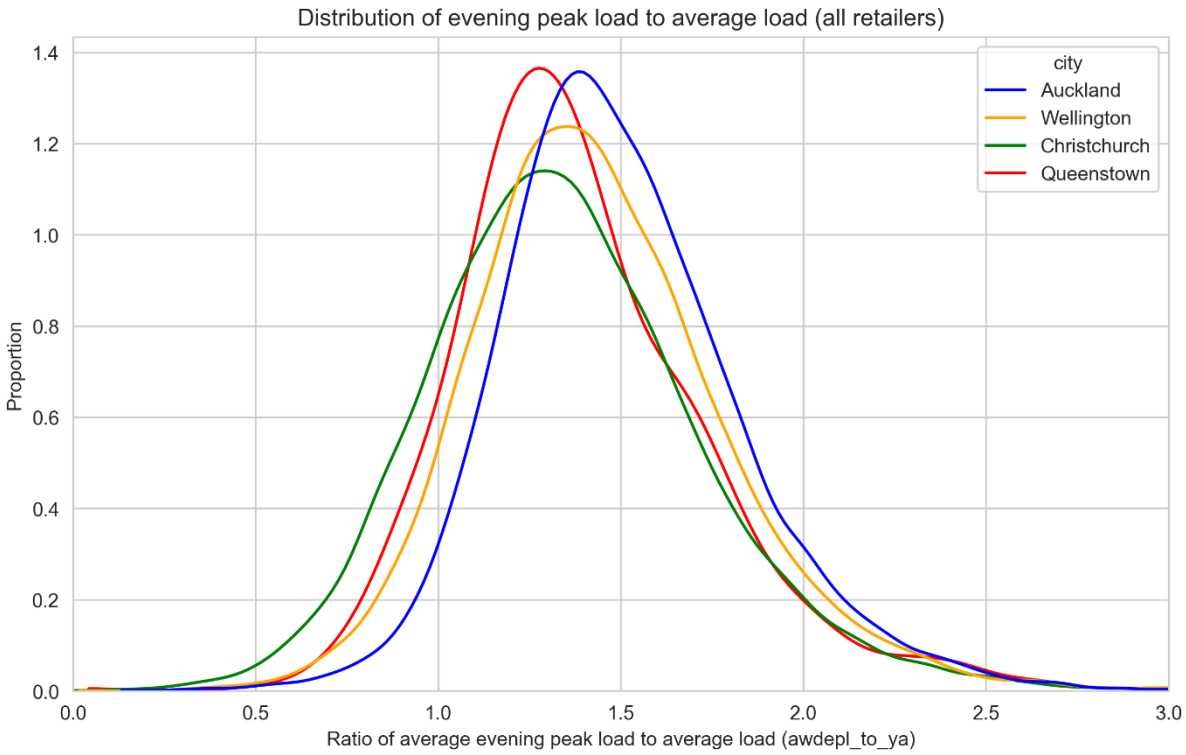


Figure 11: Ratio of average weekday evening peak load to average load over the year. The evening peak period is defined as 5pm to 9pm.

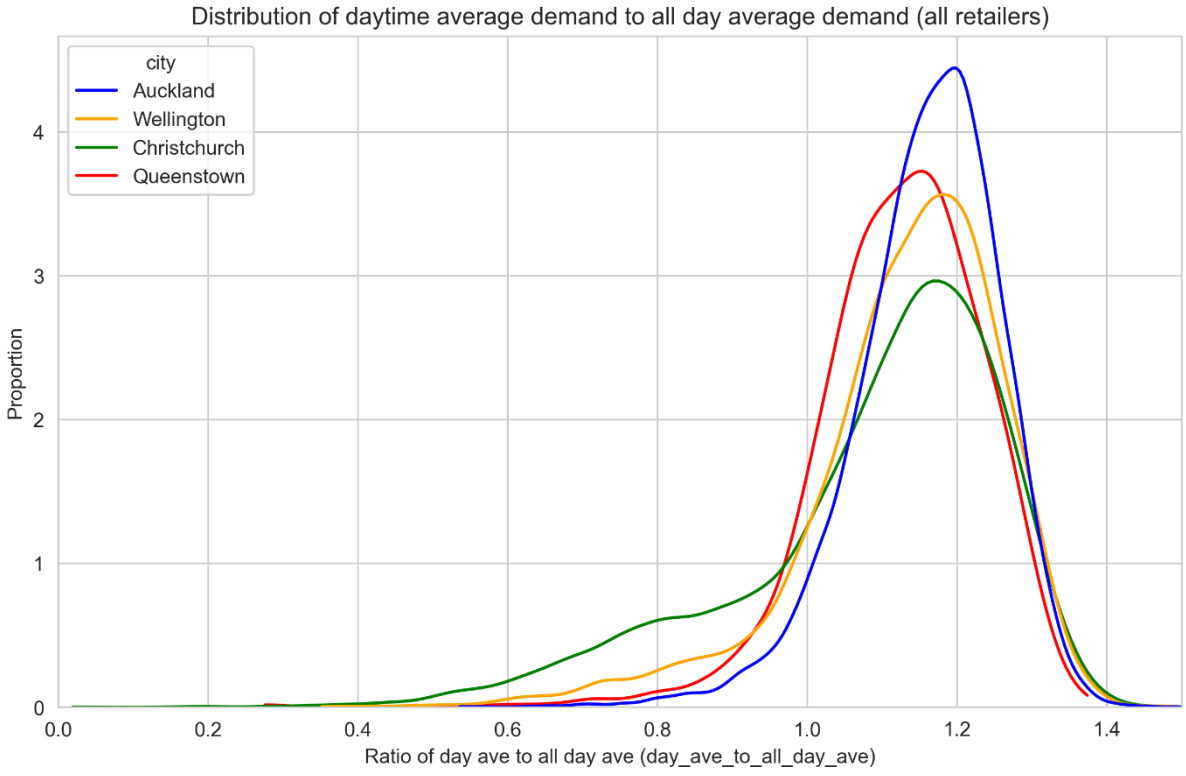


Figure 12: Ratio of daytime average load to all day average load over the year. Daytime is defined as 7am-11pm.

3.2 Frequency domain parameters

Frequency domain parameters were determined by taking the discrete Fourier transform (DFT) of each load profile over an interval of 364 days, since this is an exact multiple of one week, and should still pickup annual or twice annual events such as summer/winter. Note that taking the DFT over 365 days could introduce spectral leakage, resulting in a 'smearing' of frequency signals across multiple frequencies. Only the magnitude of the output was considered in the first instance since the strength of each frequency component was of interest, rather than its phase. A frequency domain filter was also applied so that only strong signals appear in the resulting spectrum. Each signal was then normalised to the DC value (the value at a frequency of zero, which in this case is the average annual load), so that each frequency signal could be related to the average annual load. The unit of frequency adopted was cycles per year.

Figure 13 and Figure 14 give an example spectrum from four different load profiles and point out frequency components in the spectrum. Figure 15 to Figure 18 show the load profiles of these loads, where the frequency signals identified in the spectrum are evident. Because the spectrum clearly shows frequency components that seem to relate to consumption patterns, the idea of using these in establishing clusters seemed promising. The nature of the frequency components over all load profiles was therefore investigated further, with the parameters summarised in

Table 4.

*Table 4: Frequency domain load profile parameters investigated. * Indicates parameters selected for clustering trials discussed in Section 4. ** Indicates parameters trialled and ultimately used to define the clusters used in the study.*

Period	Parameter	Reference and notes
Annual	Annual frequency component *	Figure 19. While this shows good spread between cities and within each city, it was not used, as other means of differentiating seasons and cities were used.
Biannual	Twice annual frequency component *	Figure 20. Discarded as it shows little spread between load profiles.
Quarterly	Quarterly frequency component	Figure 21. Discarded as it shows very little spread between load profiles.
Monthly	Monthly frequency component	Figure 22. Discarded as it shows very little spread between load profiles.
Weekly	Weekly frequency component *	Figure 23. Discarded as it shows very little spread between load profiles.
Daily	Daily frequency component *	Figure 24.
Twice daily	Twice daily frequency component *	Figure 25.
Three times a day	Three-times-a-day (every eight hours) frequency component	Figure 26. Discarded as it shows little spread between load profiles.
Four times a day	Four-times-a-day (every six hours) frequency component	Figure 27. Discarded as it shows little spread between load profiles.
Six times a day	Six-times-a-day (every four hours) frequency component	Figure 28. Discarded as it shows very little spread between load profiles.

The following observations are made from the distributions of frequency components in Figure 19 to Figure 28.

1. There are large differences between cities in the annual frequency component. However, difference between cities is dealt with through representative load profiles of each city selected from clusters. The annual frequency component also shows good spread. However, as discussed later, load profiles are selected from clusters where the maximum month occurs in winter, which selects for a degree of seasonality.
2. The biannual, three-times-a-day, four-times-a-day and in particular quarterly, monthly, weekly, and six-times-a-day, show little spread compared to annual, daily and twice daily. These were discarded, although biannual and weekly were used in the clustering trials discussed in Section 4.
3. Daily and twice daily show good spread and are therefore considered as potential clustering parameters.

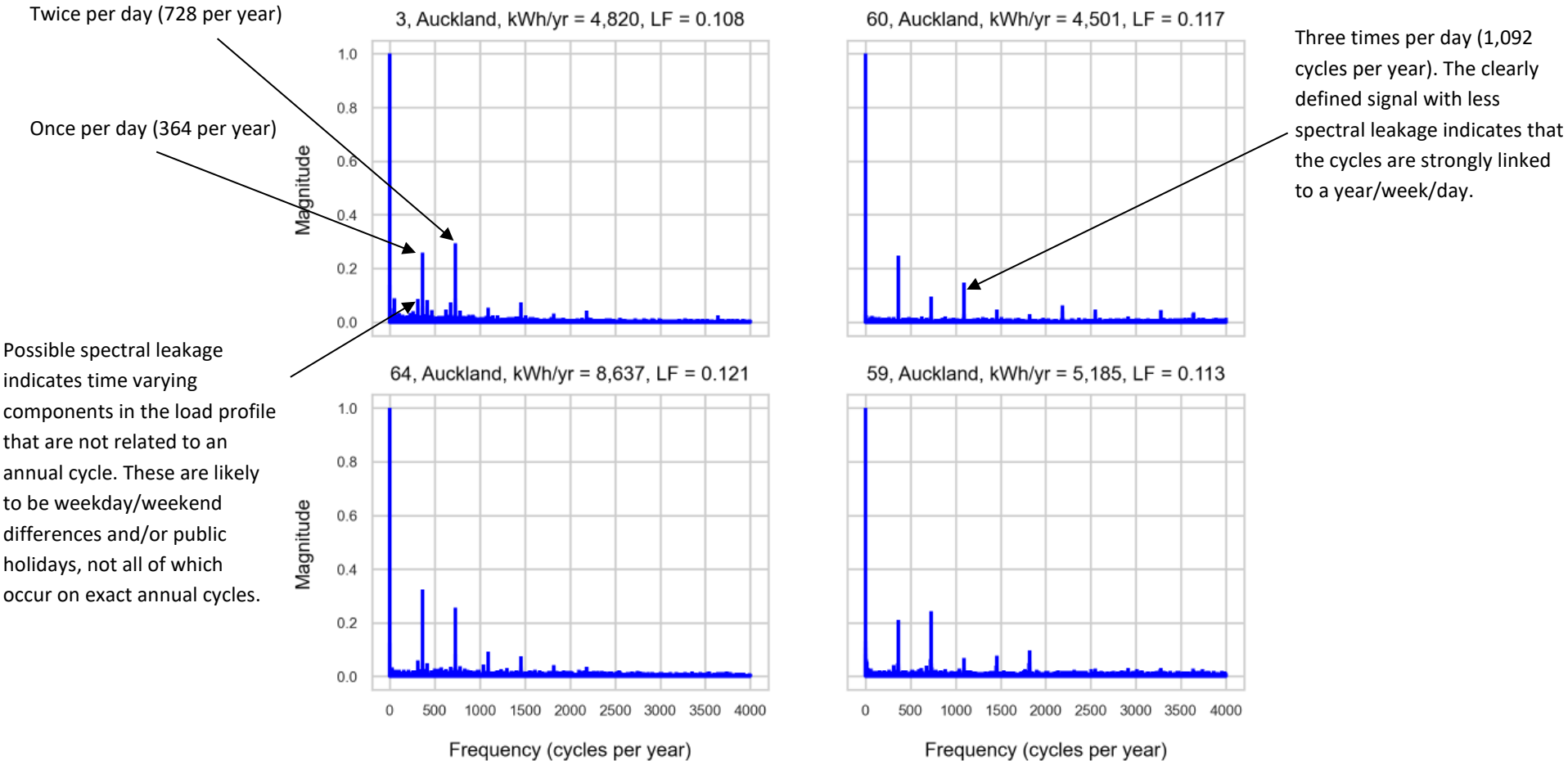


Figure 13: Load profile spectrum examples. The title of each spectrum gives the load profile identifier, the city, the annual consumption (kWh pa), and the load factor (LF).

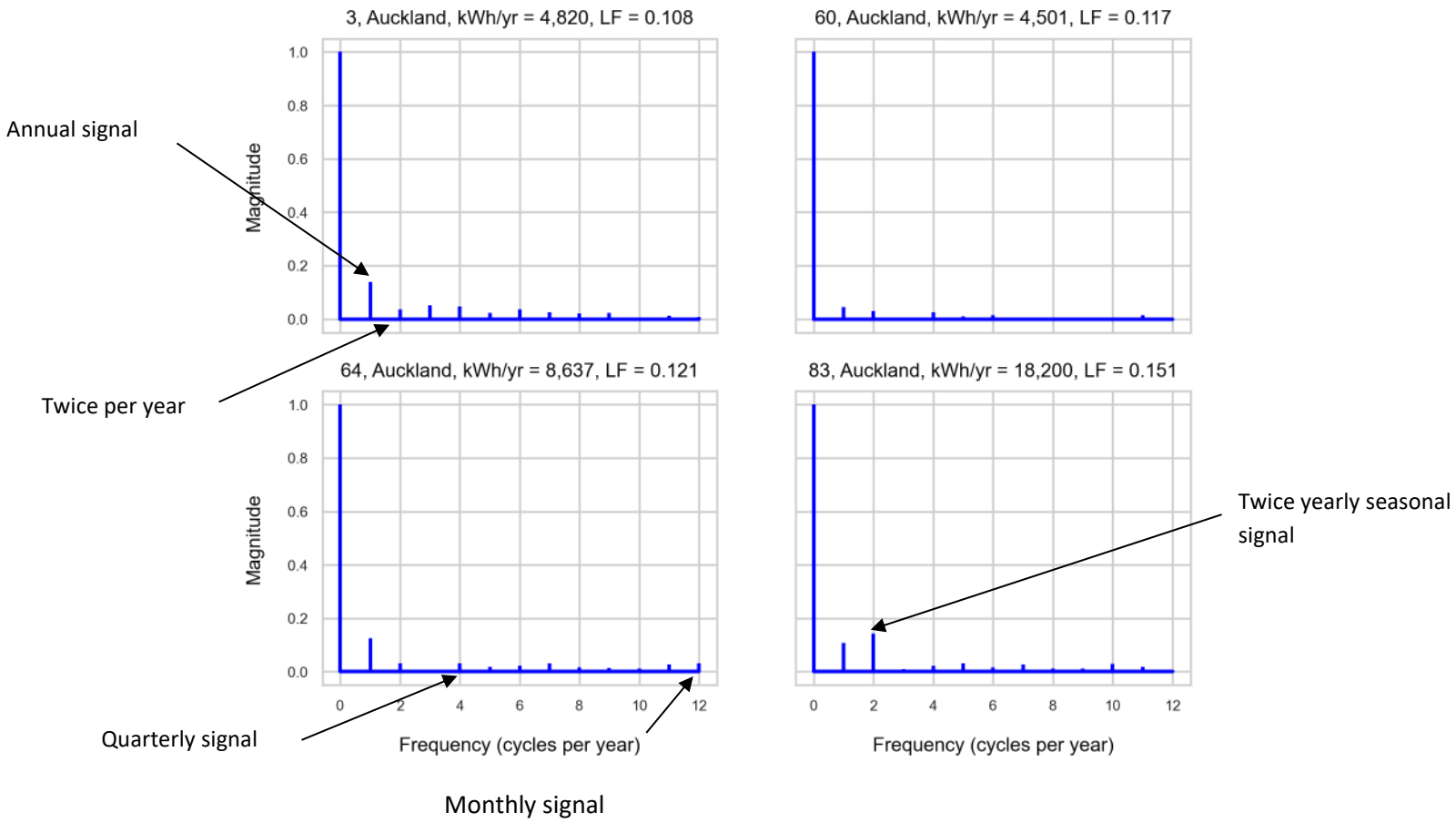


Figure 14: Load profile spectrum examples with the x-axis expanded. Load profile 59 has been substituted with load profile 83 to show some interesting seasonal patterns.

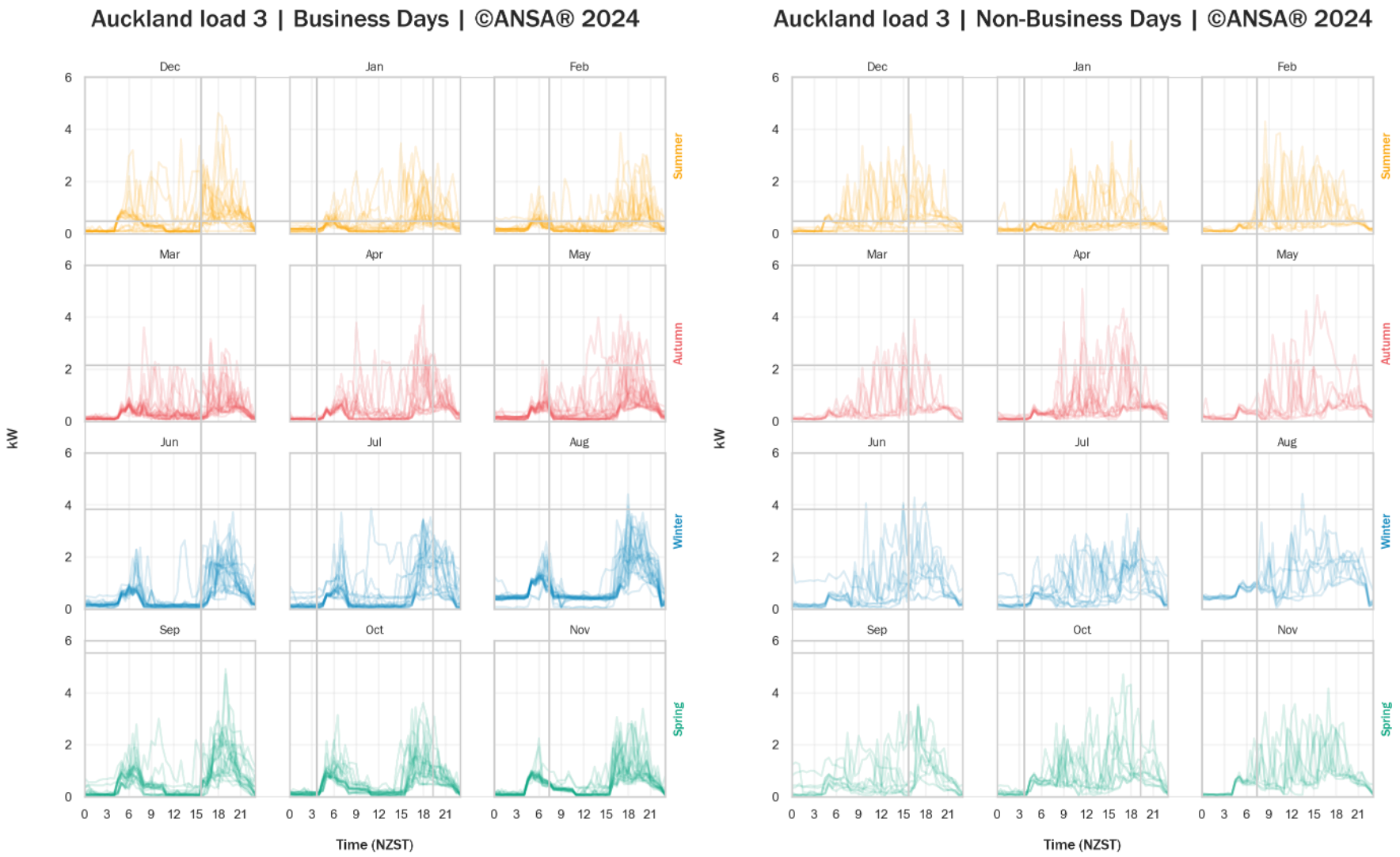


Figure 15: Load 3 showing the strong twice daily pattern.

