Synoptic weather regimes over Aoteaora New Zealand

Article

Published Version

Synoptic weather regimes over Aotearoa New Zealand

J. Williams[1] and J.Renwick[2]

[1]*NIWA-Taihoro Nukurangi, 301 Evans Bay Parade, Ākautangi, Hātaitai, Wellington, Aotearoa New Zealand*
[2]*School of Geography, Environment and Earth Sciences, Victoria University of Wellington - Te Herenga Waka, Kelburn, Wellington, Aotearoa New Zealand*

Correspondence: jonny.williams@niwa.co.nz

Key words: synoptic, regimes, clustering, reanalysis, Python.

This paper is dedicated to the memory of Dr Brett Mullan and to his contribution to weather and climate research in Aotearoa New Zealand.

Abstract

This work provides an updated set of 12 dominant geopotential height fields over Aotearoa New Zealand defined using NCEP/NCAR reanalyses and provides an initial analysis of the same using a state of the art dataset with increased temporal and spatial resolution; ERA5. These regimes were initially produced by Kidson (2000) and have provided the basis for many other subsequent studies. These maps provide a guide to the prevailing weather due to the strong relationships between circulation patterns and surface climate in New Zealand. The results presented here using the NCEP/NCAR reanalysis are broadly in agreement with previous work but with some important differences. The most notable of these differences is the need to average two regimes together to provide good agreement between this work and Kidson (2000). These differences are attributed to differences in software used, improvements to the underlying dataset itself and to the 'mixing' of statistically indistinguishable empirical orthogonal functions in different linear combinations. Using ERA5 data results in sets of weather types which are analogous in some ways to previous work but with some important differences, especially over mountainous regions. Use of ERA5 noticeably improves the ratio of intra- to inter-regime variance; a measure of the 'quality' of the cluster analyses. All data and code used in this work is publicly accessible and it is hoped that this will provide a catalyst for open discussions on this topic, particularly with relation to future perturbations to these regimes under climate change. The mathematical methods used in this study are widely used in the New Zealand weather and climate community and it is hoped that this work will prove useful as a resource for future studies.

1. Introduction

The world's meteorology and climatology is extraordinarily complex in its spatio-temporal variability. Because of this, simplifications and methodologies for reducing and communicating this complexity are essential if we are to understand it better. Examples are many and various; e.g. the Beaufort scale in wind speed, rainfall probability in a weather forecast, cloud clustering in climate models (Williams and Webb 2009), classification of El Niño or La Niña (e.g. Trenberth 1997 for a review) and the ubiquitous quasi-static approximation.

Submitted: 19 April 2021
Accepted: 8 October 2021

Although these methodologies are of course hugely different in their scope, they all have one fundamental thing in common; that of the reduction in the complexity of a system such that it can be better understood by the relevant audience. This study is limited only to the region surrounding New Zealand and to one meteorological variable, the 1000hPa geopotential height, *z*. Initially we use the data at 0000 and 1200 UTC as in Kidson (2000) before moving on to using the ERA5 dataset's results at 0000, 0600, 1200 and 1800 UTC.

The literature contains several examples of synoptic classification methods over New Zealand in addition to the most well known example: the 12 'Kidson types', which form the basis of the first part of this study. For example, in Jiang et al. (2004) and Jiang (2011) the authors use obliquely rotated T-mode principal component analysis and convergent K-means clustering to find 10 and 12 dominant weather regimes respectively. These regimes are analogous to those found in Kidson (2000) - hereafter K2K - but, as the author states, are not 'one to one'. Indeed, this foreshadows the work presented here by noting that '*... the convergent K-means clustering (in Jiang et al. 2004) did not change the (K2K) cluster centroid locations significantly, but redistributed the memberships of daily maps to form compact clusters.*'

Other authors have used self-organising maps - SOMs - for example as in Jiang et al. (2013) and Gibson et al. (2016). It is beyond the scope of this work to give a detailed explanation of the SOM generation procedure, but in brief it is a machine learning technique which produces two-dimensional representations - i.e. maps - from a higher dimensional input dataset. The aim is to preserve the essential structure of the input data whilst allowing it to be more easily interpreted. In this qualitative respect it is therefore analogous to the 'PCA + K-means' clustering approach of K2K. However, a key difference in the approaches is that SOMs add a means of topology conservation whereby the cluster means dynamically adjust - i.e. 'self organise' - throughout the process (Sheridan and Lee 2011).

There are two main motivations for this work; firstly to provide a pedagogical primer for how the widely used 'PCA + K-means' method is implemented in practice, and secondly to apply the method to the ERA5 dataset (Hersbach et al., 2020). ERA5 has 10 times the grid resolution of NCEP/NCAR (0.25° versus 2.5°) in addition to increased temporal resolution and is thus capable of a far more detailed analysis. It is, however, advantageous to use the same methodology as K2K here since it provides traceability with many previous studies and is implemented in not more than a few 10s of lines of Python code using modern software packages. We note however that methods such as SOMs and other machine learning approaches - for example Saavedra-Moreno et al. (2015) - are more rigorous in their derivations and better able to represent extremes. It is important to state at this point however that is it not the intention of this work to recommend a particular methodology for regime classification.

We have deliberately chosen to use the same variable - 1000hPa geopotential height - and domain of interest for the ERA5 analysis to maintain commonality with K2K. Other authors use clustering metrics at 850hPa (Sheridan et al., 2008), 700hPa (Michelangeli et al., 1995) and 500hPa (Moore and Dixon 2015) in their analyses. The use of variables further away from the surface has the advantage that the synoptic-scale circulation is less impacted by surface features. This is just one example of a potential extension to this work which we hope to motivate. We are not advocating that this method is superior to other methods such as SOMs, however it does have the advantage of being widely known, particularly in the New Zealand community.

The dataset used in the first part of this work is identical to that used in K2K, which uses 28,852 fields of *z* between

January 1958 and June 1997 to define 12 dominant synoptic weather types. These are frequently referred to as 'Kidson types' in their wide use in modern- (Parsons et al. 2014) and paleo-climate (Ackerley et al. 2011) studies. Similar methods have also been used to study weather regimes in other parts of the world such as South America (Solman and Menéndez 2003).

The synoptic weather types found in K2K grouped the tens of thousands of input data points into 12 types. These are further split into 3 regimes; 'trough', 'zonal' and 'blocking'. Together, these regimes express the dominant weather types over New Zealand. This is because of New Zealand's topography and exposure to strong wind streams, resulting in strong statistical relationships between synoptic-scale flows and surface climate.

It should be noted that the second level classification of the clusters into the three larger 'trough', 'zonal' and 'blocking' types is, by its nature, subjective and in our analysis of the NCEP/NCAR dataset here, we use the same overarching terms purely to promote ease of comparison. When we move on to the calculation of new clusters for the ERA5 data, we choose to dispense with these terms altogether and simply order them by their areal mean height. This is found by calculating an area-weighted mean of the final height clusters.

This study gives a more detailed account of the derivation of these synoptic types than given previously and provides an update to the types' occurrences compared to K2K, which itself builds on many other previous studies (Kidson 1999, Kidson 1997, Kidson 1994a, Kidson 1994b, Kidson and Watterson 1995 and Ward 1963). The new results presented are broadly in agreement with those from K2K, however there are some notable differences.

The K2K methodology is widely cited, however the available details on the types' calculation are somewhat opaque, particularly to those readers less familiar with statistical analysis and clustering techniques. To aid future work, the code used is freely available and a step by step guide is given below for:

- How the weather types themselves are calculated from a reference dataset, in this case the NCEP/NCAR reanalysis.

- How to assign a particular synoptic type to a new observation or model output of z.

- How to interpret the meaning and derivation of the types from mathematical and geometrical arguments.