Figure 16: Load 60 showing the strong daily pattern and weaker three times a day pattern.

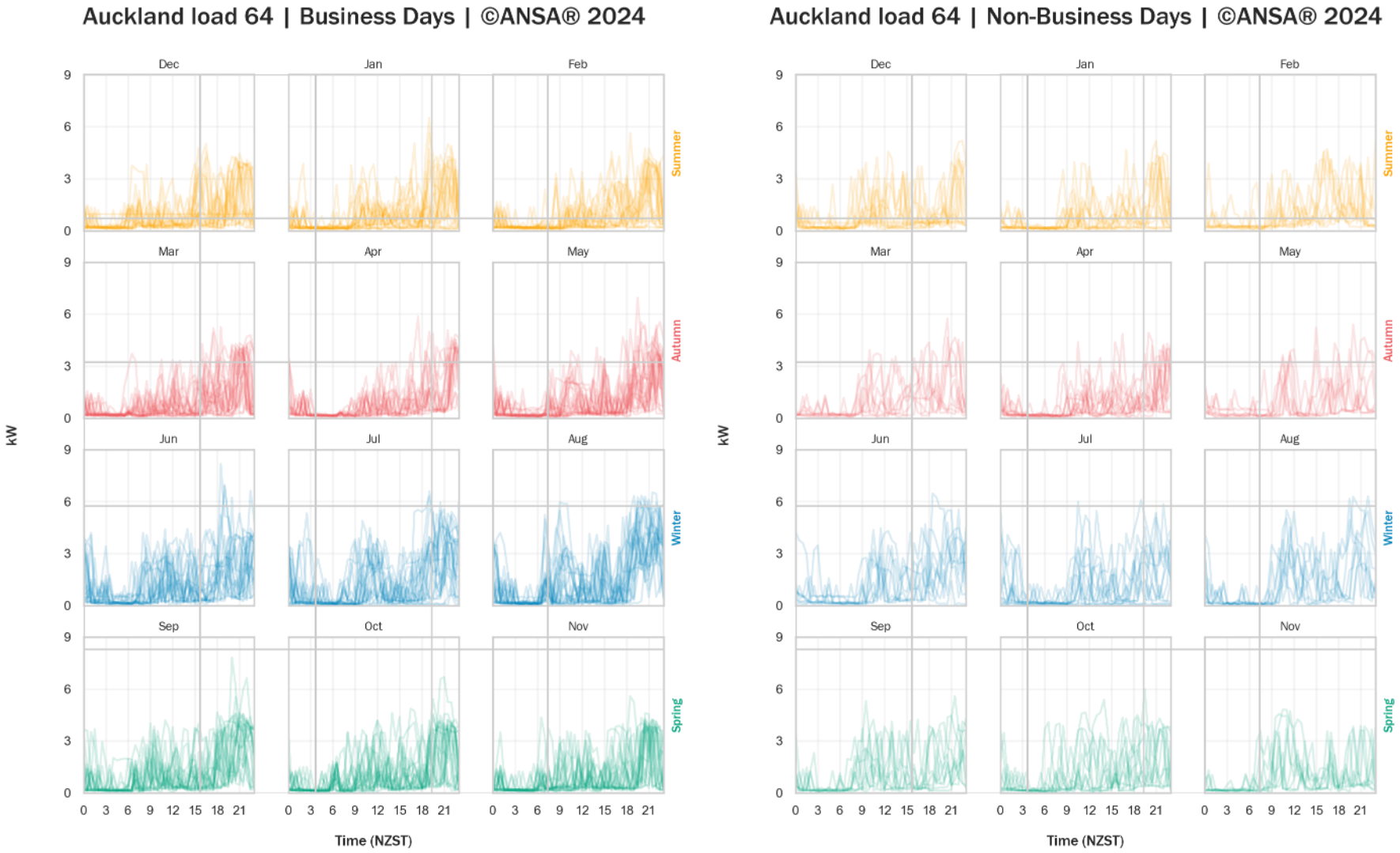


Figure 17: Load 64 showing the strong daily and to a lesser extent twice daily pattern.

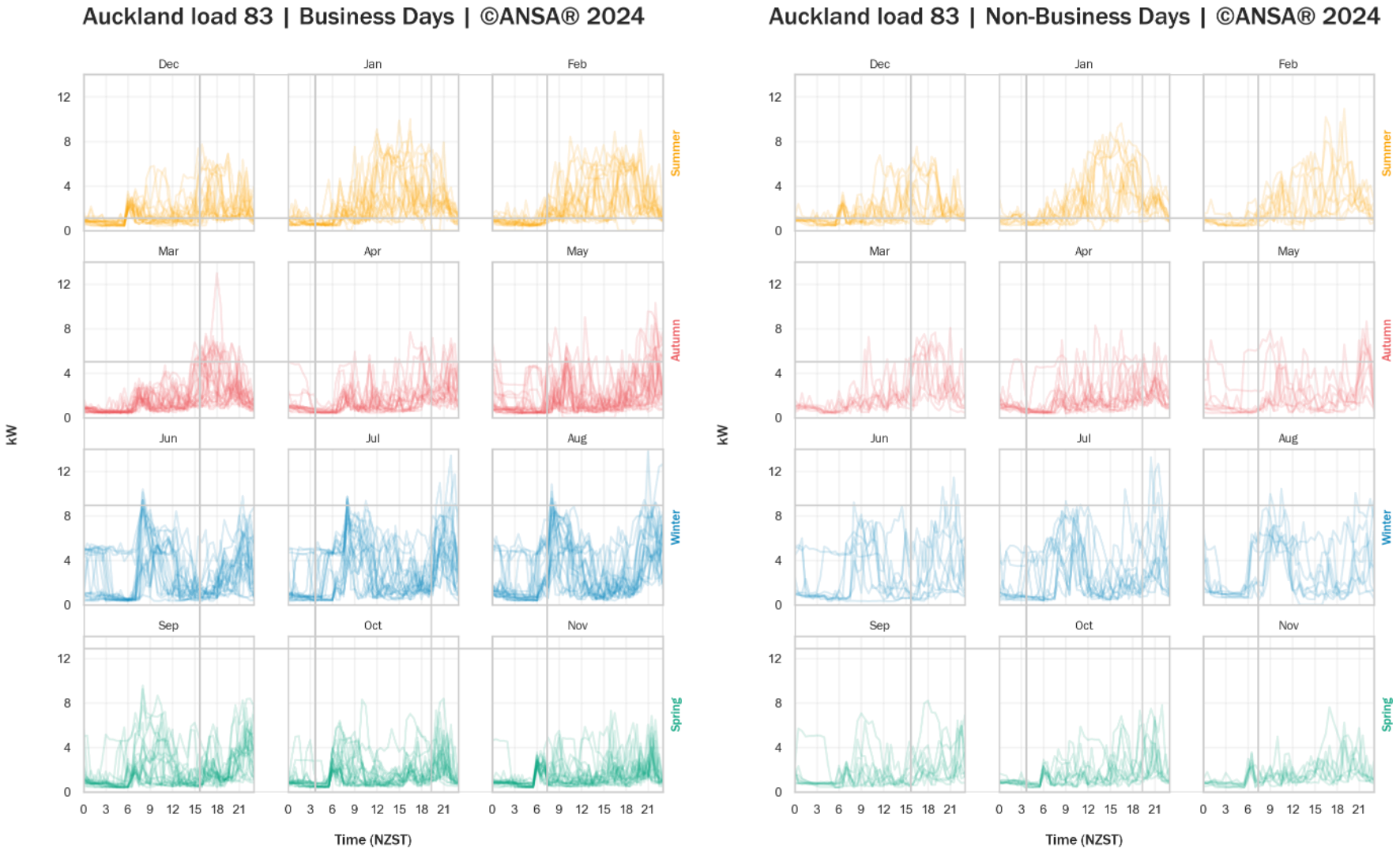


Figure 18: Load 83 with twice per year season signals evident (Autumn and Spring are generally lower than Summer and Winter).

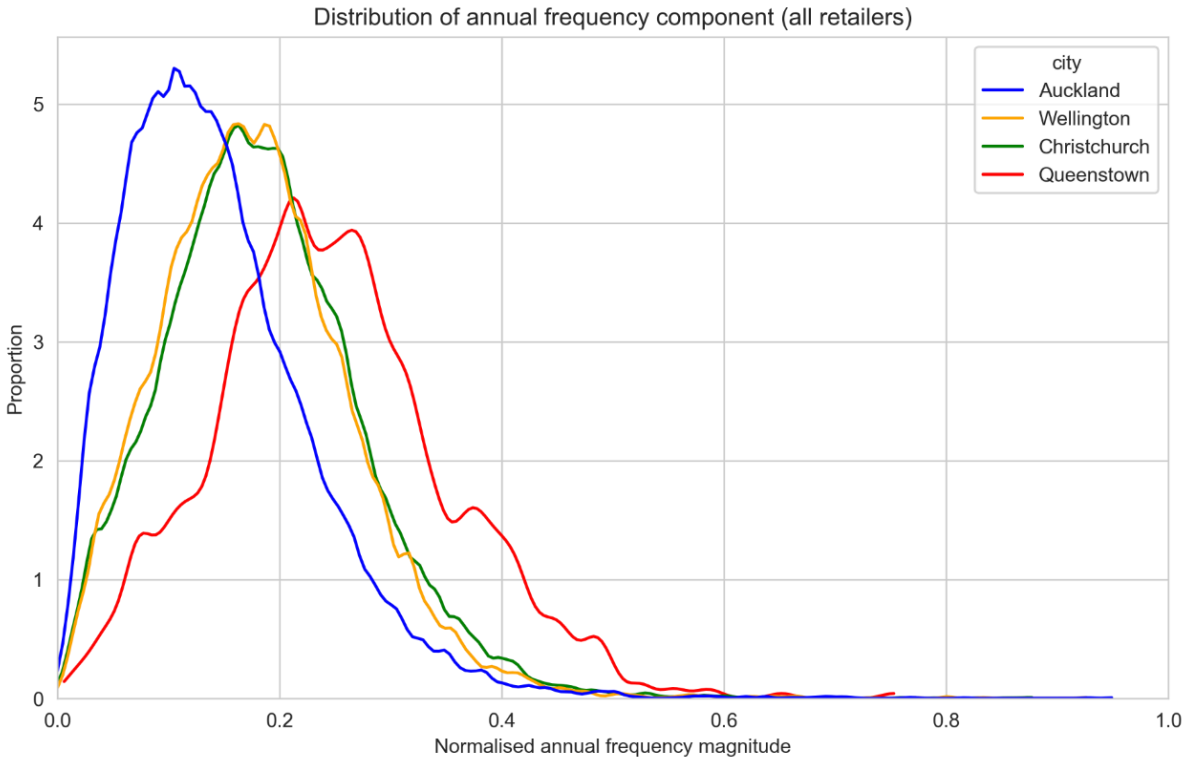


Figure 19: Distributions of the annual frequency components in the load profiles from each city. Since these represent events that happen once per year, it represents seasonality (summer/winter). The x-axis gives the normalised frequency magnitude which represents the strength of the 'once per year' signal in the spectrum of all load profiles. Zero is non-existent, while one is equivalent to the average load itself.

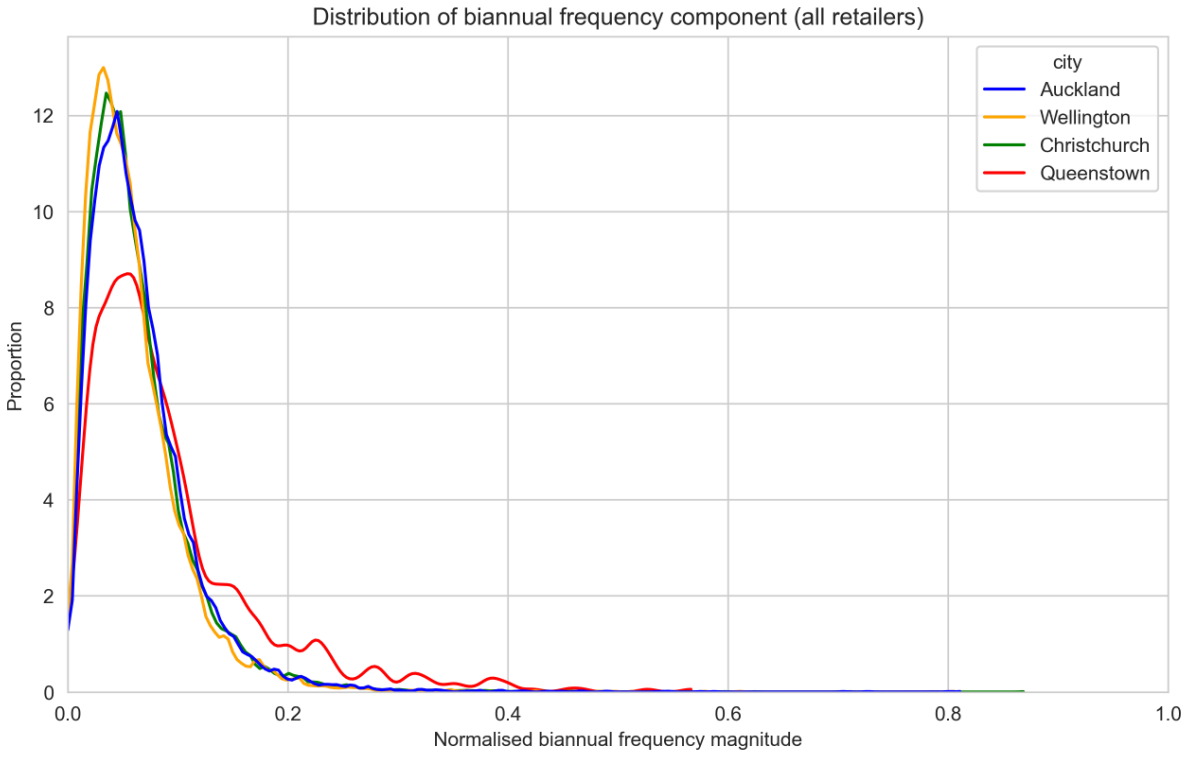


Figure 20: Distributions of the twice per year (biannual) frequency components in the load profiles from each city.

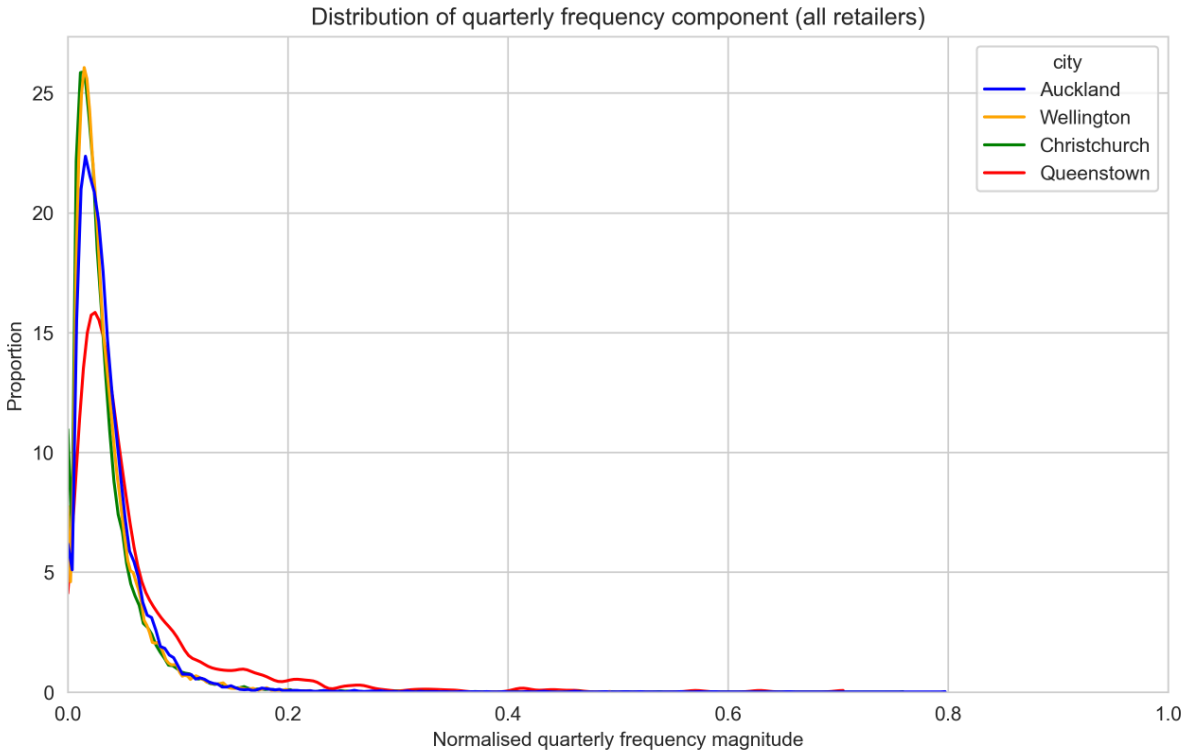


Figure 21: Distributions of the quarterly frequency components in the load profiles from each city.

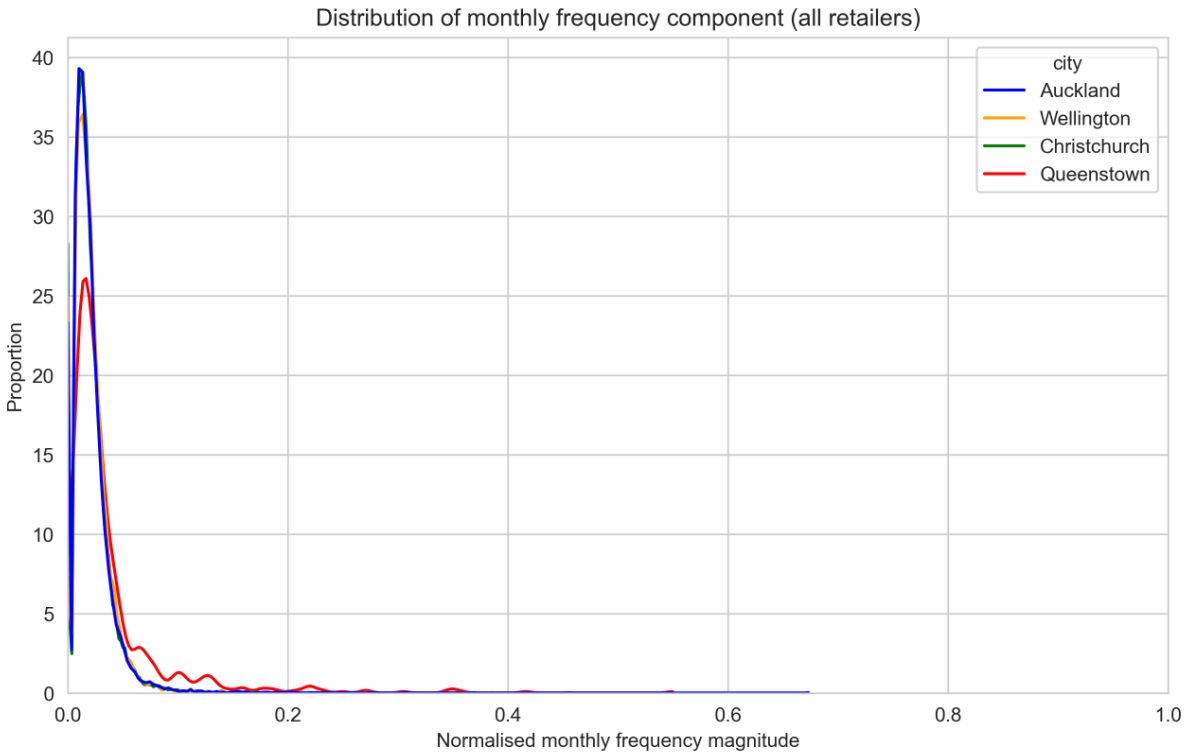


Figure 22: Distributions of the monthly frequency components in the load profiles from each city.

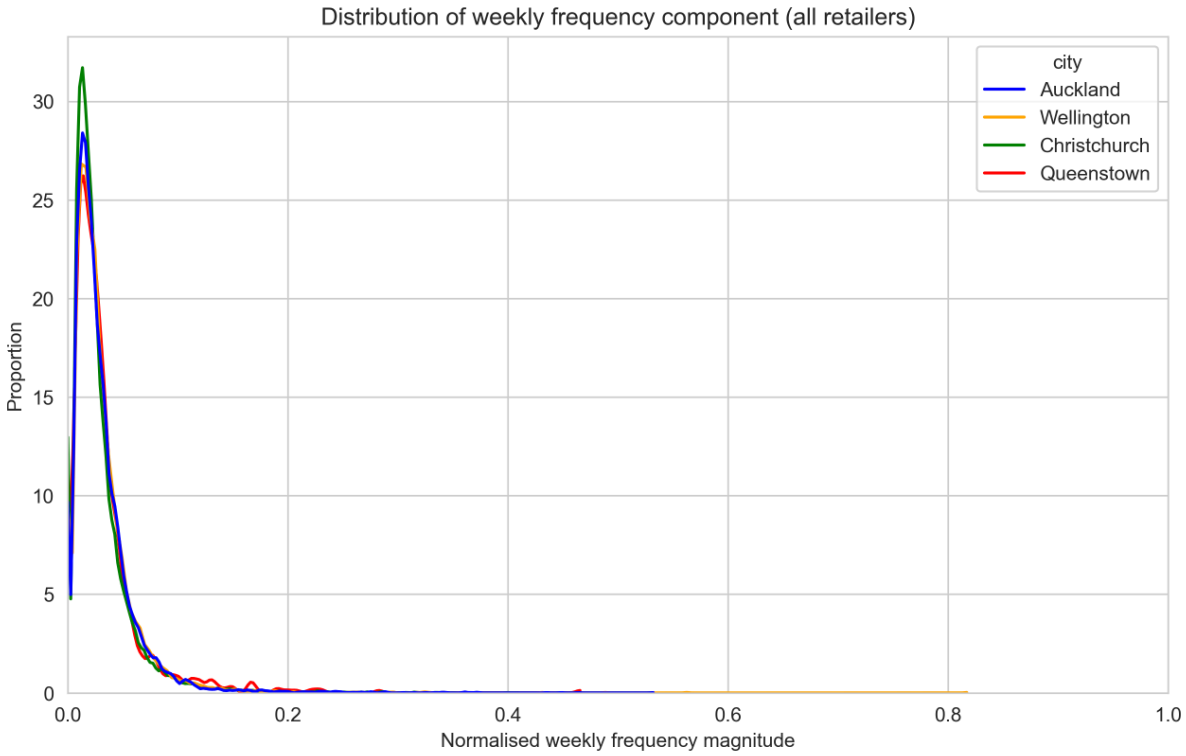


Figure 23: Distributions of the weekly frequency components in the load profiles from each city.

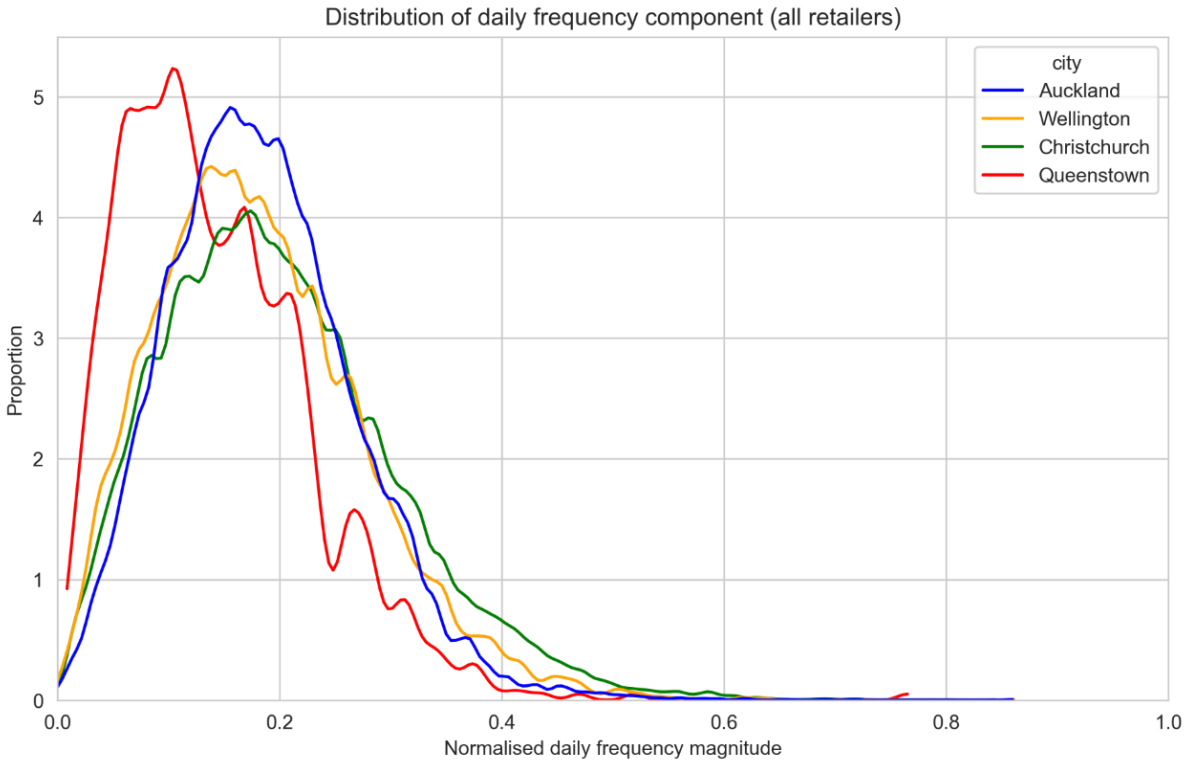


Figure 24: Distributions of the daily frequency components in the load profiles from each city.

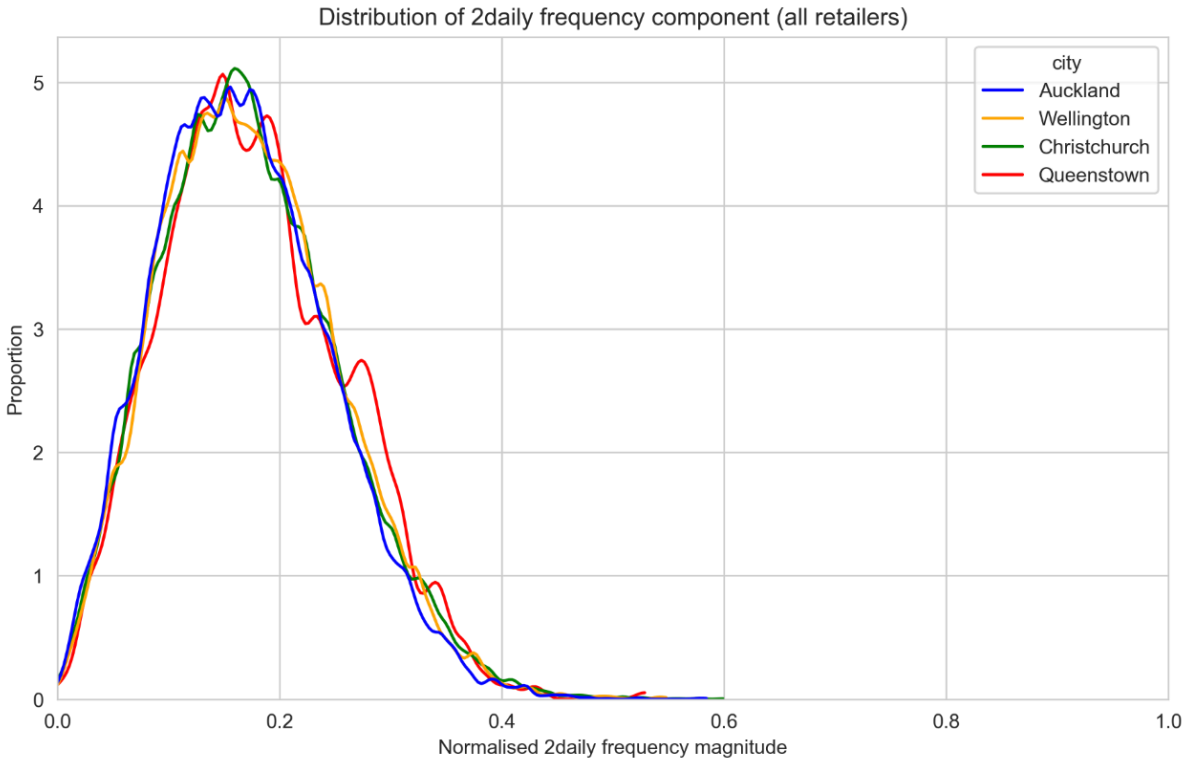


Figure 25: Distributions of the twice daily frequency components in the load profiles from each city.

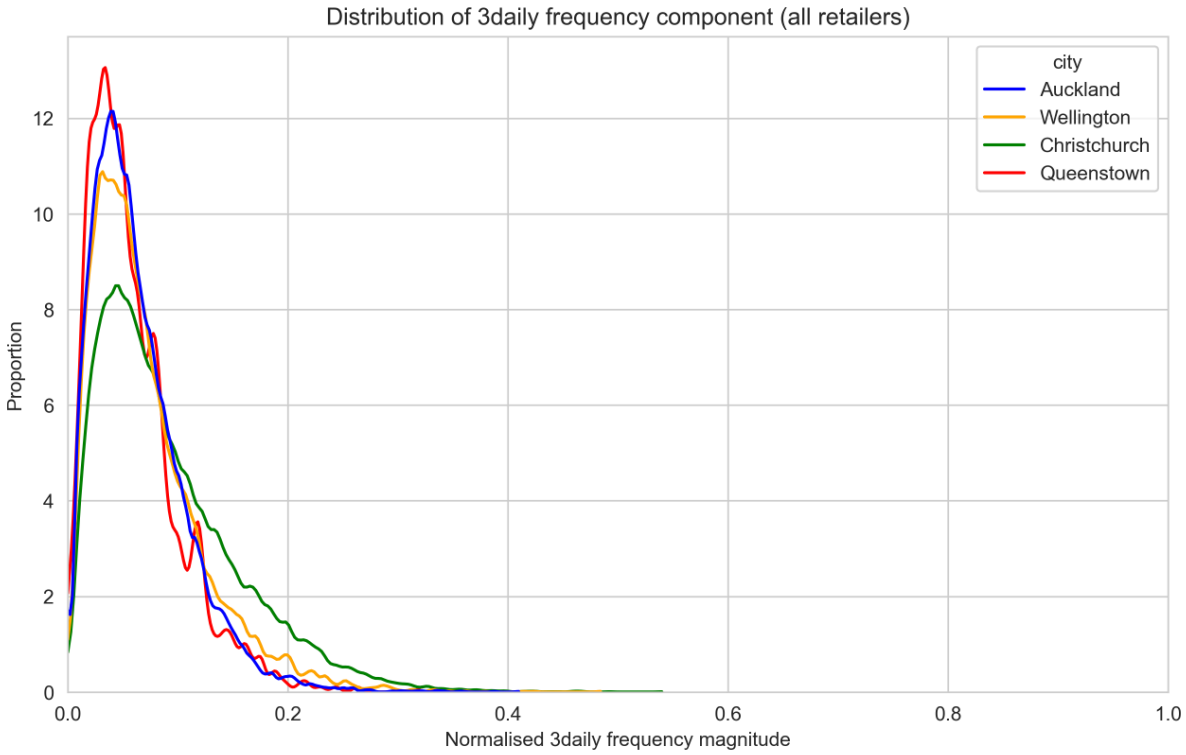


Figure 26: Distributions of the three-times-a-day (every eight hours) frequency components in the load profiles from each city.

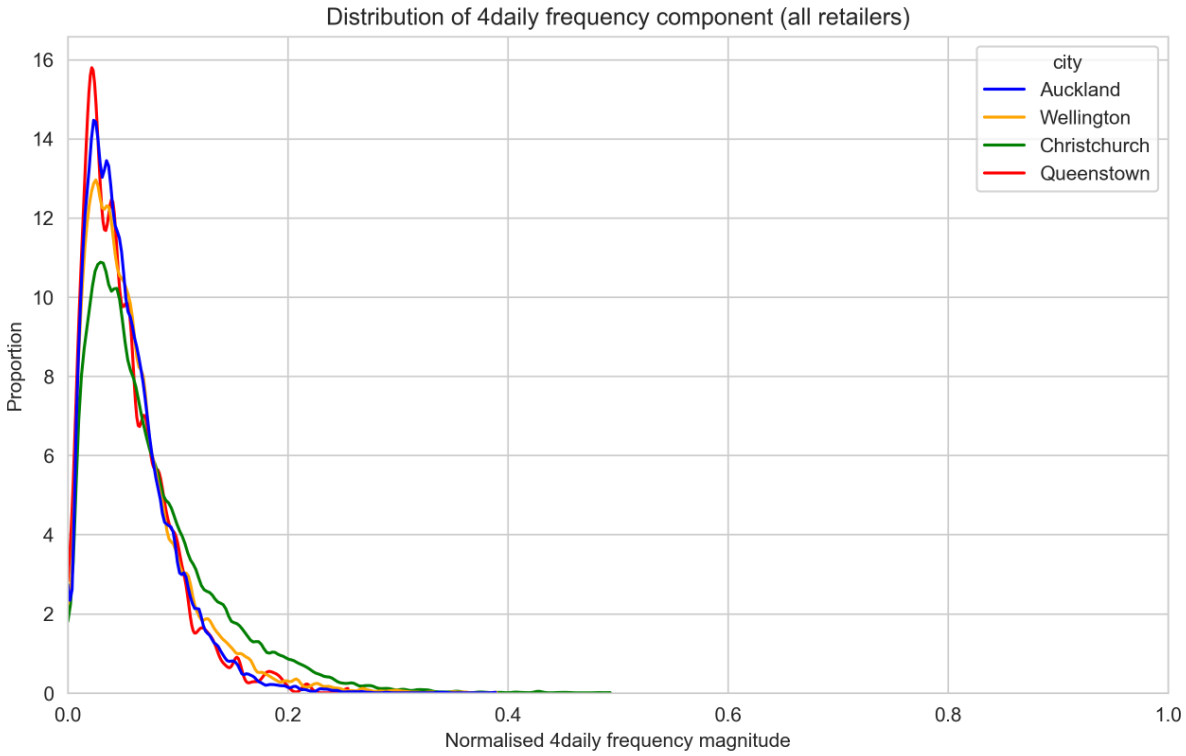


Figure 27: Distributions of the four times-a-day (every six hours) frequency components in the load profiles from each city.

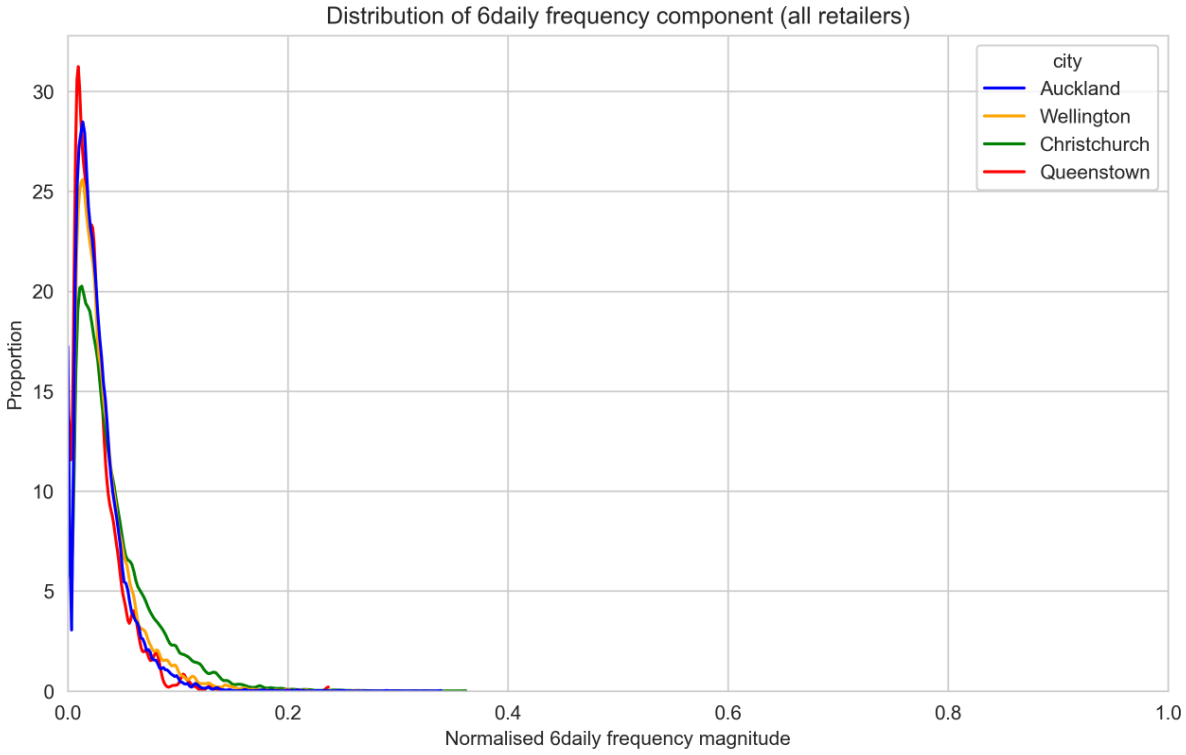


Figure 28: Distributions of the six times-a-day (every four hours) frequency components in the load profiles from each city.

4 Clustering the load profiles including selection of clustering parameters

The K-Means clustering method was used since it is a method specifically designed for classification of data. K-NN (K-nearest neighbour) and decision trees were considered and rejected as they require training data sets which did not match the problem in this case. Two challenges with K-Means clustering, and indeed any clustering, are:

- finding appropriate parameters that define features of the data to be clustered (load profiles in this study), and
- settling on an appropriate number of clusters.

If there are too many parameters, the problem becomes an enormously complex set of multidimensional clusters, with over-fitting occurring, especially if some of the parameters are correlated. In addition, categorical parameters (season for maximum load for example) can cause large swings in clusters, when in fact most load profiles clustered from other features tend to cluster towards a particular maximum/minimum season anyway. Further, selecting too many clusters results in lack of differentiation between them.

The balance is to find the number of clusters where the error between cluster members and the centroid (an intra-cluster measure) is minimal, and the distance between clusters (an inter-cluster measure) is maximised. The sum of squared errors within clusters, and a silhouette function between clusters, were used to measure these. Trials were then run with various combinations of 12 parameters noted from Section 3, and the number of clusters varied from two to 25. In each trial the number of clusters was varied, seeking to minimise the sum of squared errors (intra-cluster error) while maximising the silhouette function (inter-cluster distance).

In addition to trialling combinations of parameters, correlations between parameters were investigated, allowing more of the 12 parameters to be discarded. For example:

- Figure 29 shows good correlation between annual average load factor and annual average load, leading to discarding of annual average load factor in the cluster parameter selection, since the representative loads are selected based on average annual load as mentioned previously.
- However, annual average load factor was also eventually discarded as a parameter. This is partly due to the eventual parameters selected partially representing load factor – a high morning peak load tends to correlate with a low load factor, as does a high evening peak load, as indicated in Figure 30.

Further to this, clusters were eventually formed that represent a range of annual consumptions, with cluster samples then selected based on city and annual consumption. This then enabled annual average load and annual average load factor to be discarded as a parameters. This also ensured that a load profile used to assess solar in a particular city was actually selected from that city.

In addition, it was found that frequency domain parameters added little to the cluster selection due to good correlation with some of the time domain parameters.

This process was repeated, with a selection of results summarised in **Error! Reference source not f ound.**. In combination with the cluster trials outlined above, this led to the eventual selection of eight clusters (the last row of **Error! Reference source not found.**), determined from the following t hree parameters:

- Ratio of average daytime load to all day load (Figure 12),
- Ratio of average weekday morning peak load to average weekday load over the year (Figure 10), and
- Ratio of average weekday evening peak load to average weekday load over the year (Figure 11).