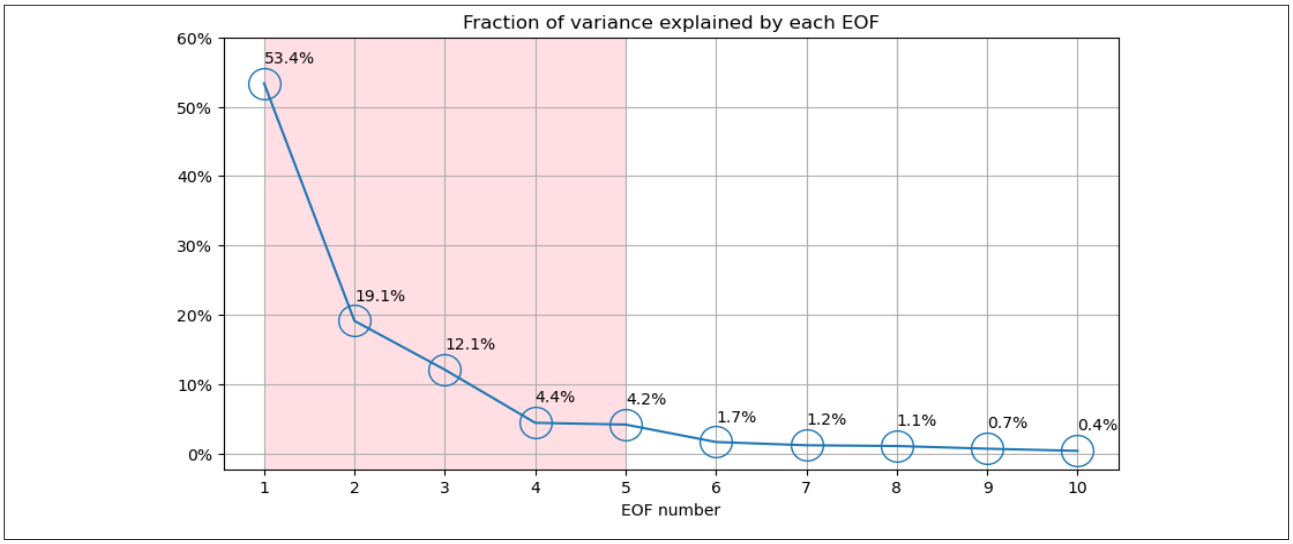

Figure 1: Fraction of variance explained by the first 5 EOFs in the time series of z_s. Although the first 10 EOFs are shown, only 5 are used in this analysis (shown by the shaded region).

2. Methodology

The most widely referenced work in the literature on this subject is K2K. This builds on earlier work (e.g. Kidson 1997 and Kidson 1994a) to construct 12 dominant synoptic weather regimes. This section gives a step-by-step guide to reproducing these weather types, i.e. Figure 2 in Kidson 2000 (and Figure 1 in Ackerley et al. 2011).

2.1 Data and software

We use the same input data as used in K2K, that is 1000hPa geopotential height data (z) from the NCEP/NCAR reanlysis (Kalnay et al. 1996) for January 1958 to June 1997 inclusive at 0000 and 1200 UTC.

We use the Python programming language exclusively for this work and make use of the open source eofs package (Dawson and Wales 2019, Dawson 2016) to calculate the principal components (PCs) and empirical orthogonal functions (EOFs). K-means clustering analysis is then carried out on the PCs to obtain the dominant weather regimes using the scikit-learn package (Pedregosa et al. 2011), which relies on NumPy (Harris et al. 2020) and SciPy (Virtanen et al. 2020) for its underlying operation and includes many more functions other than K-means clustering. A more detailed explanation of the mathematical basis of these packages and methods is beyond the scope of this paper.

2.2 Mathematical basis

Firstly, height anomalies are calculated by removing the time mean of the heights ($\overline{z}$):

$$z_s = z - \overline{z} \ . \qquad (1)$$

Now the EOFs are calculated and the first 5 are retained. The fractions of variance explained by the first 5 EOFs are given in Figure 1 and together they account for 93.2% of the observed variability. The first 5 EOFs are shown in Figure 2. The cutoff of five EOFs was used for consistency with K2K.

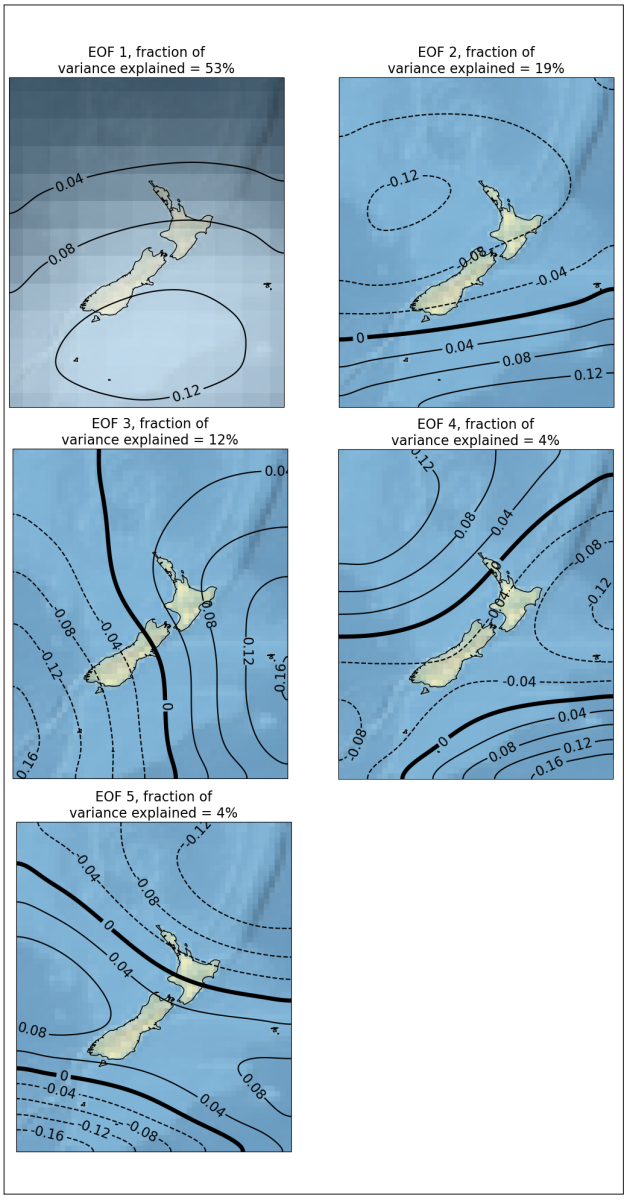

Figure 2: The first 5 EOFs, ε, of z_s. The contour lines are smoothed via linear interpolation between gridpoint values and the gridscale of the data used is shown in the subfigure for EOF 1. The background of the figures shows the local relief. The contour interval is 0.04 in all cases and varies from -0.16, to 0.16. The values are dimensionless. Positive contours are solid, negative contours are dashed and the zero line is thickened.