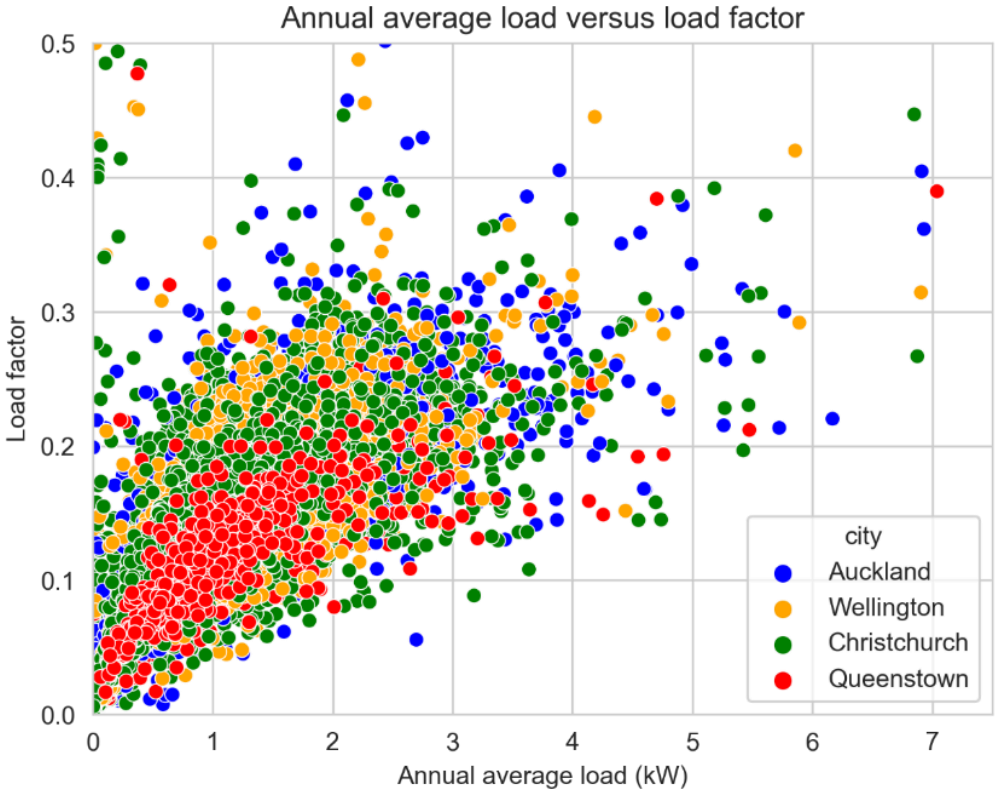
In selecting the number of clusters, diminishing returns are rapidly experienced in the sum of squared errors, as illustrated in Figure 31 for the above parameters. The method usually adopted is to choose the 'elbow' of the sum of squared errors versus the number of clusters, which gave five clusters. On further examination of the clusters there was some overlap. Hence, the number of clusters was increased to eight. Section 5 discusses the clusters selected.

Table 5: Parameter selections trialled for cluster formation. The last row gives the parameters eventually used to form the clusters.

Parameters used to form clusters	Number of clusters	Sum of squared error	Silhouette score
All time domain (those with * or ** from Table 3)	5	2,800	0.3
	8	1,900	0.26
All frequency domain (those with * or ** from Table 4)	5	1,000	0.22
	8	800	0.195
All time domain, no average annual consumption	5	1,800	0.35
	8	1,400	0.31
All frequency domain with annual average consumption	5	1,500	0.195
	8	1,200	0.17
Three time domain and three frequency domain[1,2]	5	1,000	0.21
	8	1,800	0.18
Three time domain and three frequency domain with seasons	5	2,300	0.35
	8	1,800	0.27
Three time domain and four frequency domain with seasons[3]	5	1,600	0.34
	8	1,900	0.24
Three time domain and four frequency domain with seasons and annual average demand	5	3,100	0.25
	8	2,200	0.22
All time domain and all frequency domain with annual average demand removed	5	3,500	0.23
	8	3,000	0.205
All time domain and all frequency domain with annual average demand removed and minimum season removed	5	900	0.295
	8	650	0.26
All time domain and all frequency domain with annual average demand removed, minimum season removed, and evening peak removed	5	550	0.31
	8	393	0.299
Just ratio of average daytime load to all day load, ratio of average weekday morning peak load to average weekday load over the year, and ratio of average weekday evening peak load to average weekday load over the year	5	332	0.294
	8	**270**	**0.277**

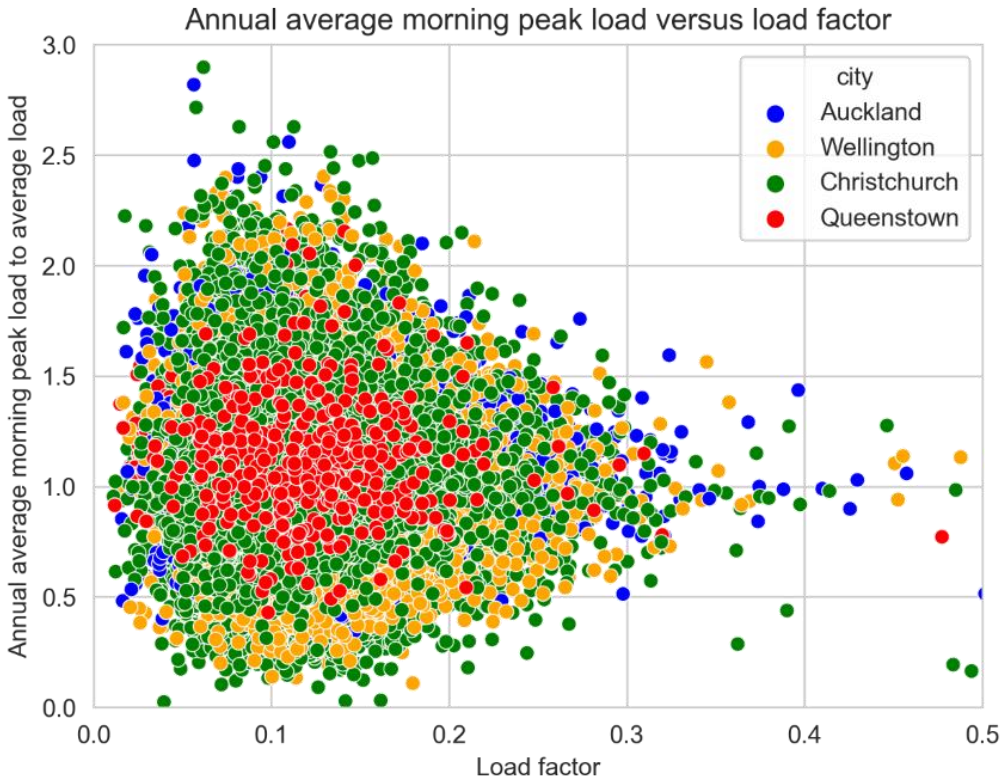
[1] Time domain: average annual load factor, average weekday morning peak load to annual average load, and average daytime load to average all day load. Frequency domain: Annual, weekly, and daily.
[2] Silhouette score indicates high overlap between clusters, even though the within cluster sum of squared errors is relatively low.
[3] Biannual frequency signal is added.

(a)

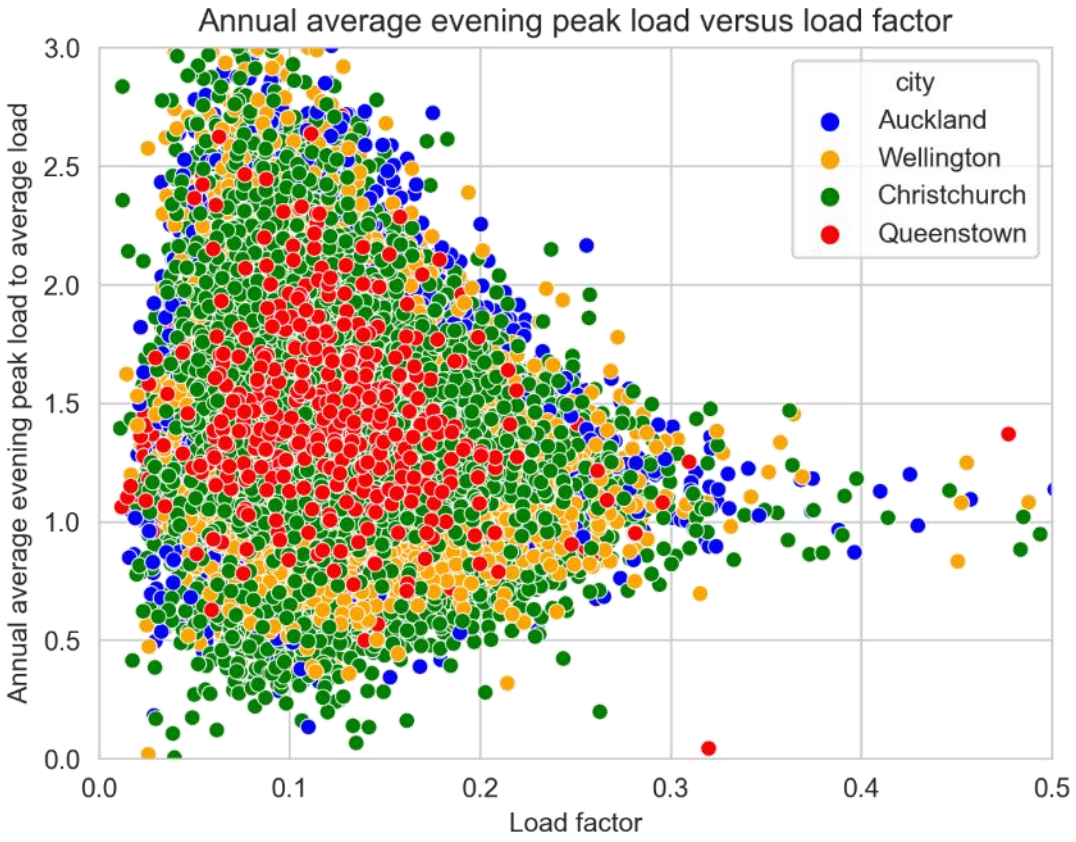


(b)

Figure 29: Average annual load factor versus average annual load, showing a degree of correlation between them. The strong diagonal line in (a) is the result of discarding samples with load above 15 kW, annual load above 31,000 kWh (average load of 3.54 kW), and annual load below 900 kWh. (b) shows the full sample set before these samples were discarded.



(a)



(b)

Figure 30: (a) Average annual weekday morning peak load to average load versus average annual load factor; and (b) Average annual weekday morning peak load to average load versus average annual load factor.

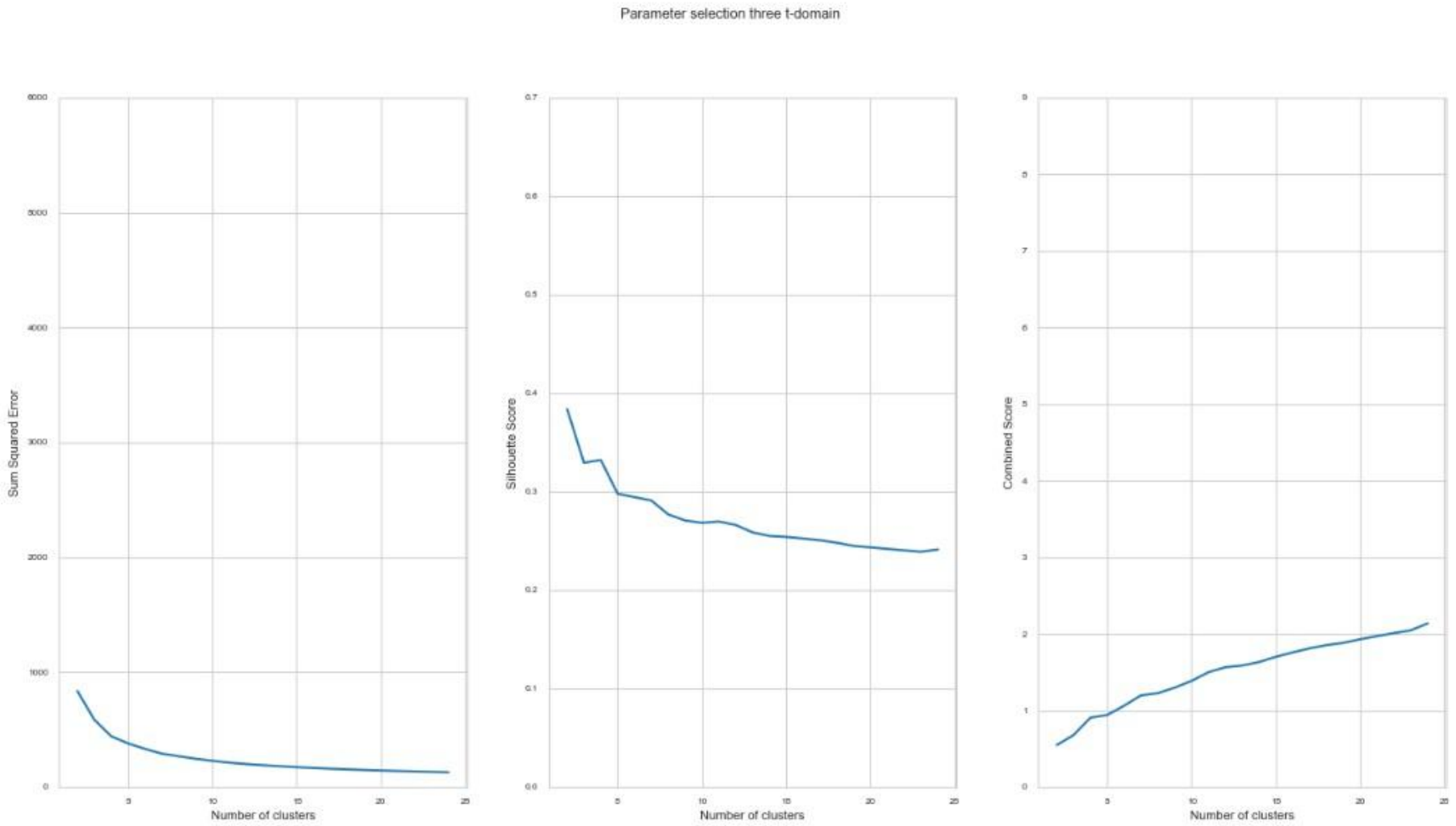


Figure 31: Sum of squared errors, silhouette score, and a combined score of the cluster parameters used (the last two rows of Table 5).

5 Summary of the clusters determined and used in the study

As summarised in the previous section, the three parameters shown in Table 6 were used to establish eight clusters. Figure 32 to Figure 34 show the clusters visually by these parameters, showing the size of each cluster and the separation between them.

Table 6: Clustering parameters used to determine the clusters used in the study.

Parameter	Name used in the main report	Symbol in this appendix	Figure giving the details of the parameters
Ratio of average daytime load to average all day load over the year	*daytime demand ratio* (summarised further to just 'daytime' in graphs and tables.)	day_ave_to_all_day_ave	Figure 12
Ratio of average weekday morning peak load to average weekday load over the year	*morning peak ratio* (summarised further to just 'morning' in graphs and tables.)	awdmpl_to_ya	Figure 10
Ratio of average weekday evening peak load to average weekday load over the year	*evening peak ratio* (summarised further to just 'evening' in graphs and tables.)	awdepl_to_ya	Figure 11

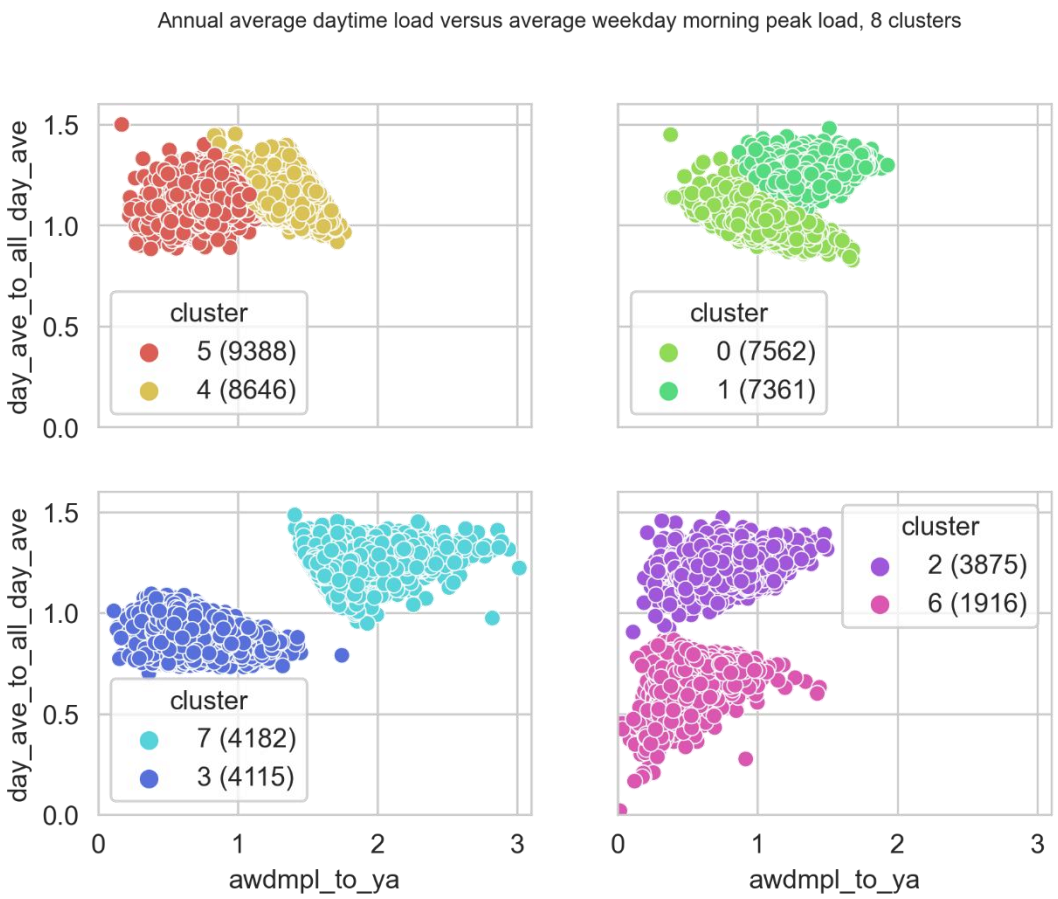


Figure 32: Daytime demand ratio versus morning peak ratio for the eight clusters. Clusters are ordered from the largest (the greatest number of load profiles) to smallest (the lowest number of load profiles).

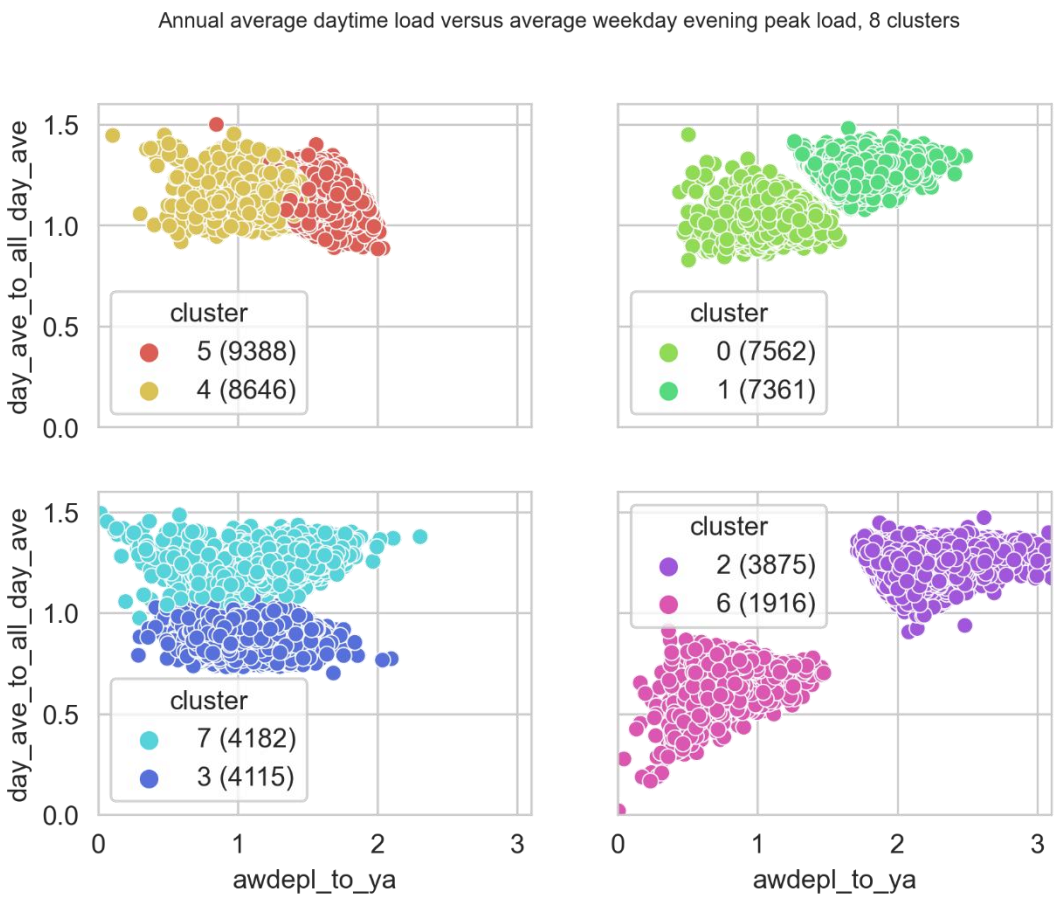


Figure 33: Daytime demand ratio versus evening peak ratio for the eight clusters.

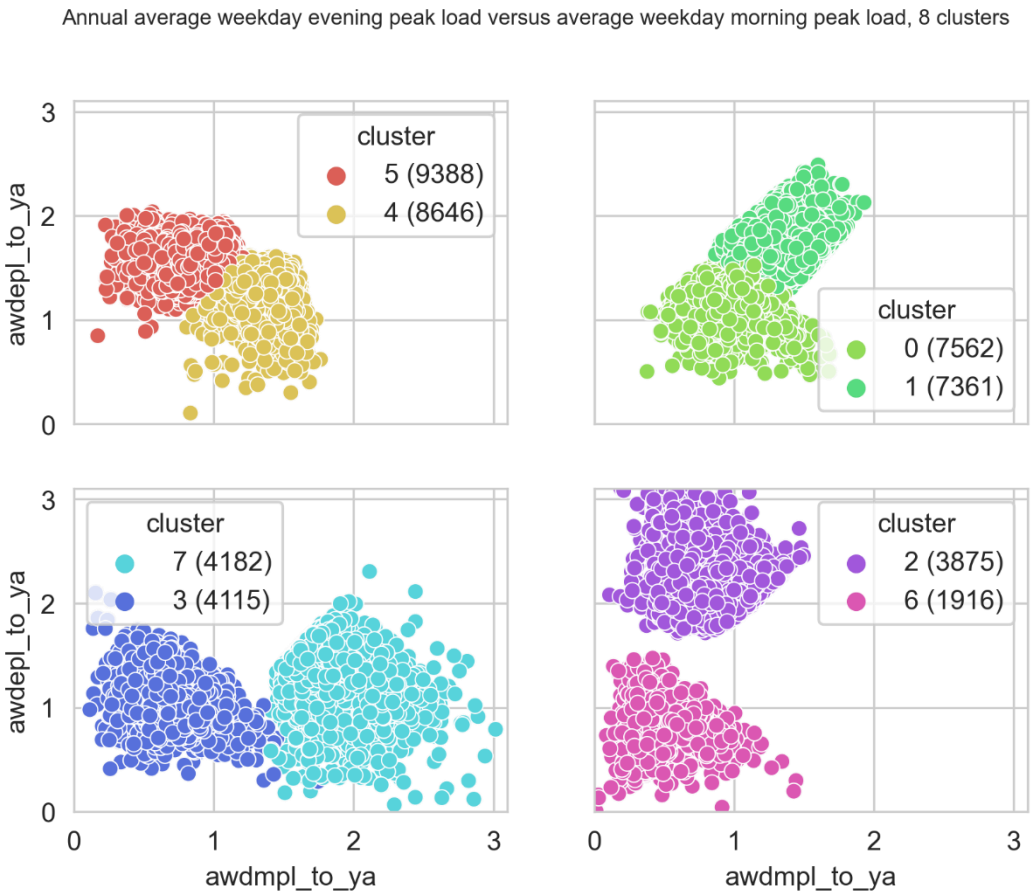


Figure 34: Evening peak ratio versus morning peak ratio for the eight clusters.

Table 7: Load profiles in each cluster, and the number representing each city, with the cell shading indicating the number of load profiles.

Cluster	Total number of load profiles selected	Number of load profiles selected from			
		Auckland	Wellington	Christchurch	Queenstown
0	7,562	2,227	1,758	3,358	219
1	7,361	2,726	1,721	2,794	120
2	3,875	1,529	897	1,387	62
3	4,115	473	783	2,821	38
4	8,646	2,800	2,028	3,573	245
5	9,388	3,784	2,112	3,359	133
6	1,916	47	272	1,588	9
7	4,182	987	943	2,158	94
Total	47,045	14,573	10,514	21,038	920

Table 8: A description of each cluster, with the median values of each parameter given for each cluster.

Cluster	Description	Cluster parameter medians (load ratios within the time periods given)		
		Day average to all day average	Average weekday morning peak load to average load	Average weekday evening peak load to average load
0	medium daytime & medium morning & medium evening	1.05	1.01	1.19
1	very high daytime & high morning & very high evening	1.24	1.26	1.67
2	very high daytime & low morning & very high evening	1.23	0.84	2.08
3	low daytime & low morning & medium evening	0.89	0.76	1.05
4	high daytime & high morning & high evening	1.15	1.30	1.30
5	high daytime & low morning & very high evening	1.14	0.90	1.56
6	very low daytime & very low morning & low evening	0.69	0.57	0.84
7	very high daytime & very high morning & high evening	1.25	1.75	1.29

6 Selection of representative loads for the study from the clusters

As alluded to earlier, a subset of two individual load profiles that represent each cluster were selected from each cluster. In addition, these were selected to represent each of the four cities, and annual consumptions of 8,000 kWh pa and 12,000 kWh pa. Hence there are up to 128 load profiles in the study that should represent all consumers (four cities x two annual consumption levels x two samples per city x eight clusters equates to up to 128 individual load profiles). Practically, there are fewer due to not every combination being present in each cluster for each city and annual consumption level – particularly the case with smaller clusters and cities with fewer load profiles (Queenstown for example).

The following process was used to select the representative load profiles of each cluster. The various error parameters discussed were determined to provide sufficient samples, while also ensuring that the samples were sufficiently representative of each cluster (close to the cluster centres and target annual consumptions).

1. Find all load profiles within a defined distance of the cluster median.

   This distance was set at 6% of the cluster median for each of the three cluster parameters for all clusters, and 7% for Cluster 6. Naturally this resulted in fewer load profiles than the numbers in Table 7 representing each cluster, but load profiles closer to the cluster centres. The resulting numbers of load profiles representing each cluster are set out in Table 9.

2. For each of the target annual consumptions (8,000 kWh pa and 12,000 kWh pa), determine the distance of the actual consumption from the target as a proportion of the actual value.

   Likewise, determine the distance of each cluster parameter from the cluster median as a proportion of the cluster median value.

3. Take the mean sum of squares of the resulting four values for all load profiles summarised in Table 9.

4. Apply a further filter to ensure that the annual energy of each cluster sample is within 5% of the target annual energy (8,000 kWh pa and 12,000 kWh pa).

5. Apply a further filter to ensure that the maximum demand occurs in winter and minimum demand occurs in summer.

6. Following steps 3, 4, and 5, find and select the two load profiles with minimum sum of squared error values. If there is only one load profile per cluster from the above filtering process, select that. In some cases, there were no load profiles resulting.

Table 10 summarises the number of load profiles selected to represent each cluster, city, and 8,000 kWh pa. Table 11 summarises the number of load profiles selected to represent each cluster, city and 12,000 kWh pa.

Table 9: Load profiles in each cluster, and the number representing each city, prior to filtering further by annual consumption.

Cluster	Total number of load profiles selected	Number of load profiles selected from				Selection criteria applied to obtain selected load profiles (% from cluster median for each of the three cluster parameters)
		Auckland	Wellington	Christchurch	Queenstown	
0	546	152	140	236	18	6.0%
1	791	275	208	300	8	6.0%
2	205	85	48	68	4	6.0%
3	79	17	23	39	0	6.0%
4	815	226	208	355	26	6.0%
5	609	230	167	203	9	6.0%
6	39	0	4	34	1	7.0%
7	349	75	89	182	3	6.0%
Total	3,433	1,060	887	1,417	69	

Table 10: The number of load profiles selected to represent each cluster and city that are within 5% of 8,000 kWh per annum. The goal was two load profiles per city and cluster, giving 16 per city and 64 in total.

Cluster	Total number of load profiles selected	Number of load profiles selected from			
		Auckland	Wellington	Christchurch	Queenstown
0	7	2	2	2	1
1	7	2	2	2	1
2	5	2	2	1	0
3	4	2	2	0	0
4	7	2	2	2	1
5	7	2	2	2	1
6	2	0	0	2	0
7	7	1	2	2	2
Total	46	13	14	13	6

Table 11: The number of load profiles selected to represent each cluster and city that are within 5% of 12,000 kWh per annum. The goal was two load profiles per city and cluster, giving 16 per city and 64 in total.

Cluster	Total number of load profiles selected	Number of load profiles selected from			
		Auckland	Wellington	Christchurch	Queenstown
0	7	2	2	2	1
1	8	2	2	2	2
2	5	2	0	2	1
3	4	1	1	2	0
4	8	2	2	2	2
5	8	2	2	2	2
6	1	0	0	1	0
7	6	2	2	2	0
Total	47	13	11	15	8

7 Summary of the representative loads used in the study and their cluster parameter values

Table 12 and Table 13 list the load profiles selected to represent each cluster, city and consumption level (8,000 kWh pa in Table 12 and 12,000 kWh pa in Table 13).

Examples of 12,000 kWh pa load profiles selected to represent each cluster and city are given in

- Figure 35 to Figure 38 for Auckland,
- Figure 39 to Figure 42 for Wellington,
- Figure 43 to Figure 46 for Christchurch, and
- Figure 47 to Figure 50 for Queenstown.

All figures refer to the three clustering parameters using the same levels and descriptions as Table 8. The parameters themselves are shortened, with the full parameter descriptions given in Table 6. In some instances, 12,000 kWh pa samples were unavailable. In these cases, and where available, 8,000 kWh pa samples have been substituted in. Otherwise, the load profile is omitted with the caption noting why. It is clear from these, and Table 10 and Table 11, that Clusters 3 and 6 are not representative of all cities. Indeed, Cluster 6 appears to only represent Christchurch, which may be due to Orion's historically strong day-night price differentiation, or simply that Christchurch had more load profile samples available. Cluster 3 also appears to mainly represent all cities except Queenstown, most likely due to there being fewer load profiles available from Queenstown.

When viewing these load profiles, recall that the clustering parameters are ratios of load within time periods to load in other time periods, and within the load profile itself. Clusters were not selected based on annual consumption – that was selected for after clustering. Therefore, comparing absolute power levels between graphs in terms of cluster selection is not possible. Instead, they should be compared within time periods within each individual load profile.

Table 12: Load profile selected to represent each cluster, city, and 8,000 kWh pa consumption. Those with blanks in the annual consumption and cluster parameters did not have load profiles within the distances specified previously from the annual consumption values and cluster centres.

City	Cluster	Order	Target kWh	Actual annual consumption (kWh)	Distance from cluster centre			Distance from target kWh
					Day average to all day	Average weekday morning peak load to average load	Average weekday evening peak load to average load	
Auckland	0	1	8,000	7,903	-0.8%	1.3%	0.7%	-1.2%
Auckland	0	2	8,000	8,295	0.9%	-1.6%	3.1%	3.7%
Wellington	0	1	8,000	8,146	-3.7%	-0.9%	-0.5%	1.8%
Wellington	0	2	8,000	7,860	-2.2%	-1.3%	3.3%	-1.7%
Christchurch	0	1	8,000	7,962	-3.1%	-0.4%	-3.5%	-0.5%
Christchurch	0	2	8,000	7,736	-0.4%	3.0%	-1.6%	-3.3%
Queenstown	0	1	8,000	7,633	1.9%	0.7%	-2.5%	-4.6%
Queenstown	0	2	8,000					
Auckland	1	1	8,000	8,160	-0.1%	-1.4%	3.1%	2.0%
Auckland	1	2	8,000	7,994	-0.4%	-4.1%	0.4%	-0.1%
Wellington	1	1	8,000	8,055	0.7%	-0.5%	0.6%	0.7%
Wellington	1	2	8,000	7,995	-1.5%	-3.5%	-0.8%	-0.1%
Christchurch	1	1	8,000	7,980	0.1%	-2.5%	0.2%	-0.3%
Christchurch	1	2	8,000	8,052	-2.6%	1.1%	-2.2%	0.6%
Queenstown	1	1	8,000	8,236	-0.1%	0.2%	-3.6%	3.0%
Queenstown	1	2	8,000					
Auckland	2	1	8,000	8,144	0.5%	-0.2%	-3.3%	1.8%
Auckland	2	2	8,000	8,402	-0.3%	-2.0%	1.2%	5.0%
Wellington	2	1	8,000	8,087	2.9%	-0.5%	3.8%	1.1%
Wellington	2	2	8,000	8,312	-2.2%	4.8%	4.3%	3.9%
Christchurch	2	1	8,000	7,624	-5.8%	0.1%	1.3%	-4.7%
Christchurch	2	2	8,000					
Queenstown	2	1	8,000					
Queenstown	2	2	8,000					
Auckland	3	1	8,000	8,092	5.3%	-4.4%	2.3%	1.2%
Auckland	3	2	8,000	7,892	-3.8%	5.6%	-3.9%	-1.4%
Wellington	3	1	8,000	7,963	-3.0%	3.2%	-0.9%	-0.5%
Wellington	3	2	8,000	8,291	5.9%	-4.6%	3.5%	3.6%
Christchurch	3	1	8,000					
Christchurch	3	2	8,000					
Queenstown	3	1	8,000					
Queenstown	3	2	8,000					
Auckland	4	1	8,000	8,006	-3.7%	0.3%	-0.5%	0.1%
Auckland	4	2	8,000	7,829	-2.7%	-0.4%	-2.1%	-2.1%
Wellington	4	1	8,000	8,163	-1.3%	-1.6%	2.1%	2.0%
Wellington	4	2	8,000	8,193	0.4%	-4.7%	-0.6%	2.4%
Christchurch	4	1	8,000	8,053	-2.4%	-0.2%	1.4%	0.7%
Christchurch	4	2	8,000	7,907	-1.0%	1.9%	2.0%	-1.2%
Queenstown	4	1	8,000	8,397	1.2%	-4.4%	5.1%	5.0%
Queenstown	4	2	8,000					
Auckland	5	1	8,000	7,981	2.6%	3.8%	3.3%	-0.2%
Auckland	5	2	8,000	8,011	-0.7%	-2.6%	-5.0%	0.1%
Wellington	5	1	8,000	8,106	-1.8%	2.1%	2.0%	1.3%
Wellington	5	2	8,000	8,127	-2.6%	-2.4%	2.1%	1.6%
Christchurch	5	1	8,000	8,039	-0.9%	-0.3%	0.9%	0.5%
Christchurch	5	2	8,000	7,825	-2.6%	0.8%	-2.3%	-2.2%
Queenstown	5	1	8,000	8,386	-3.4%	0.0%	2.5%	4.8%
Queenstown	5	2	8,000					
Auckland	6	1	8,000					
Auckland	6	2	8,000					
Wellington	6	1	8,000					
Wellington	6	2	8,000					
Christchurch	6	1	8,000	8,324	-0.3%	4.5%	4.8%	4.0%
Christchurch	6	2	8,000	7,931	-3.4%	-1.8%	7.2%	-0.9%
Queenstown	6	1	8,000					
Queenstown	6	2	8,000					
Auckland	7	1	8,000	7,922	5.8%	4.8%	2.2%	-1.0%
Auckland	7	2	8,000					
Wellington	7	1	8,000	7,916	1.2%	3.8%	-2.9%	-1.0%
Wellington	7	2	8,000	7,953	-2.3%	4.0%	-2.3%	-0.6%
Christchurch	7	1	8,000	8,039	-0.7%	0.4%	-1.6%	0.5%
Christchurch	7	2	8,000	8,186	1.6%	-0.1%	2.7%	2.3%
Queenstown	7	1	8,000	8,242	1.0%	3.5%	-4.4%	3.0%
Queenstown	7	2	8,000	7,858	3.3%	-3.8%	-5.4%	-1.8%

Table 13: Load profile selected to represent each cluster, city, and 12,000 kWh pa consumption. Those with blanks in the annual consumption and cluster parameters did not have load profiles within the distances specified previously from the annual consumption values and cluster centres.

City	Cluster	Order	Target kWh	Actual annual consumption (kWh)	Distance from cluster centre			Distance from target kWh
					Day average to all day	Average weekday morning peak load to average load	Average weekday evening peak load to average load	
Auckland	0	1	12,000	11,931	-1.7%	-3.7%	-1.0%	-0.6%
Auckland	0	2	12,000	11,504	1.3%	0.6%	0.6%	-4.1%
Wellington	0	1	12,000	11,982	3.6%	1.7%	0.8%	-0.1%
Wellington	0	2	12,000	11,914	-0.9%	-4.4%	2.8%	-0.7%
Christchurch	0	1	12,000	12,439	1.6%	0.5%	0.0%	3.7%
Christchurch	0	2	12,000	11,970	1.6%	3.6%	-1.2%	-0.3%
Queenstown	0	1	12,000	11,715	-2.3%	2.5%	5.3%	-2.4%
Queenstown	0	2	12,000					
Auckland	1	1	12,000	11,605	-1.3%	-0.7%	-0.1%	-3.3%
Auckland	1	2	12,000	11,933	-1.0%	-3.1%	2.2%	-0.6%
Wellington	1	1	12,000	11,544	1.8%	0.1%	-0.1%	-3.8%
Wellington	1	2	12,000	11,996	-1.9%	-2.5%	-3.1%	0.0%
Christchurch	1	1	12,000	12,053	-2.9%	0.7%	-3.0%	0.4%
Christchurch	1	2	12,000	12,256	0.4%	-4.6%	0.4%	2.1%
Queenstown	1	1	12,000	12,379	-2.2%	4.0%	2.0%	3.2%
Queenstown	1	2	12,000	12,631	-1.2%	6.0%	2.6%	5.3%
Auckland	2	1	12,000	12,041	1.0%	-1.0%	-3.0%	0.3%
Auckland	2	2	12,000	12,016	-5.1%	-4.1%	-5.4%	0.1%
Wellington	2	1	12,000					
Wellington	2	2	12,000					
Christchurch	2	1	12,000	12,028	-5.5%	-4.4%	-2.1%	0.2%
Christchurch	2	2	12,000	11,451	-2.7%	-5.9%	4.2%	-4.6%
Queenstown	2	1	12,000	12,425	3.1%	4.2%	-3.9%	3.5%
Queenstown	2	2	12,000					
Auckland	3	1	12,000	12,598	1.2%	0.0%	1.2%	5.0%
Auckland	3	2	12,000					
Wellington	3	1	12,000	12,552	2.4%	3.6%	-3.7%	4.6%
Wellington	3	2	12,000					
Christchurch	3	1	12,000	11,743	2.0%	-1.9%	-2.7%	-2.1%
Christchurch	3	2	12,000	11,617	-2.6%	4.4%	-3.2%	-3.2%
Queenstown	3	1	12,000					
Queenstown	3	2	12,000					
Auckland	4	1	12,000	11,887	-3.1%	1.5%	0.1%	-0.9%
Auckland	4	2	12,000	12,198	-1.4%	3.7%	-0.6%	1.7%
Wellington	4	1	12,000	11,937	2.0%	1.8%	-2.7%	-0.5%
Wellington	4	2	12,000	11,697	-0.8%	1.4%	2.7%	-2.5%
Christchurch	4	1	12,000	12,447	-2.3%	0.1%	-2.7%	3.7%
Christchurch	4	2	12,000	12,093	-0.1%	-4.1%	3.2%	0.8%
Queenstown	4	1	12,000	11,773	-0.3%	2.4%	2.5%	-1.9%
Queenstown	4	2	12,000	11,989	-0.3%	-3.5%	-2.1%	-0.1%
Auckland	5	1	12,000	12,236	-2.0%	-0.7%	-2.9%	2.0%
Auckland	5	2	12,000	11,588	2.9%	1.9%	2.6%	-3.4%
Wellington	5	1	12,000	12,258	-1.2%	0.6%	0.1%	2.2%
Wellington	5	2	12,000	12,415	-1.0%	0.4%	2.0%	3.5%
Christchurch	5	1	12,000	11,910	-2.1%	2.0%	-1.2%	-0.8%
Christchurch	5	2	12,000	12,385	2.1%	1.8%	0.2%	3.2%
Queenstown	5	1	12,000	11,758	2.7%	5.4%	0.0%	-2.0%
Queenstown	5	2	12,000	11,795	-5.4%	6.1%	-5.8%	-1.7%
Auckland	6	1	12,000					
Auckland	6	2	12,000					
Wellington	6	1	12,000					
Wellington	6	2	12,000					
Christchurch	6	1	12,000	12,319	-1.4%	6.6%	3.7%	2.7%
Christchurch	6	2	12,000					
Queenstown	6	1	12,000					
Queenstown	6	2	12,000					
Auckland	7	1	12,000	12,184	5.6%	-2.1%	-0.3%	1.5%
Auckland	7	2	12,000	11,867	-3.7%	-4.3%	2.7%	-1.1%
Wellington	7	1	12,000	12,480	0.3%	4.2%	-2.4%	4.0%
Wellington	7	2	12,000	11,837	-4.0%	1.8%	-5.6%	-1.4%
Christchurch	7	1	12,000	12,057	-0.9%	-1.1%	-2.8%	0.5%
Christchurch	7	2	12,000	12,554	0.1%	-0.8%	0.7%	4.6%
Queenstown	7	1	12,000					
Queenstown	7	2	12,000					