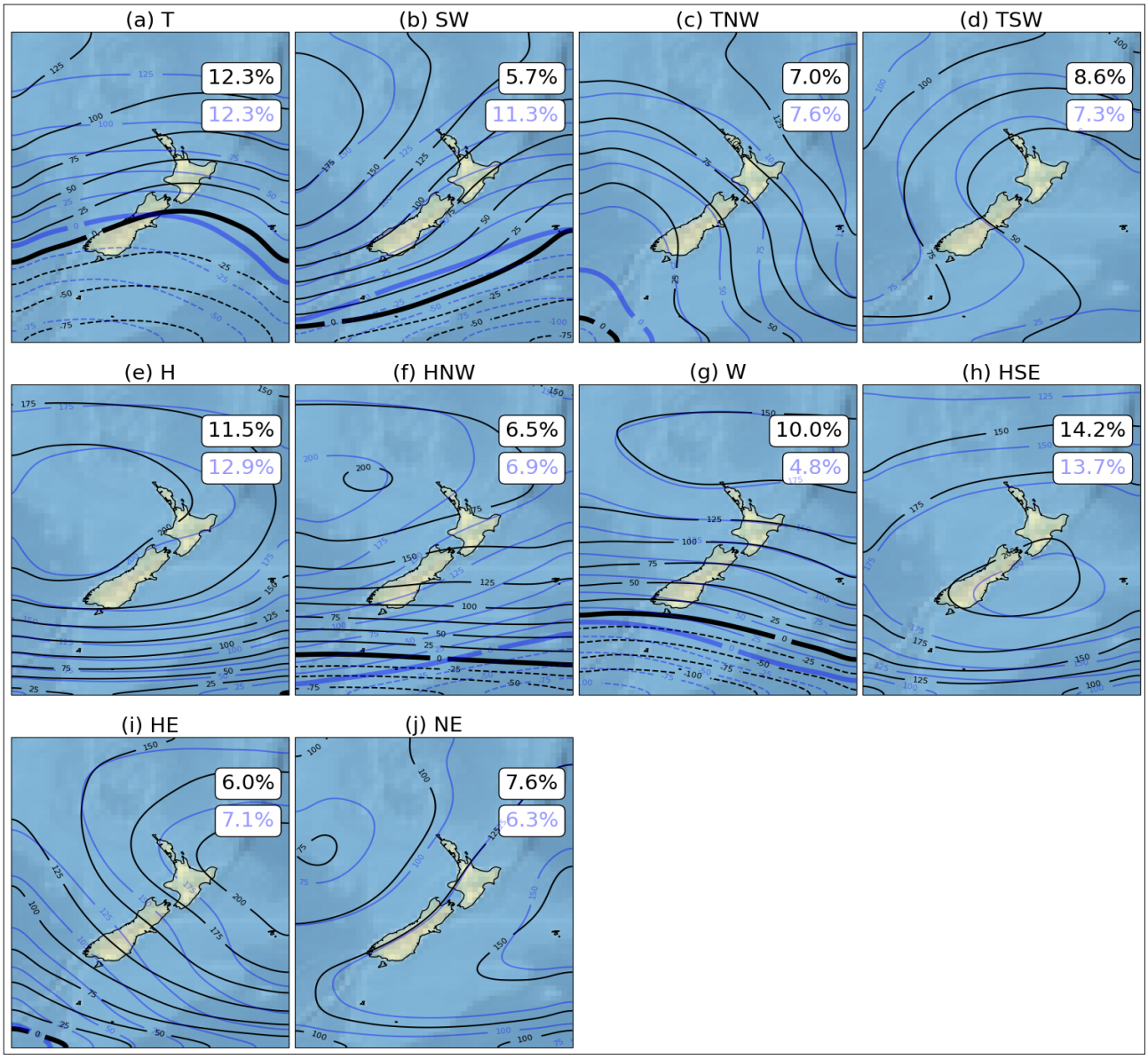

Figure 3: The 10 clusters which are deemed to be in close enough agreement with those of K2K to be assigned the same label (in black). Again the contour lines are smoothed and the original K2K clusters are shown in blue. The inset boxes show the fraction of time spent in each regime with the font colours matching the contour line colours. The background of the figures shows the local relief.

Next the PCs, P_n, are calculated, represening the amplitude of each EOF in a given height field. The fundamental operation here involves projecting the height field onto the EOFs and this is discussed further in the context of assigning regimes to arbitrary datasets in Appendix A. For the cluster analysis, the PCs are normalised to give:

$$\hat{P}_n$$

$$\hat{P}_n = \frac{(P_n - \overline{P_n})}{\sigma_{P_n}}, \qquad (2)$$

where σ_p is the standard deviation of P across time and $1 \leq n \leq 5$.

Now, the K-means clustering assigns each z_s field to one of 12 clusters. The K-means approach assigns the fields to an ever-reducing number of clusters, from an arbitrary starting point, to a final step of one cluster holding all observations. Our analysis has chosen to stop the clustering at 12 clusters, for consistency with K2K. This

order of these clusters is arbitrary and is chosen purely to match Kidson's original work. Once each z_s field has been assigned a value between 1 and 12, the final cluster means, C_z are the time mean of the z fields assigned to each cluster, that is,

$$C_{z,i} = \overline{z_i} \ , \tag{3}$$

where $1 \leq i \leq 12$.

The first 10 of these final clusters are shown in Figure 3, along with their equivalents from Kidson (2000). The associated winds from the same reanalysis product at the same twice daily sampling frequency are shown in Figure 4.

Although the agreement between the 2 analyses is generally good for the 10 clusters shown in Figure 3, the 2 remaining ones are not shown due to their pronounced differences with the HW and R blocking types from K2K. These differences are discussed in the next section.

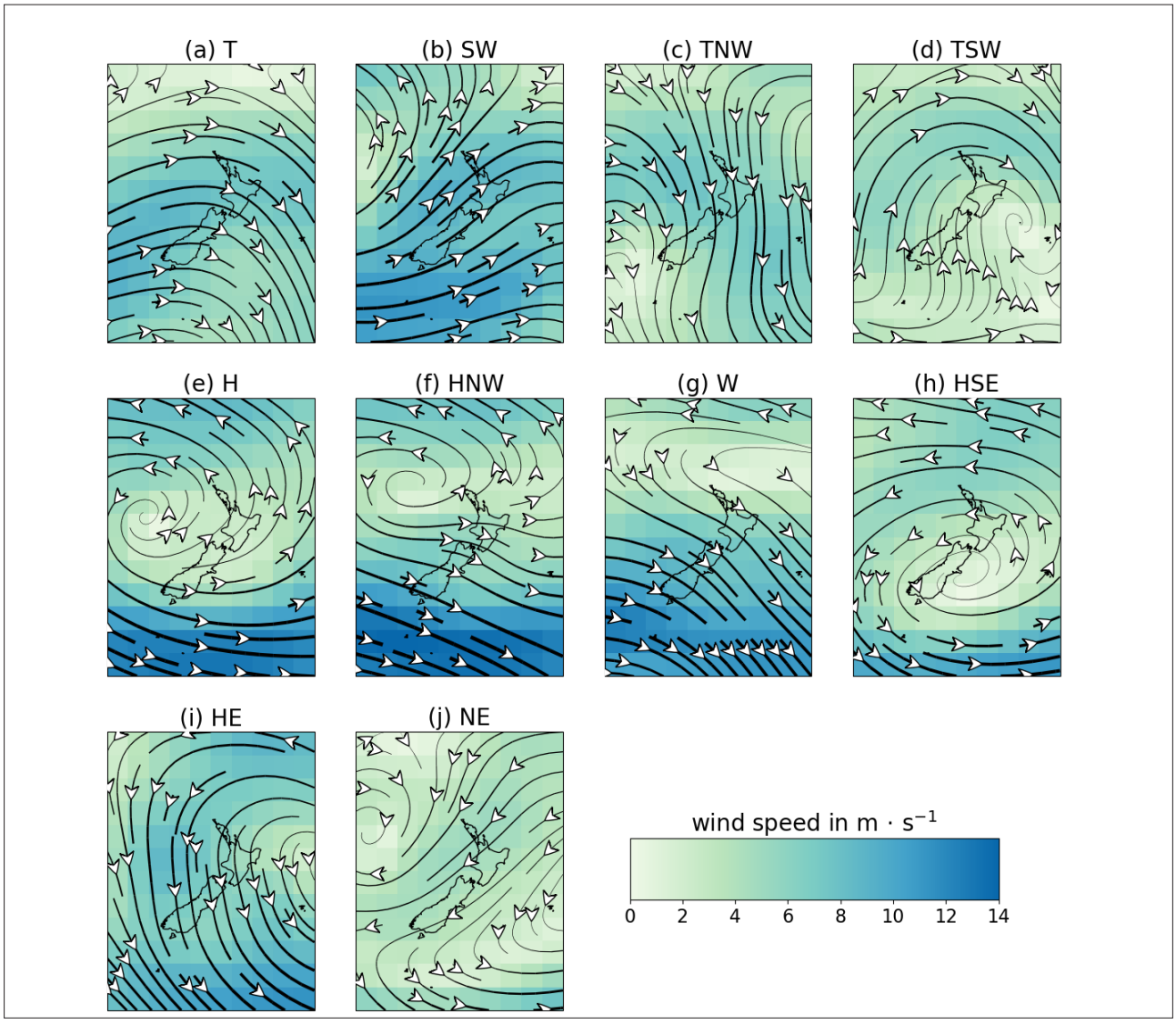

Figure 4: Wind streamlines and speeds (ms^{-1}) for the synoptic regimes shown in Figure 3. The colours show the windspeed at the gridscale. The thickness of the lines is proportional to the local speed and is consistent across subfigures.

3. How do the new synoptic types differ from Kidson's and why?

It is clear from Figure 3 that 10 of the synoptic types obtained here are generally in good agreement with those of K2K. That being said, the HNW regime is considerably more zonal in the new case, especially at southernmost latitudes. The SW regime from K2K is displaced north of the equivalent one from this analysis by approximately 100km and the opposite for W. This likely accounts for the factor of approximately 0.5 and 2 difference in occurrence frequency in the new case with respect to K2K.

There are however notable structural differences between the remaining two clusters obtained in this work and the HW and R blocking clusters in K2K. So much so in fact, that it is dubious to assign the same synoptic weather type in these cases. Figure 5 shows the HW and R blocking regimes from K2K along with the two remaining clusters from this work. The averages of the two remaining regimes are also shown.