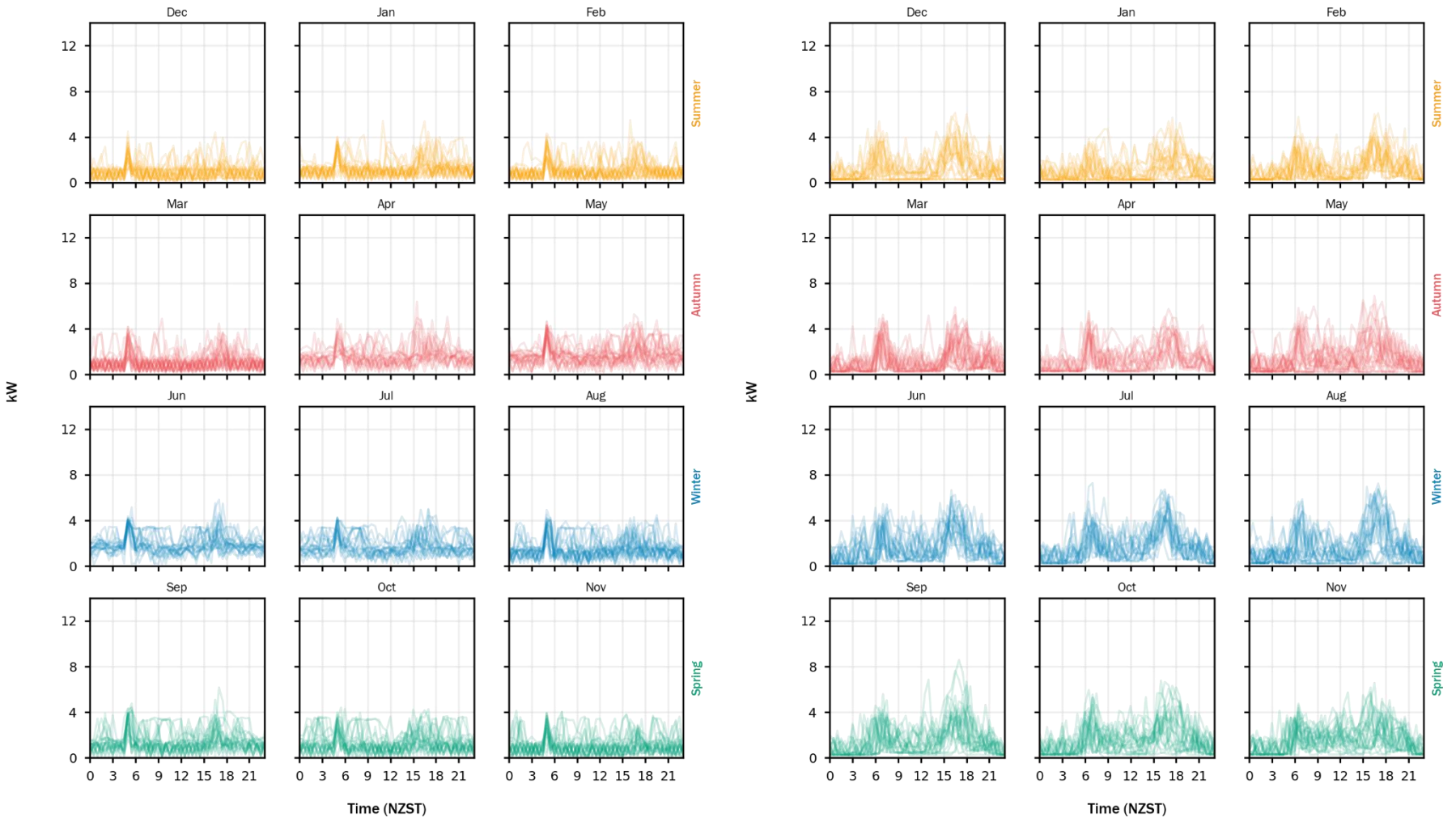


Figure 35: Auckland Cluster 0 (medium daytime & medium morning & medium evening) and Cluster 1 (very high daytime & high morning & very high evening).

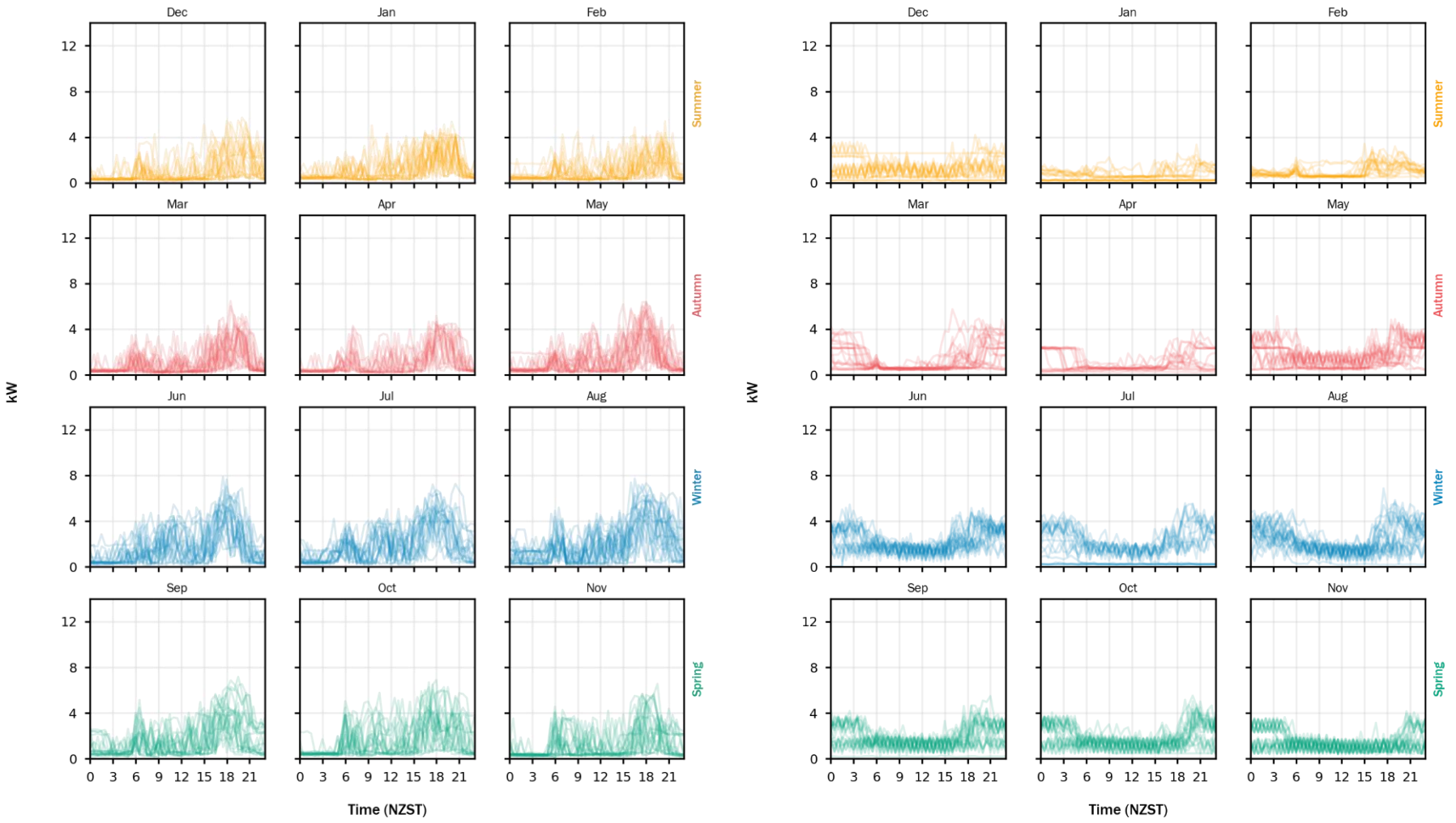


Figure 36: Auckland Cluster 2 (very high daytime & low morning & very high evening) and Cluster 3 (low daytime & low morning & medium evening).

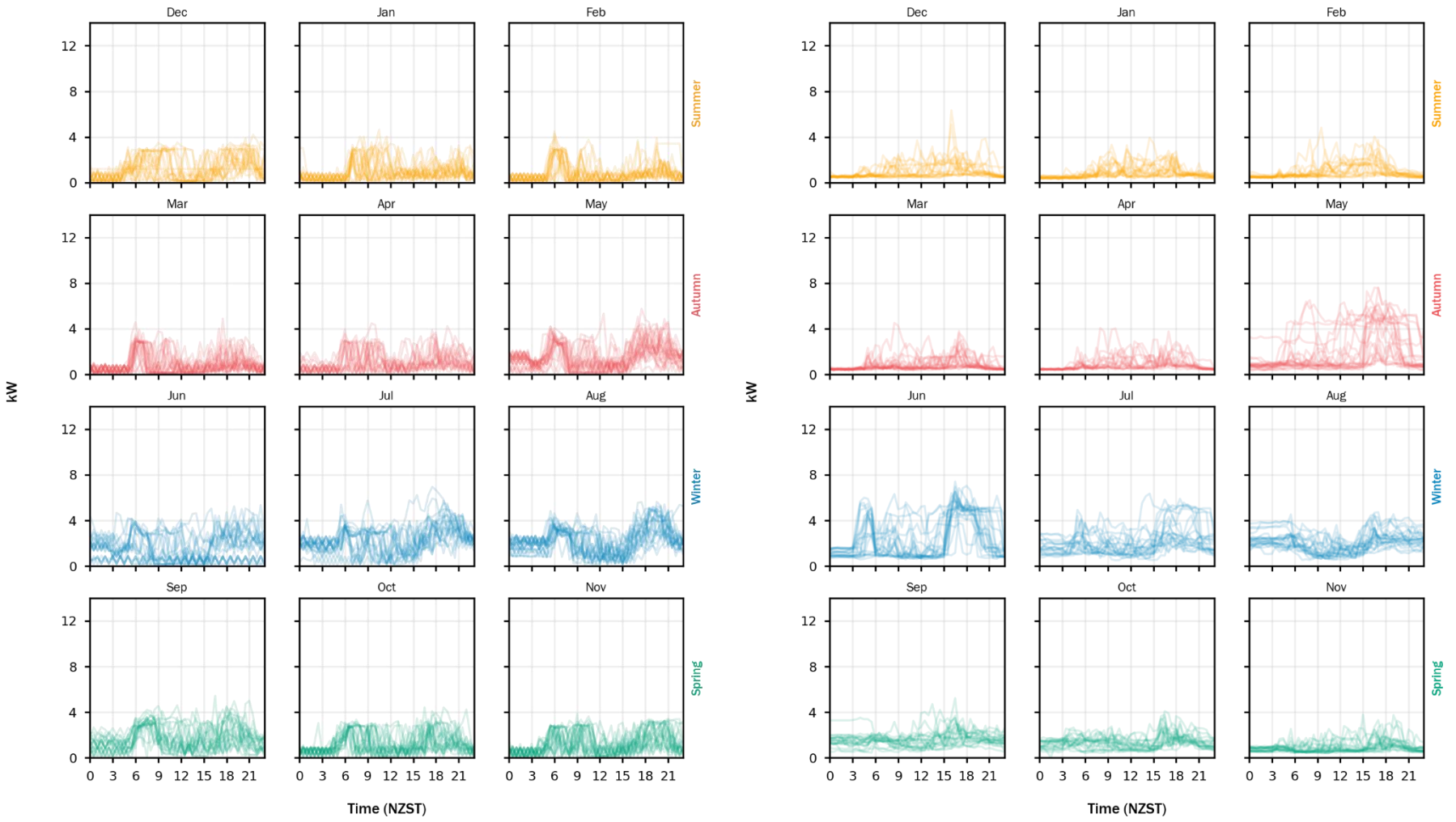


Figure 37: Auckland Cluster 4 (high daytime & high morning & high evening) and Cluster 5 (high daytime & low morning & very high evening).

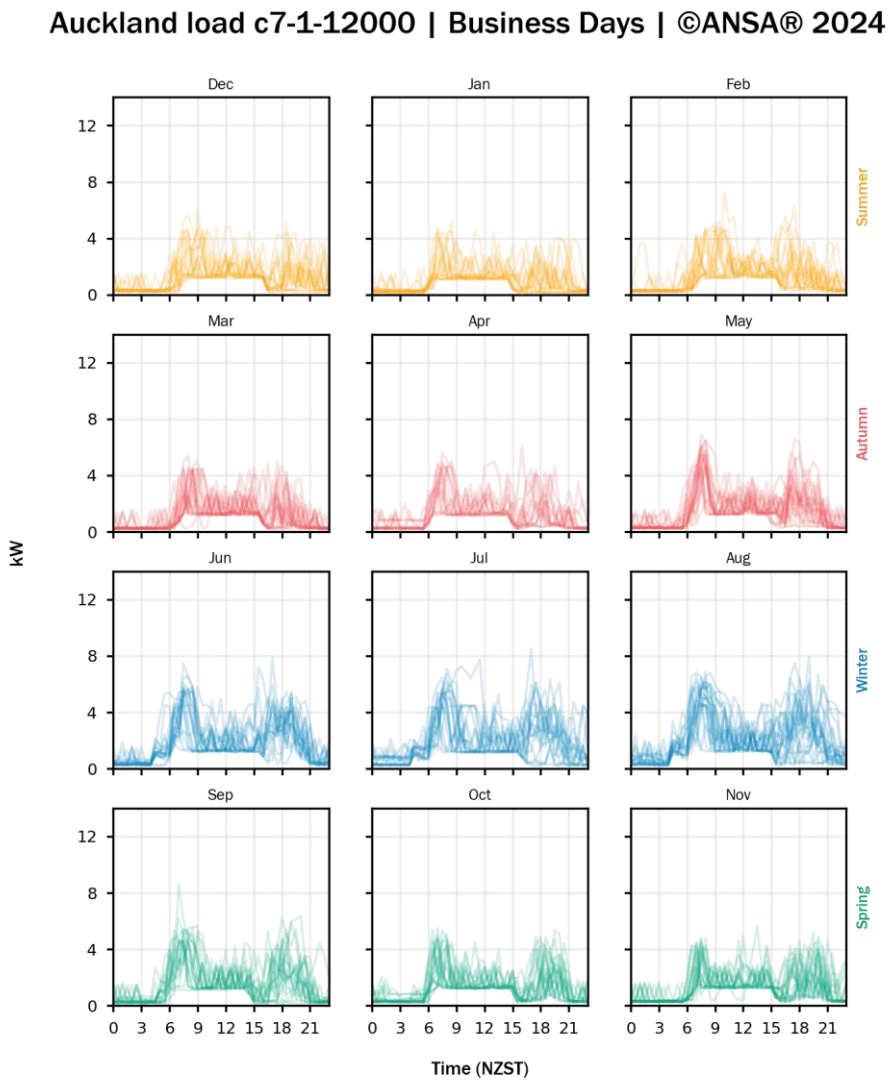


Figure 38: Auckland Cluster 7 (very high daytime & very high morning & high evening). Neither an 8,000 kWh pa nor a 12,000 kWh pa Cluster 6 sample was available.

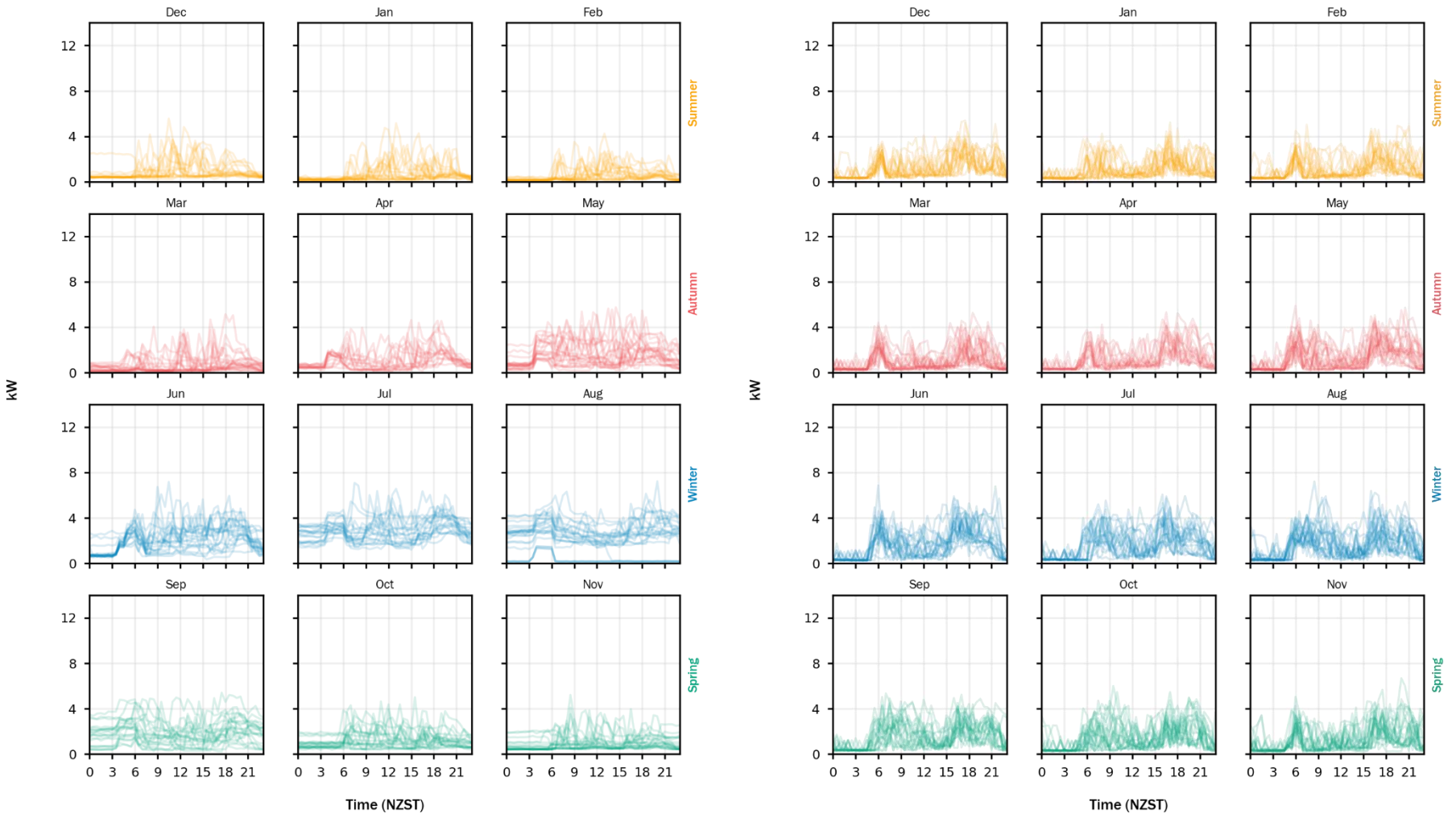


Figure 39: Wellington Cluster 0 (medium daytime & medium morning & medium evening) and Cluster 1 (very high daytime & high morning & very high evening).

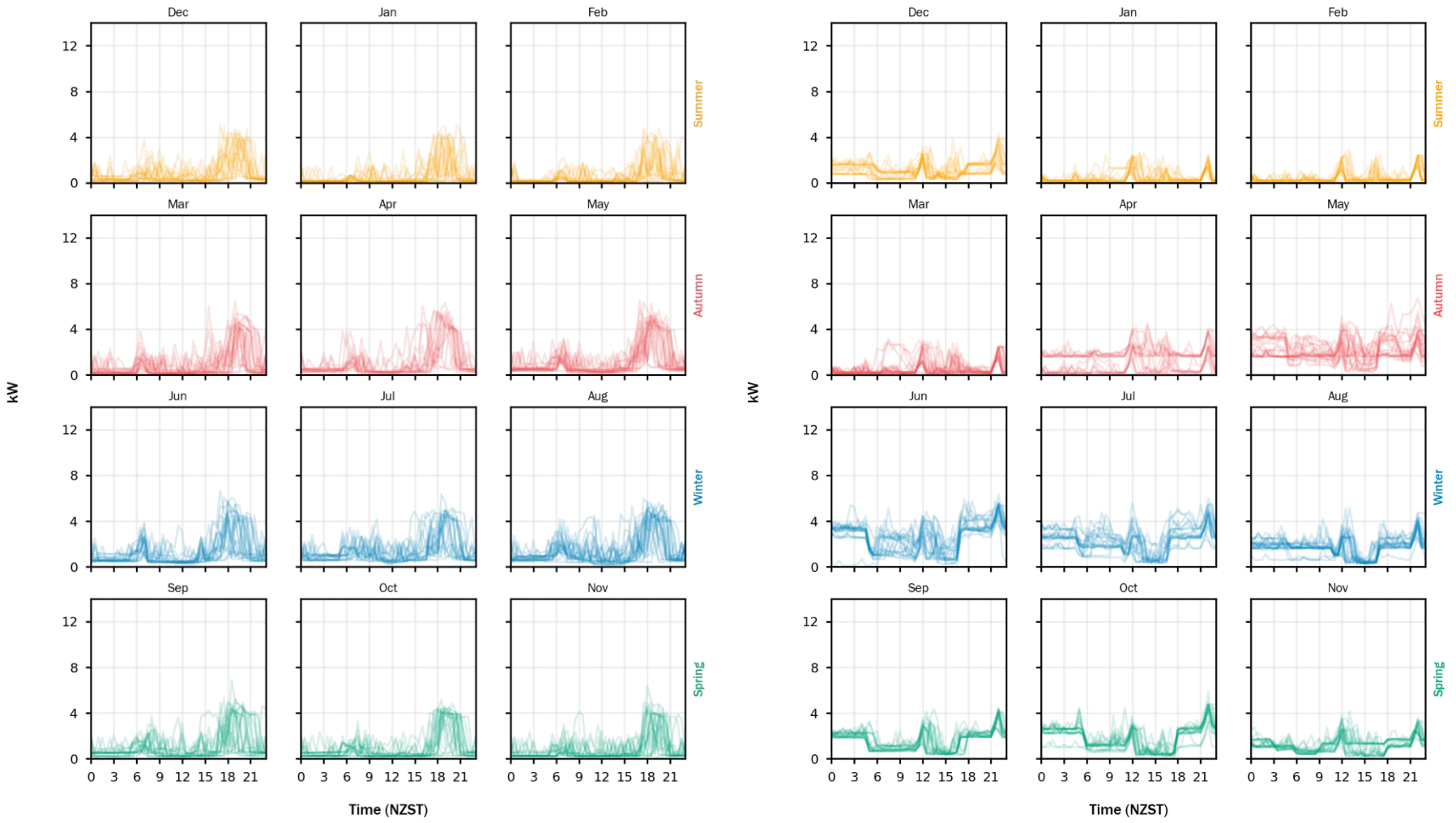


Figure 40: Wellington Cluster 2 (very high daytime & low morning & very high evening) 8,000 kWh pa substituted and Cluster 3 (low daytime & low morning & medium evening). No 12,000 kWh pa Cluster 2 sample was available.

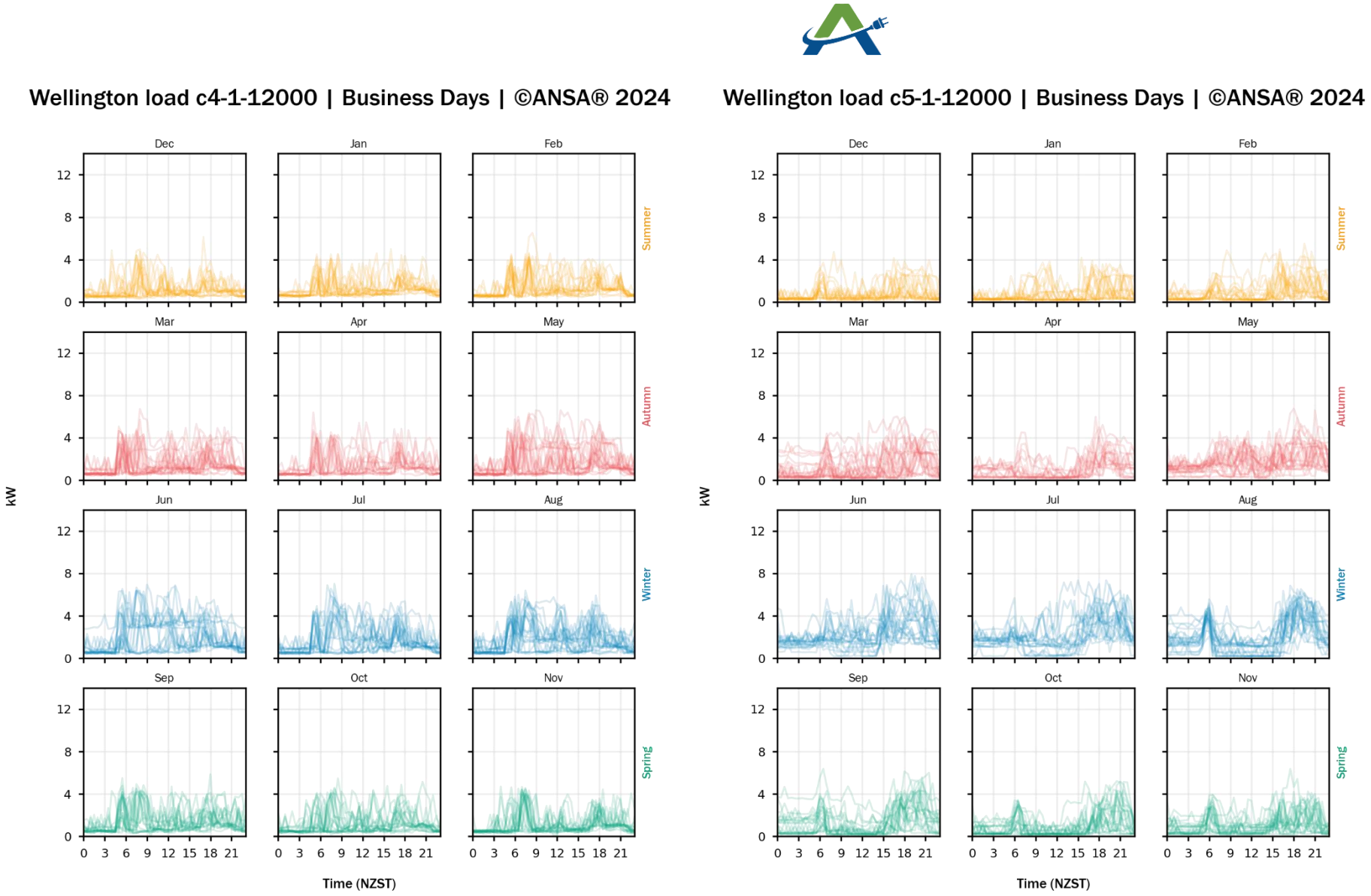


Figure 41: Wellington Cluster 4 (high daytime & high morning & high evening) and Cluster 5 (high daytime & low morning & very high evening).

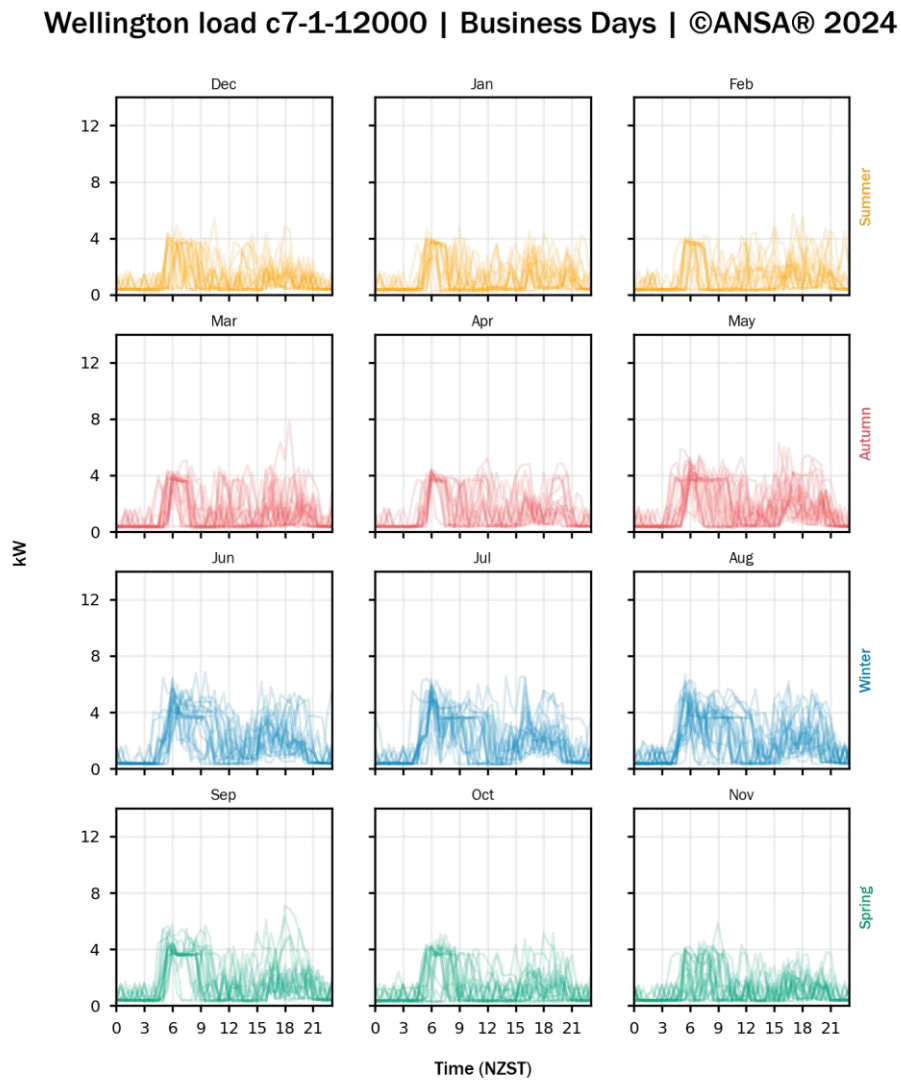


Figure 42: Wellington Cluster 7 (very high daytime & very high morning & high evening). No Cluster 6 sample was available.

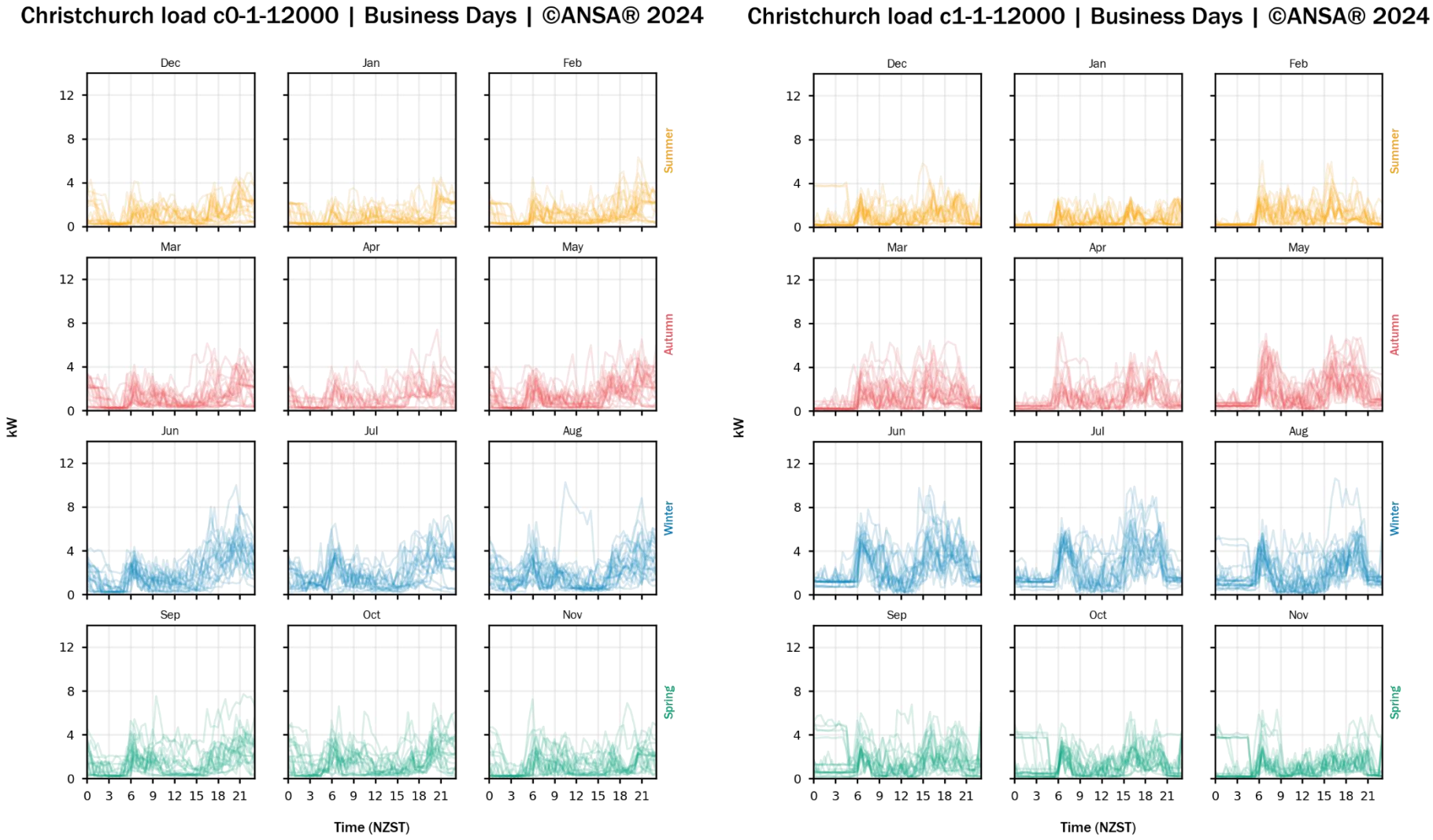


Figure 43: Christchurch Cluster 0 (medium daytime & medium morning & medium evening) and Cluster 1 (very high daytime & high morning & very high evening).

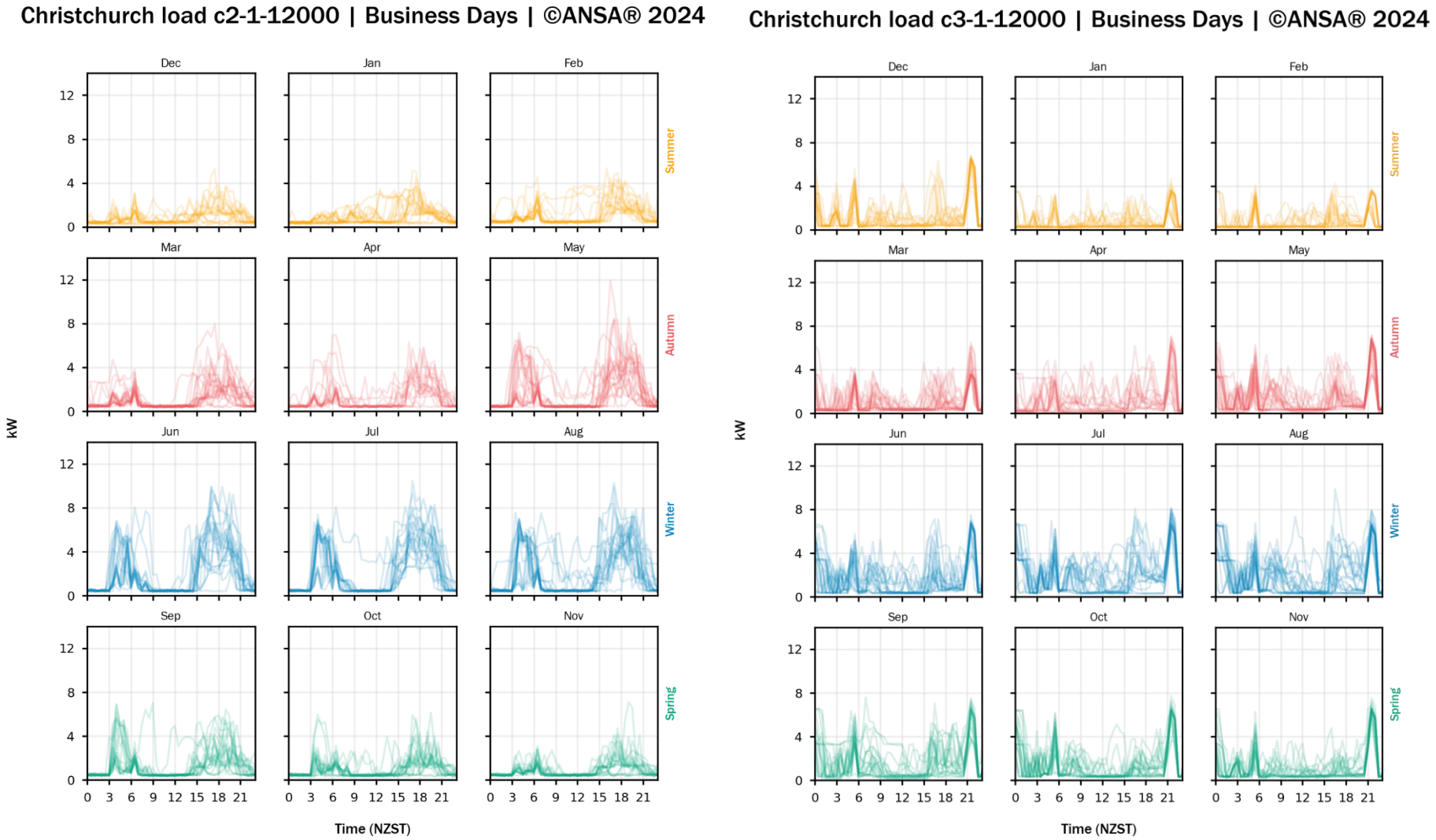


Figure 44: Christchurch Cluster 2 (very high daytime & low morning & very high evening) and Cluster 3 (low daytime & low morning & medium evening).