The two new types shown in Figure 5 (a) and (b) are quite different to the HW and R regimes in K2K (Figure 5 (d) and (e)) yet their mean is strikingly similar, as is their combined fractional occurrence of 10.6% versus 10.1%.

As to why the new weather types found here are somewhat different to those of K2K, we have performed sensitivity analysis of the parameters used in the K-means clustering and have found the fractional occurrence of the regimes obtained here to be robust. For example, by default the K-means solver is run 10 times using different initial estimates and each of these estimates is iterated up to

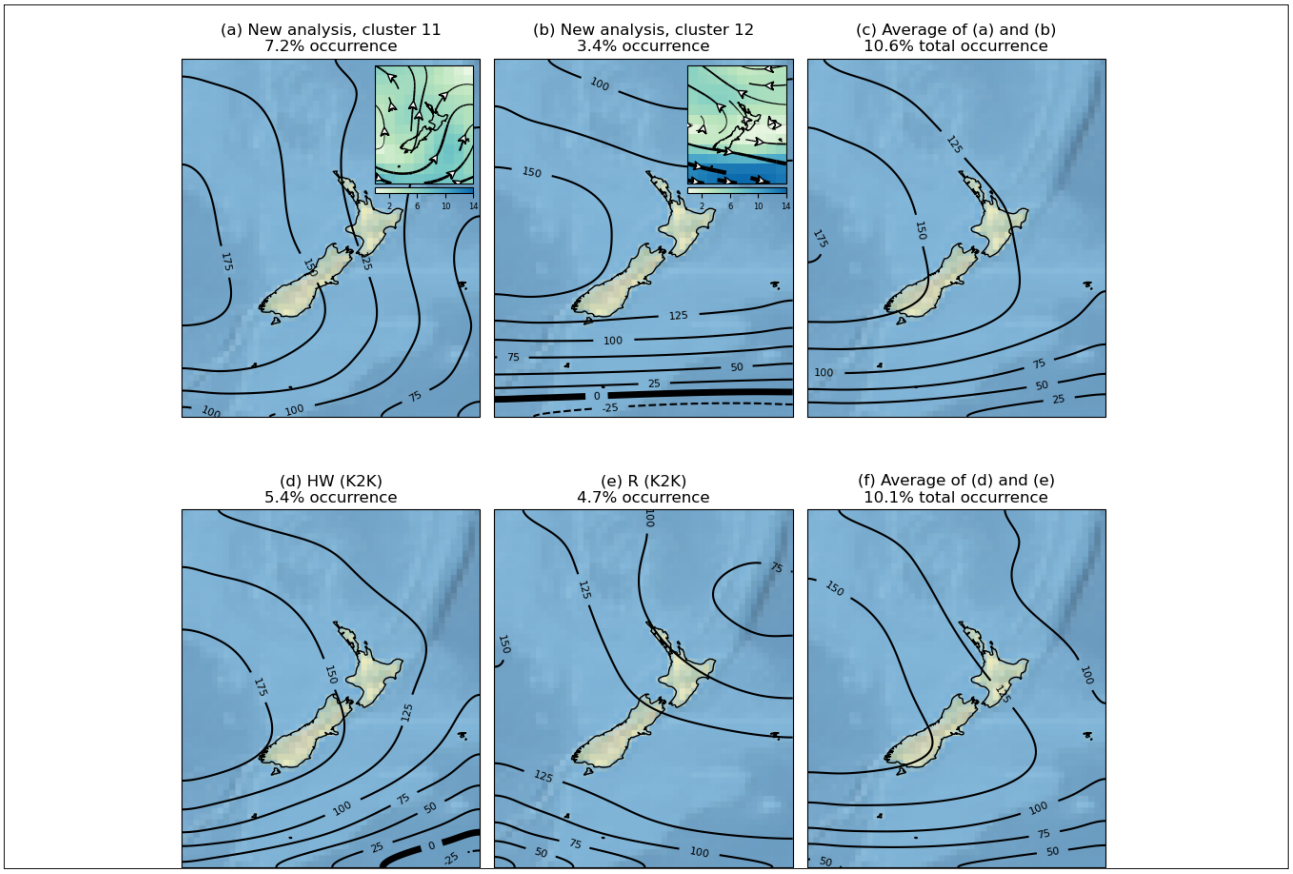

Figure 5: (a)-(b) The remaining two regimes from the new analysis which are not shown in Figure 3. Subfigures (a) and (b) also show the wind fields for the relevant cluster shown. The colour scale is the same as for Figure 4 and the horizontal density of the streamlines is halved to improve legibility. Subfigure (c) is the average of (a) and (b). (d)-(e) the HW and R regimes from K2K. Subfigure (f) shows the average of (d) and (e).

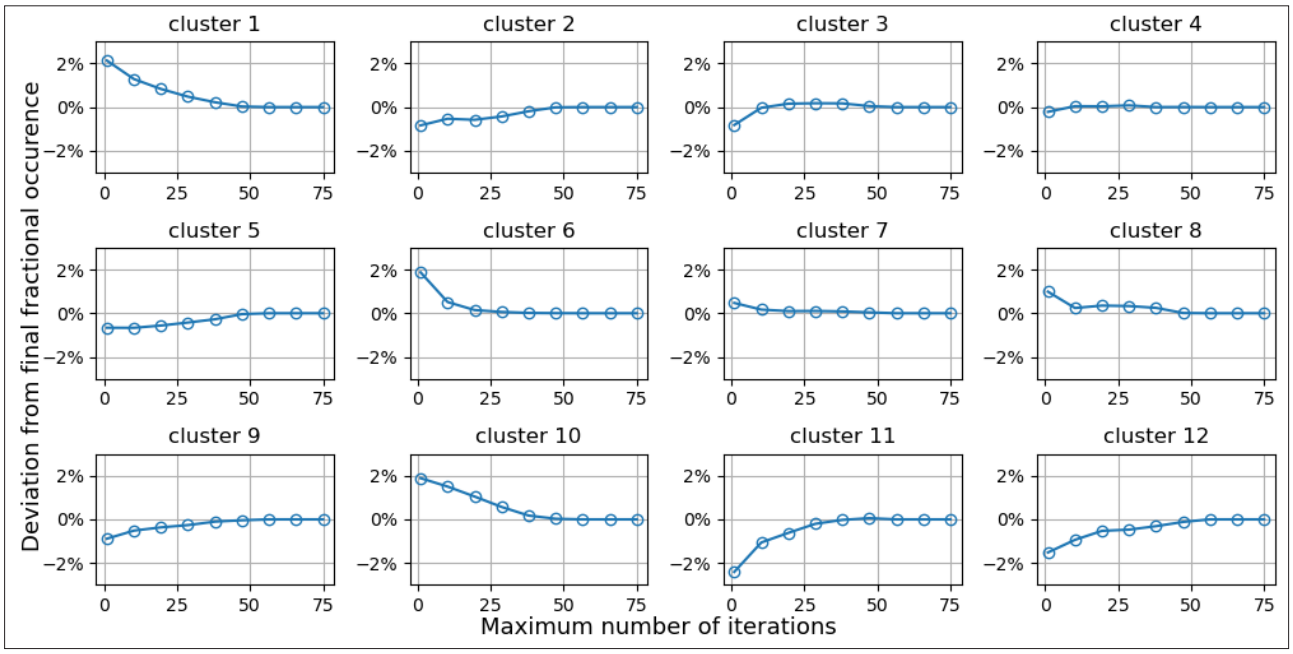

Figure 6: Deviation of the fractional occurrence of each regime as a function of the maximum number of iterations used in each pass of the K-means clustering algorithm. The largest difference is substantially smaller than the differences seen in Figure 3 (b) and (g) - which show the largest differences between the analyses - and therefore cannot account for the differences seen in the occurrence fractions.

300 times to ensure convergence. Figure 6 shows the relationship between the maximum number of iterations and the deviation of the final fractional occurrence of the regimes shown in Figures 3 and 5.

Figure 6 shows that the largest change in any of the cluster occurrences as the number of iterations is increased is approximately 2%. This therefore cannot account for the larger differences in the SW and W regimes found here compared to K2K (Figure 3 (b) and (g)). It is also possible that the dataset used in K2K was affected by the assimilation of incorrect pseudo-observations - or PAOBs - in early versions of the reanalysis (e.g. Kidson 1999).