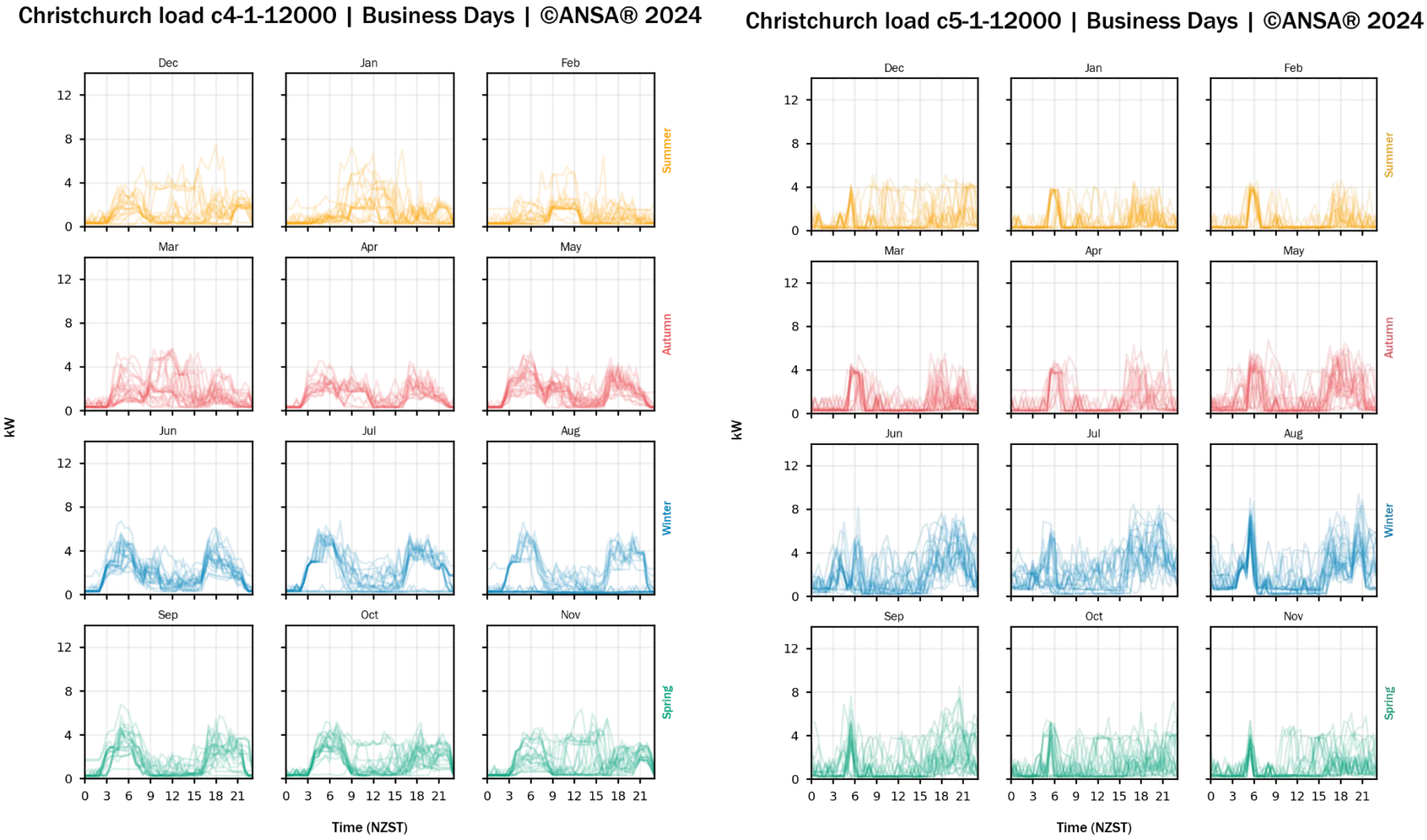


Figure 45: Christchurch Cluster 4 (high daytime & high morning & high evening) and Cluster 5 (high daytime & low morning & very high evening).

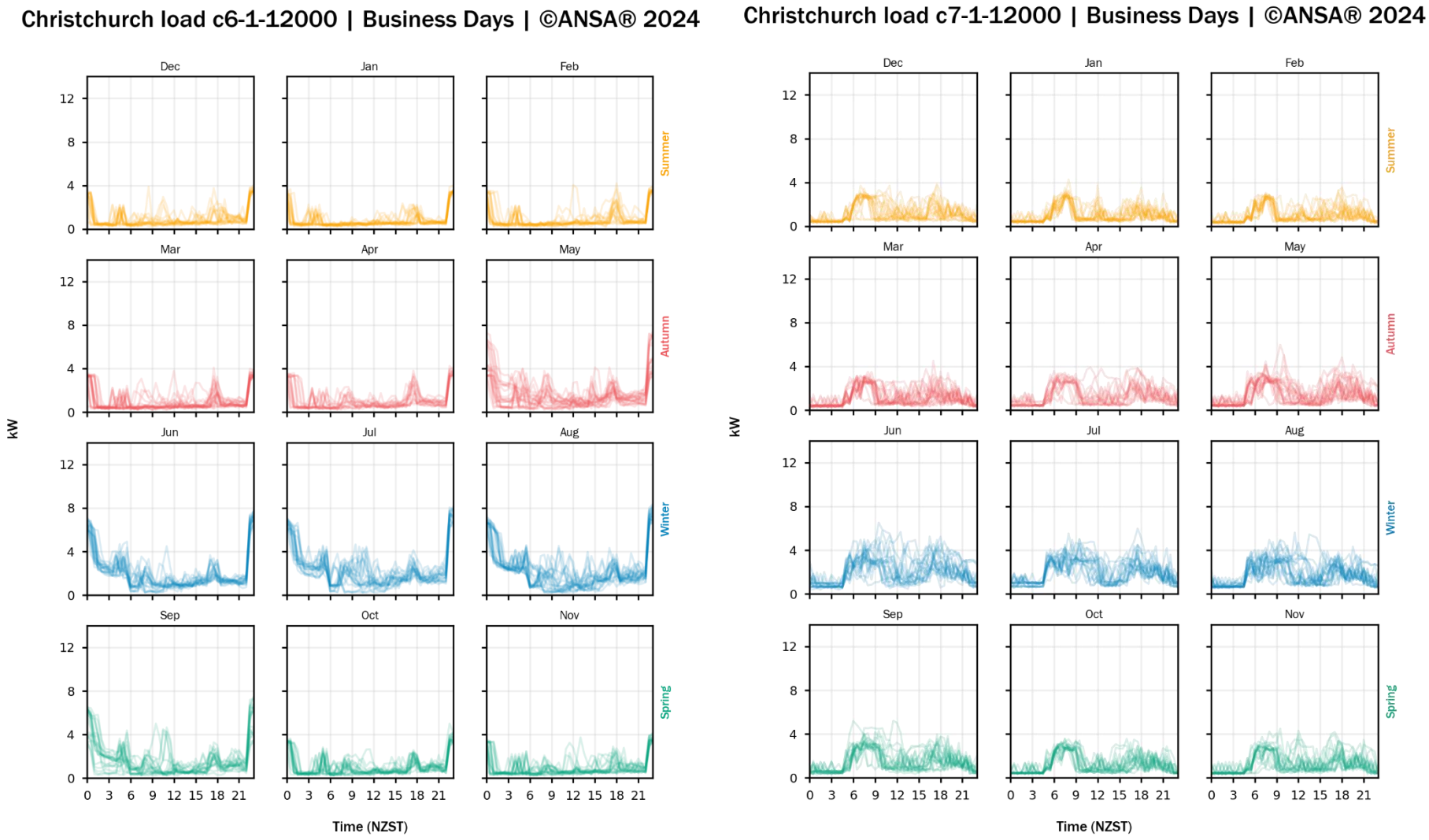


Figure 46: Christchurch Cluster 6 (very low daytime & very low morning & low evening) and Cluster 7 (very high daytime & very high morning & high evening).

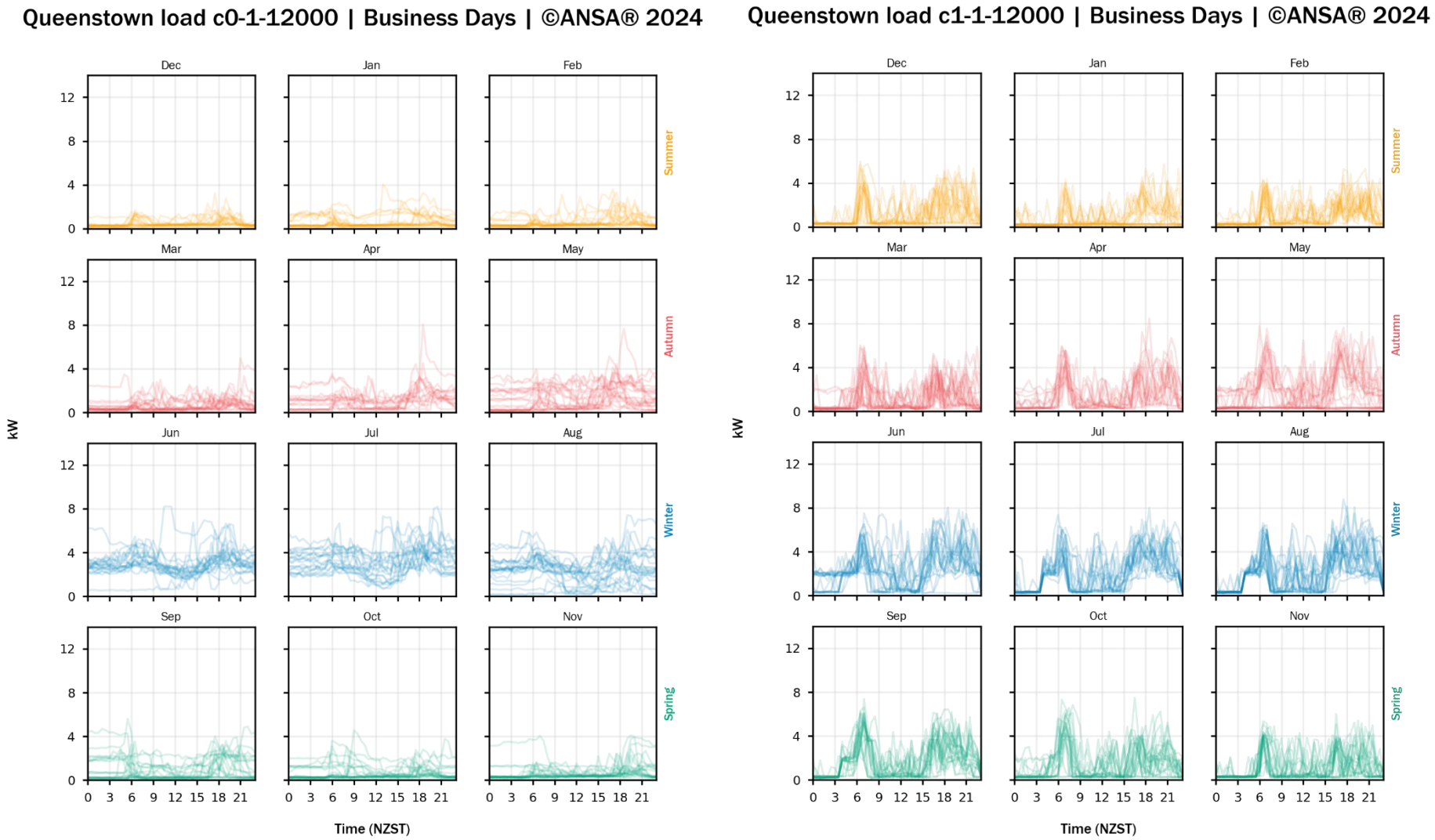


Figure 47: Queenstown Cluster 0 (medium daytime & medium morning & medium evening) and Cluster 1 (very high daytime & high morning & very high evening).

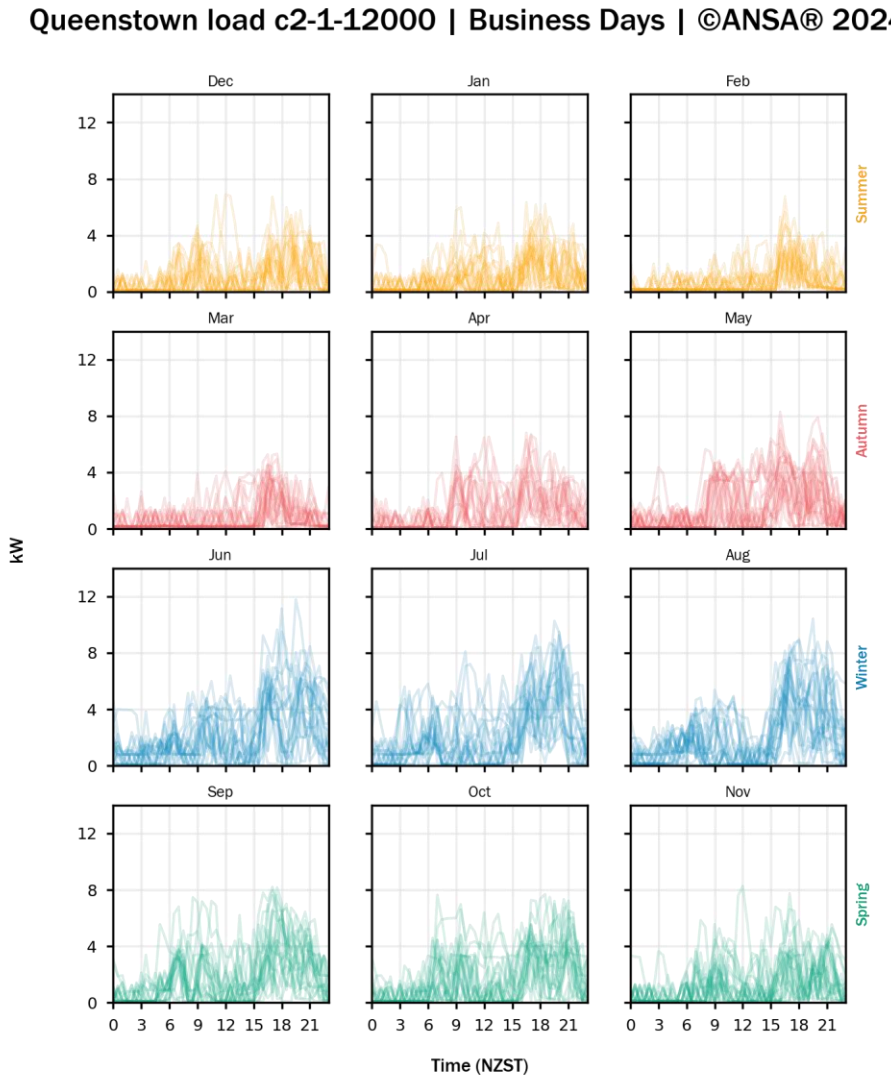


Figure 48: Queenstown Cluster 2 (very high daytime & low morning & very high evening). No Cluster 3 sample was available.

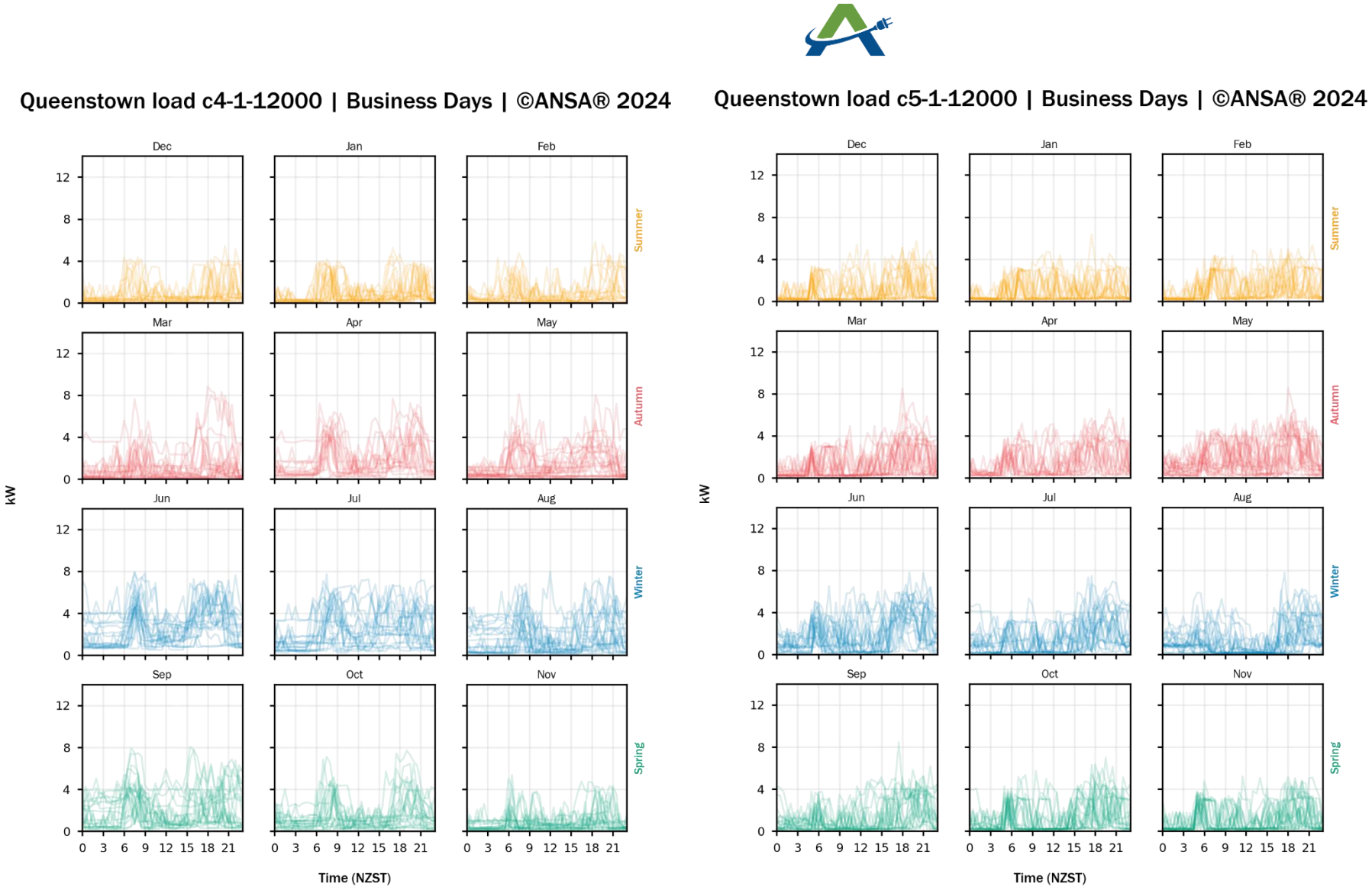


Figure 49: Queenstown Cluster 4 (high daytime & high morning & high evening) and Cluster 5 (high daytime & low morning & very high evening).

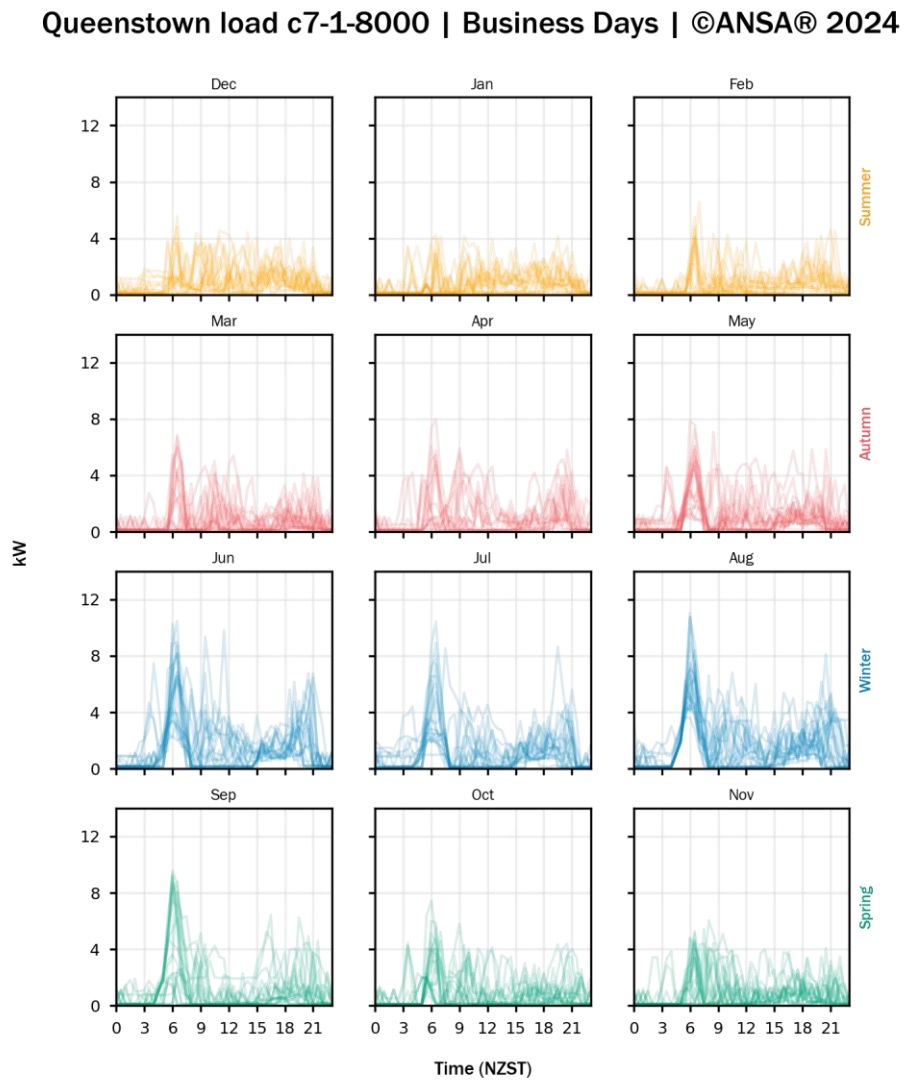


Figure 50: Queenstown Cluster 7 (very high daytime & very high morning & high evening). Only an 8,000 kWh pa Cluster 7 sample was available, and no Cluster 6 sample was available.