It should also be acknowledged that no calculation is perfect and that different implementations of common algorithms will inevitably lead to some element of disagreement. This is noted in the documentation for the K-means software used here (Pedregosa et al., 2011): *"Given enough time, K-means will always converge, however this may be to a local minimum"*. For this reason - and also since the EOF/PCA methods are highly robust - we attribute the differences in the clusters obtained here predominantly to the K-means clustering method used. That said, the precise convergence criteria in K2K are not known (i.e. exacting details of the mathematical formalism or its application). Further discussion of cluster convergence is described in Michelangeli et al. (1995), using their concept of 'classifiability' indices.

The result shown in Figure 5 is reminiscent of the 'mixing' of EOFs or principal components subject to high sampling variability (e.g. Cheng et al., 1995). When the eigenvalues of consecutive EOFs are not statistically separate, different linear combinations of those EOFs can appear, as sample size changes. In this case, a similar thing seems to have occurred within the clustering algorithm and the definition of the cluster means.

At this stage, it is natural to ask if the new regimes obtained using ostensibly the same method are somehow more robust than those obtained in K2K. One way of doing this in a statistical fashion by comparing the ratio of intra- to inter-cluster variance in each of the analyses.

The intra-cluster variance for each cluster is calculated by calculating the variance through time at each gridpoint and then summing them. The inter-cluster variance is found in the same manner except that the dimension that the variance is calculated over is in 'cluster space' - i.e. over the 12 clusters - rather than in the time dimension. We then add the 12 intra-cluster variances together and divide by the inter-cluster variance to give one number for each analysis. By this method, the lower the ratio the better since decreased intra-cluster variance and increased inter-cluster variance indicate more 'distinct' clusters. Using this method, we find that the ratio in the new analysis is marginally 'better' than K2K but only by ≈ 0.4%; 8.74 for this analysis and 8.77 for K2K. Dividing by 12 to give these ratios on a 'per cluster' basis gives 0.728 and 0.731 respectively.

Although 10 of the 12 regimes found here are in close agreement with those from K2K - at least spatially - this is no guarantee that the two analyses put the same observations into the same category.

To investigate this further, it is instructive to construct a contingency table showing how the regimes assigned in K2K compare to those from this analysis. Figure 7 shows this table which compares the types assigned to every individual 12-hourly observation in the NCEP/NCAR dataset to the same classifications from this work.

Along with Figure 3 we can immediately see that simply because - for example - regime T occurs 12.3% of the time in both analyses, this is not because they both assign the same observations to the same cluster. In fact in this example we see that there are significant contributions to the T regime in this work from observations assigned to TSW, TNW and SW regimes in K2K.

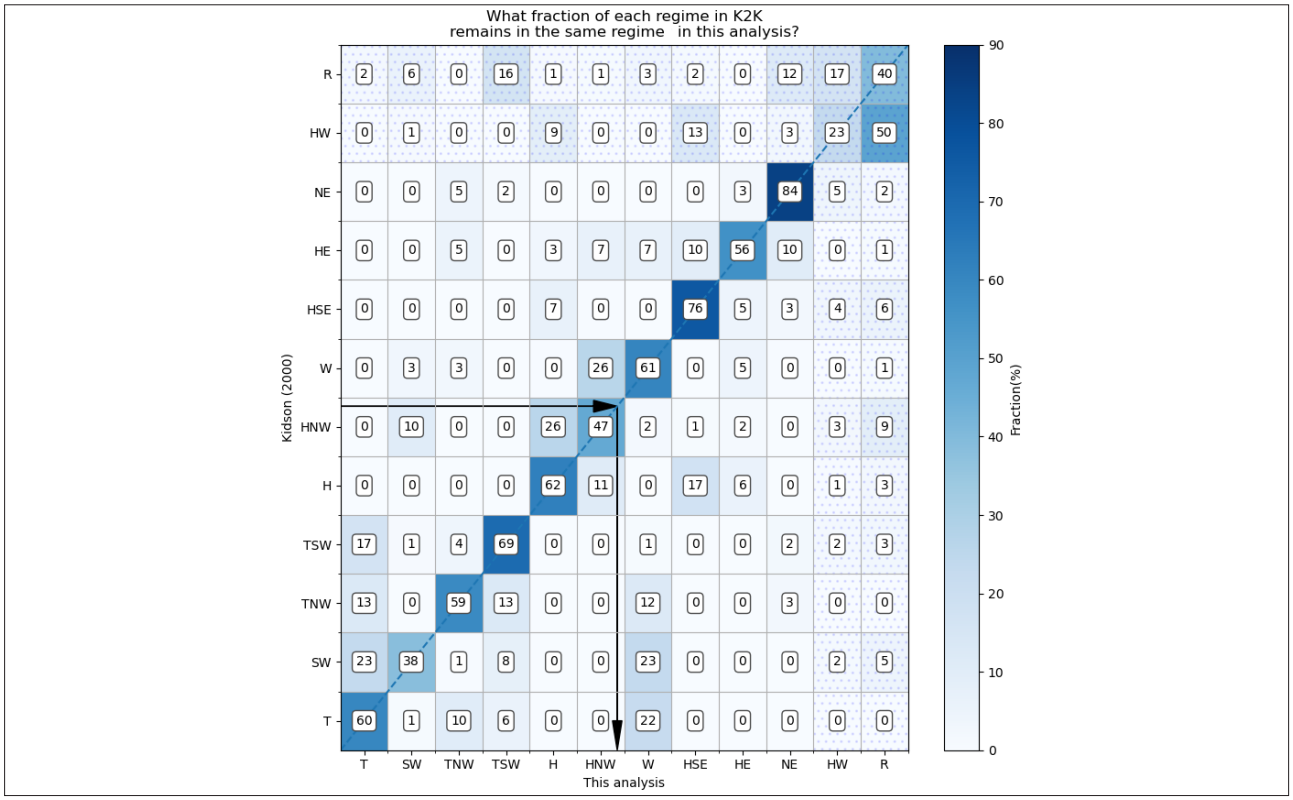

Figure 7: A contingency table showing the fraction of each regime in K2K which remains in the same regime in the new analysis, calculated on a point-by-point basis. The black arrows show - for example - that only 47% of the points in K2K which are assigned the HNW type, remain in this type in the new analysis. The hatching in the regions representing the HW and R regimes is to show the ambiguity in their assignment in this work; see Figure 5. Every row sums to 100% although this is not necessarily reflected in the integer values shown here due to rounding ambiguities.

4. New regimes using the ERA5 reanalysis

We now use the same method as described above but using the ERA5 reanalysis (Hersbach et al. 2020). The data is available hourly but in order to make the algorithms used here amenable to calculation on 1 processor in a time frame of ≈ 30 minutes, we have used 6-hourly data. At the time of writing, the ERA5 data is available from January 1979 and we use this data up until the end of 2014. The horizontal resolution is 0.25° compared to the 2.5° of the NCEP/NCAR data. Along with the doubling of time resolution used, this means that the overall number of points upon which to base this new analysis is 200 times greater than K2K.

Figure 8 shows the first five EOFs for the ERA5 dataset and is directly comparable with Figure 2 for NCEP/NCAR. There are some aspects to note immediately. Firstly, the broad structure of EOFs 1-3 on synoptic lengthscales is very similar to that given in Figure 2 for NCEP/NCAR. This is to be expected since the underlying physical state is the same, albeit represented by a different reanalysis scheme for different - but overlapping - time periods.

Secondly, compared to Figure 2, the 'order' of EOFs 4 and 5 has been reversed; broadly speaking EOF 4 for NCEP/NCAR has the same structure as EOF 5 for ERA5, and vice versa. This result illustrates the lack of statistical separation between EOFs 4 and 5.

Thirdly, the effect of the land is much more apparent, causing large local gradients and distortions in height contours. This is seen in all panels in Figure 8 but is particularly apparent for EOFs 3 and 4 as the zero contour crosses the Southern Alps and North Island Central Plateau respectively.

Figure 9 shows the 12 types obtained for the ERA5 dataset, ordered by their areal average mean geopotential height. There are some notable differences and similarities between the regimes in Figures 3, 5 and 9. For example Figure 3(j) and Figure 9(j) both show the same shape distinctive of the NE K2K-type, albeit with significant influence from the land in the latter - which is true for all of the new clusters. In addition the high pressure HSE type is seen in both analyses - Figures 3(h) and 9(l).

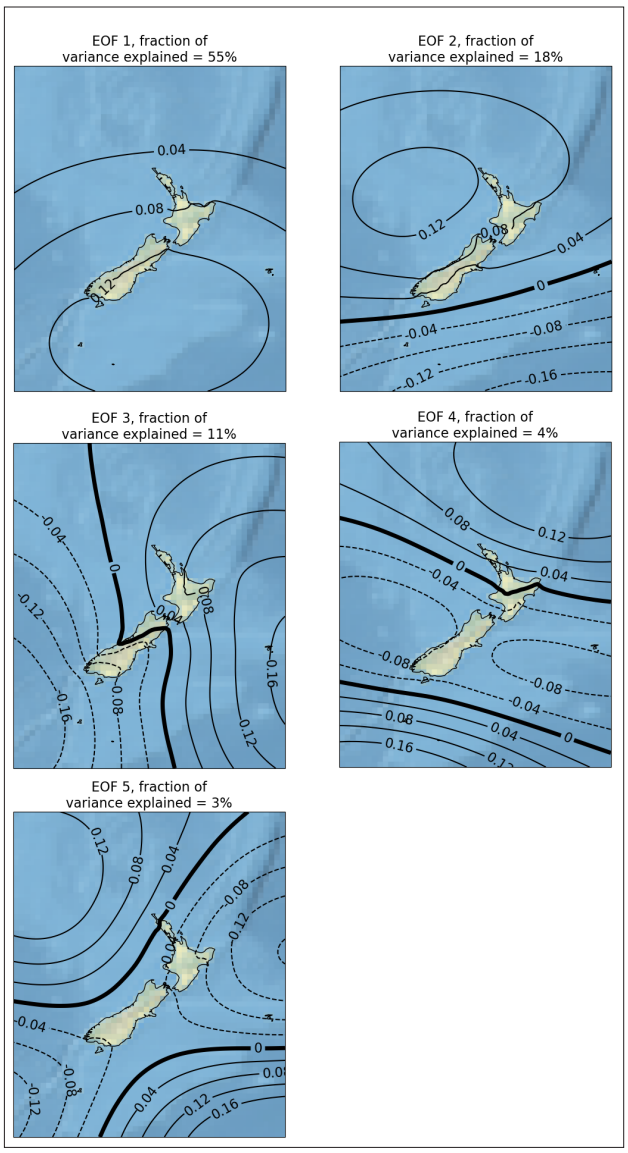

Figure 8: The first 5 EOFs, ε, of z_s for the ERA5 dataset. The contour lines are not smoothed in this case and the background of the figures shows the local relief. The EOFs are multiplied by 10 to aid comparison with Figure 2. This is due to the increase in the number of grid points and hence the total variance. The contour interval is 0.04 in all cases and varies from -0.16, to 0.16. The values are dimensionless. Positive contours are solid, negative contours are dashed and the zero line is thickened.

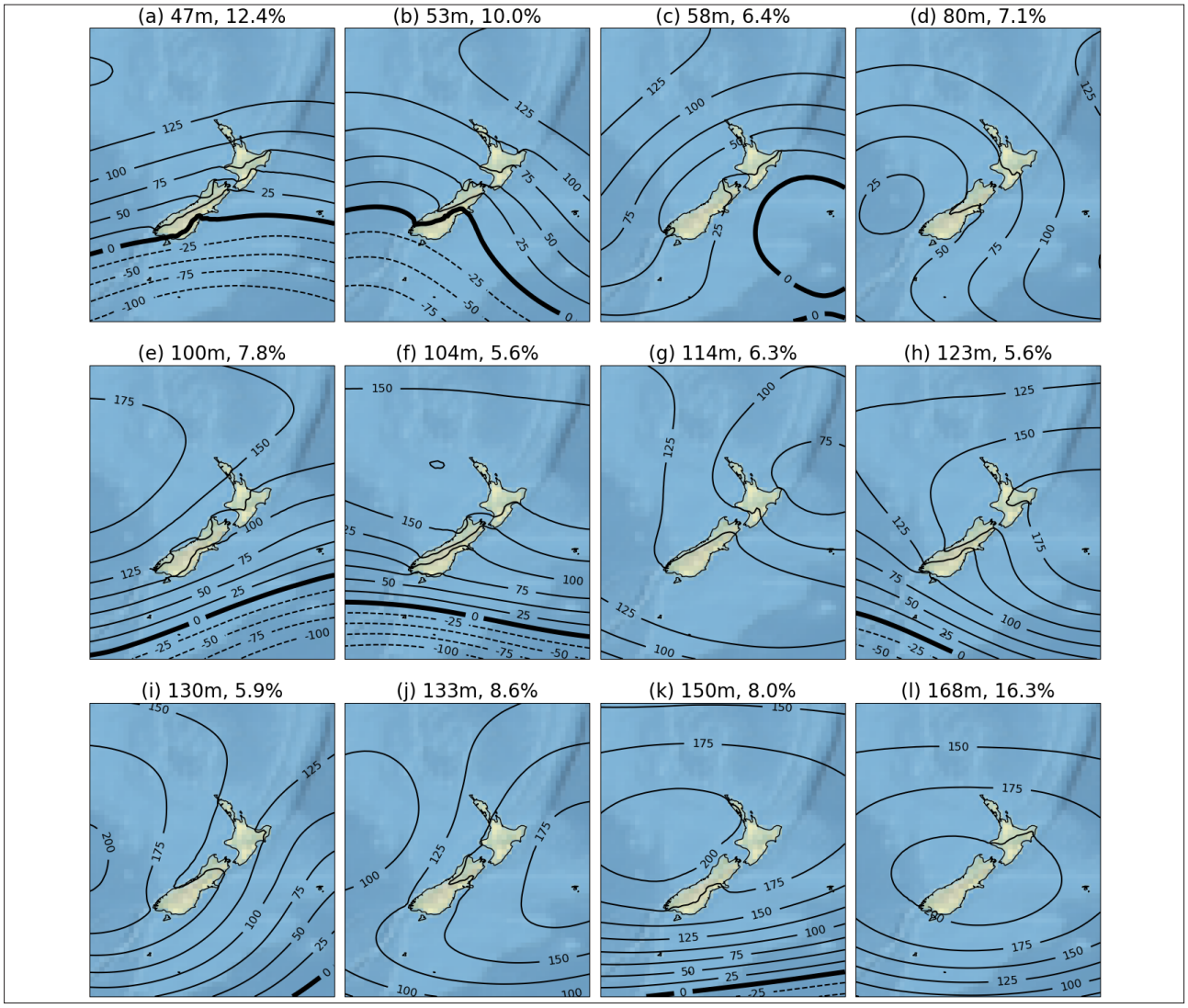

Figure 9: The 12 geopotential height clusters obtained for the ERA5 dataset. They are ordered from top left to bottom right by their areal mean geopotential height value. The frequency of occurrence for each regime is also given.

In contrast, some of the new ERA5 regimes do not have such a close analogue in K2K. For example, although Figure 9(g) is qualitatively similar to the R type in K2K, it has substantial local modifications.

Finally, revisiting the question of the ratio of intra- to inter-cluster variance for this new analysis, we find that using the 4-times-daily ERA5 data and 12 clusters gives an improvement in this ratio of ≈7% compared to the new NCEP/NCAR analysis. The individual values are 8.74 cf. 8.15 or - on a 'per cluster' basis - 0.728 cf. 0.679. This clearly illustrates the utility of using this improved dataset given the fact that the ratios for the two analyses of the NCEP/NCAR data were within ≈ 0.4% of one another and that we have only used 4 of the 24 available daily observations in ERA5.

It would be instructive in future work to construct a contingency table analogous to Figure 7 for the two reanalysis datasets used here and to examine our ERA5 results in the context of the recent study of Pohl et al., (2021).

5. Conclusions

In this work we have sought to provide not only a reassessment of dominant synoptic weather types over Aotearoa New Zealand from Kidson (2000), but also to provide an initial cluster analysis using the same methodology for a more recent dataset, ERA5. This product has significantly increased spatial and temporal resolution compared to NCEP/NCAR and indeed gives a substantially improved intra- to inter-regime variance ratio when 12 clusters are obtained. Widely-used peer-reviewed software packages in the Python programming language were used to calculate the EOFs and K-means clusters. Our results are broadly in line with K2K but with some notable differences. The key findings of this part of the work are:

- Although 10 of the 12 clusters identified in K2K are well reproduced in 'shape' in this work, 2 of the blocking regimes from K2K (HW and R) do not have recognisable analogues in the clusters found in this work. See Figure 3.

- Two of the regimes from K2K differ by factors of approximately 0.5 and 2 respectively with their spatially similar regimes from this work. This is accompanied by meridional shifts in the cluster mean height contours of about 100km in both cases, although in opposite directions.

- The average of the HW and R clusters from K2K is in striking agreement with the equivalent average of clusters 11 and 12 from this work, see Figure 5. This is attributed to different levels of EOF mixing in this work and in K2K and to subtle differences in the K-means clustering algorithms used in each case.

- Use of a contingency table for partitioning the results of this study and comparing them to those from Kidson 2000 shows that although some clusters are produced with very similar frequencies, this is not purely because the same observations are assigned the same type in the two analyses. Indeed for the T regime, even though both analyses show that this regime occurs 12.3% of the time, only 60% of the observations in this type in K2K are also assigned to this group in this work.

Our initial analysis of the clusters obtained from the ERA5 dataset show some interesting similarities and differences to previous work. Some of the dominant weather types identified in Kidson (2000) are very similar to those obtained in this new work, reflecting the overall synoptic dominance of some particular weather types, such as high pressure centred over Aotearoa New Zealand. There are however some types in the 12 cluster analysis which have some substantial local differences to those in K2K.

Future work will apply this methodology to differences in synoptic weather regimes over Aotearoa New Zealand in the UK Earth System Model (UKESM, e.g. Sellar et al. 2020) and in the NZESM (e.g. Behrens et al. 2020). This will be especially pertinent with regard to climate change and its effect on the dominant weather types that can be expected to occur in the future. It would also be of interest to examine the spatial and temporal occurrence of clusters from a selection of the CMIP6 models (see e.g. Eyring et al. 2016). This type of analysis over New Zealand has previously been reported for CMIP3 data (Parsons et al. 2014). Construction of a contingency table (Figure 7) comparing the clusters obtained from the two different reanalysis datasets used here would also be of interest.

6. Code and data availability

The analysis and plotting code used here is publicly available as a Jupyter notebook via GitHub (Williams 2021), the NCEP/NCAR reanalysis data is available at https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.pressure.html and the ERA5 data is at https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview.

Acknowledgements

JW would like to thank Dr Nicolas Fauchereau of NIWA for the idea to use the scikit-learn package at the outset of this work (https://github.com/nicolasfauchereau/ACRE_workshop) and therein noting the problems associated with reproducing the regimes found in K2K, thus stimulating the pedagogical aspect of this work. We would also like to acknowledge Dr Andrew Lorrey, also of NIWA, for useful discussions on the use and interpretation of synoptic regimes. JW is funded through the Deep South National Science Challenge, under project C01X1902 from the New Zealand Government Department for Business, Innovation and Employment (MBIE). The authors wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high performance computing facilities, consulting support and training services as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. URL https://www.nesi.org.nz. The authors also wish to acknowledge the comments from two anonymous reviewers whose comments and suggestions have greatly improved this work.

Appendix A: Manually assigning new data to a cluster set

In this appendix we use the term 'observation' to encompass any new dataset to which clusters are assigned.

There have been many studies based on the Kidson types and several of these involve fitting new datasets to the Kidson types (e.g. Ackerley et al. 2011, Parsons et al. 2014). In this section the mathematical basis for this assignment is given.

The first step is to calculate the 'projection' of the z_s onto the individual EOFs. The dimensions of the EOFs are 5 × 13 × 11, that is, 5 EOFs over a region with 13 latitude values and 11 longitude values. The z_s values have dimensions of Nt × 13 × 11, where N_t is the number of timesteps considered. The projection, $\mathcal{P}$, is defined as,

$$\mathcal{P} = \sum_{jk} z_{s,jk} \mathcal{E}_{jk} \ , \qquad \text{(A.1)}$$

and therefore, the dimensions of $\mathcal{P}$ are N_t × 5.

For each observation, we now have a 5 element array ($\mathcal{P}$) and a 12 × 5 element array representing the time mean of the PCs (P_n) for each index calculated by the K-means clustering algorithm,

$$C_{P,i} = \overline{P_i} \ , \qquad \text{(A.2)}$$

where $1 \leq i \leq 12$. The $C_{P,i}$ are often referred to in the literature as 'cluster means'.

To find out which of the 12 clusters the projection should be assigned to for each observation, the Euclidean distance, d, between the projection and each of the 12 height clusters is calculated. The projection and PC cluster arrays are normalised and are given by $\hat{\mathcal{P}}$ and $\hat{C}_{P,n}$ respectively. The minimum of these 12 numbers (i.e. the Euclidean distance in principal component space) gives the index and therefore the weather regime of each observation and its fractional occurrence. This can then be directly compared with the values obtained in Figure 3.

The Euclidean distances for each cluster, d_i, are given by,

$$d_i = \sqrt{\sum_{n=1}^{N_\mathcal{E}} \left| \hat{\mathcal{P}}_{n,i} - \hat{C}_{n,i} \right|^2} \ , \qquad \text{(A.3)}$$

where each of $\hat{\mathcal{P}}_{n,i}$ and $\hat{C}_{n,i}$ are 1 × 5 arrays and N_ε is the number of EOFs; i.e. 5 in this work.

A geometrical illustration of what the 'minimum Euclidean distance' means is shown in Figure A.1 for an arbitrary observation of z in the NCEP/NCAR reanalysis. It is clear that the figure showing the lowest d value - highlighted lines in Figure A.1(g) - has lines of $\hat{\mathcal{P}}$ and $\hat{C}$ which are closest to one another. The value of d can therefore be interpreted as a measure of the 'similarity' of $\hat{\mathcal{P}}$ and $\hat{C}$.

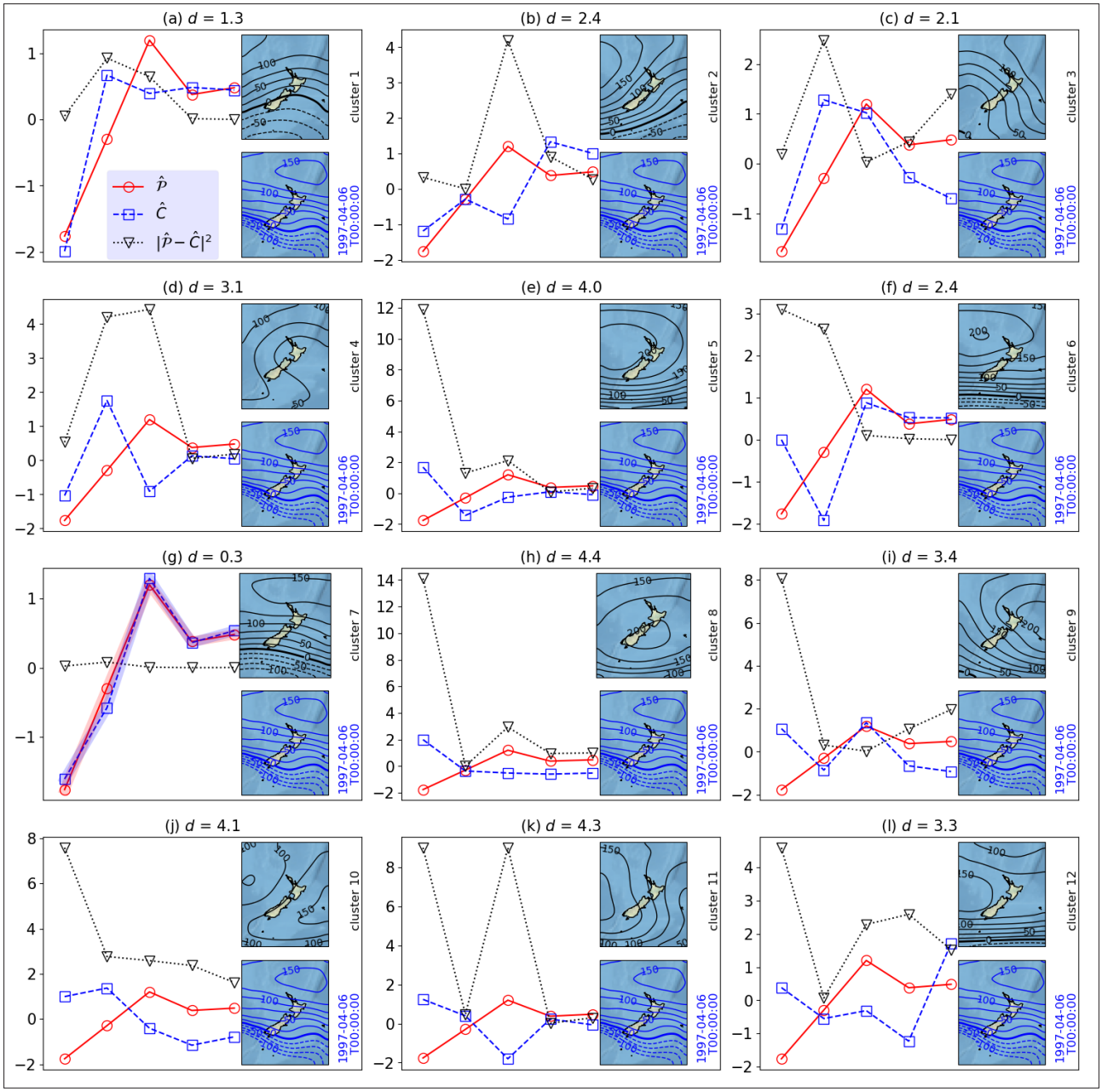

Figure A.1: Geometrical illustration of how an arbitrary observation of z (here at 0000 UTC on the 6th of April 1997) can be fitted to pre-existing clusters. Each subfigure shows the same arbitrary observation of geopotential height in the blue contours and the black contours show the 12 separate clusters obtained from this analysis (Figure 3 and Figures 5 (a) and (b)). As shown in the text, we want to find the smallest Euclidean distance, d, between the projection of the observation onto the EOFs, $\hat{\mathcal{P}}$ (○), and the cluster means, $\hat{C}$(□) as given by Equation 6. Note the simplified nomenclature used - i.e. lack of subscripts - to improve clarity. The individual $\left|\hat{\mathcal{P}}-\hat{C}\right|^2$ values are also shown (▽). The square root of the sum of the $\left|\hat{\mathcal{P}}-\hat{C}\right|^2$ values gives the Euclidean distance, d, shown in the subfigure titles and (g) highlights the similarity of the blue and red lines and hence the lowest d value.

References

Ackerley, Duncan et al. (2011). "Using synoptic type analysis to understand New Zealand climate during the Mid-Holocene". In: Climate of the Past 7.4, pp. 1189–1207.

Behrens, Erik et al. (2020). "Local grid refinement in New Zealand's earth system model: Tasman Sea ocean circulation improvements and super-gyre circulation implications". In: Journal of Advances in Modeling Earth Systems 12.7, e2019MS001996.

Cheng, Xinhua, Gregor Nitsche, and John M Wallace (1995). "Robustness of low-frequency circulation patterns derived from EOF and rotated EOF analyses". In: Journal of Climate 8.6, pp. 1709–1713.

Dawson, Andrew (2016). "eofs: A library for EOF analysis of meteorological, oceanographic, and climate data". In: Journal of Open Research Software 4.1.

Dawson, Andrew and Scott Wales (May 2019). ajdawson/eofs: Version 1.4.0. Version v1.4.0. doi: 10.5281/zenodo.2661604. url: https://doi.org/10.5281/zenodo.2661604.

Eyring, V. et al. (2016). "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization". In: Geoscientific Model Development 9.5, pp. 1937–1958. doi: 10.5194/gmd-9-1937-2016. url: https://gmd.copernicus.org/articles/9/1937/2016/.

Gibson, Peter B, Sarah E Perkins-Kirkpatrick, and James A Renwick (2016). "Projected changes in synoptic weather patterns over New Zealand examined through self-organizing maps". In: International Journal of Climatology 36.12, pp. 3934–3948.

Harris, Charles R. et al. (Sept. 2020). "Array programming with NumPy". In: Nature 585.7825, pp. 357–362. doi: 10.1038/s41586-020-2649-2. url: https://doi.org/10.1038/s41586-020-2649-2.

Hersbach, Hans et al. (2020). "The ERA5 global reanalysis". In: Quarterly Journal of the Royal Meteorological Society 146.730, pp. 1999–2049.

Jiang, Ningbo (2011). "A new objective procedure for classifying New Zealand synoptic weather types during 1958–2008". In: International Journal of Climatology 31.6, pp. 863–879.

Jiang, Ningbo, Kim N Dirks, and Kehui Luo (2013). "Classification of synoptic weather types using the self-organising map and its application to climate and air quality data visualisation". In: Weather and Climate 33, pp. 52–75.

Jiang, Ningbo, John E Hay, and Gavin W Fisher (2004). "Classification of New Zealand synoptic weather types and relation to the Southern Oscillation Index". In: Weather and Climate 23, pp. 3– 23.

Kalnay, Eugenia et al. (1996). "The NCEP/NCAR 40-year reanalysis project". In: Bulletin of the American Meteorological Society 77.3, pp. 437–472.

Kidson, John W (1994a). "An automated procedure for the identification of synoptic types applied to the New Zealand region". In: International Journal of Climatology 14.7, pp. 711–721.

Kidson, John W (1994b). "Relationship of New Zealand daily and monthly weather patterns to synoptic weather types". In: International Journal of Climatology 14.7, pp. 723–737.

Kidson, John W (1997). "The utility of surface and upper air data in synoptic climatological specification of surface climatic variables". In: International Journal of Climatology: A Journal of the Royal Meteorological Society 17.4, pp. 399–413.

Kidson, John W (1999). "Principal modes of Southern Hemisphere low-frequency variability obtained from NCEP– NCAR reanalyses". In: Journal of Climate 12.9, pp. 2808–2830.

Kidson, John W (2000). "An analysis of New Zealand synoptic types and their use in defining weather regimes". In: International Journal of Climatology: A Journal of the Royal Meteorological Society 20.3, pp. 299– 316.

Kidson, John W and Ian G Watterson (1995). "A synoptic climatological evaluation of the changes in the CSIRO nine-level model with doubled CO2 in the New Zealand region". In: International Journal of Climatology 15.11, pp. 1179–1194.

Michelangeli, Paul-Antoine, Robert Vautard, and Bernard Legras (1995). "Weather regimes: Recurrence and quasi stationarity". In: Journal of the Atmospheric Sciences 52.8, pp. 1237–1256.

Moore, Todd W and Richard W Dixon (2015). "Patterns in 500 hPa geopotential height associated with temporal clusters of tropical cyclone tornadoes". In: Meteorological Applications 22.3, pp. 314– 322.

Parsons, Simon, Adrian J McDonald, and James A Renwick (2014). "The use of synoptic climatology with general circulation model output over New Zealand". In: International journal of climatology 34.12, pp. 3426–3439.

Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". In: Journal of Machine Learning Research 12, pp. 2825–2830.

Pohl, Benjamin et al. (2021). ""Beyond Weather Regimes": Descriptors Monitoring Atmospheric Centers of Action—A Case Study for Aotearoa New Zealand". In: Journal of Climate 34.20, pp. 8341– 8360. doi: 10.1175/JCLI-D-21-0102.1. url: https://journals.ametsoc.org/view/journals/ clim/34/20/JCLI-D-21-0102.1.xml.

Saavedra-Moreno, B et al. (2015). "Surface wind speed reconstruction from synoptic pressure fields: machine learning versus weather regimes classification techniques". In: Wind Energy 18.9, pp. 1531– 1544.

Sellar, Alistair A et al. (2020). "Implementation of UK Earth system models for CMIP6". In: Journal of Advances in Modeling Earth Systems 12.4, e2019MS001946.

Sheridan, Scott C and Cameron C Lee (2011). "The self-organizing map in synoptic climatological research". In: Progress in Physical Geography 35.1, pp. 109–119.

Sheridan, Scott C, Helen C Power, and Jason C Senkbeil (2008). "A further analysis of the spatiotemporal variability in aerosols across North America: Incorporation of lower tropospheric (850hPa) flow". In: International Journal of Climatology: A Journal of the Royal Meteorological Society 28.9, pp. 1189–1199.

Solman, SA and CG Menéndez (2003). "Weather regimes in the South American sector and neighbouring oceans during winter". In: Climate Dynamics 21.1, pp. 91–104.

Trenberth, Kevin E (1997). "The Definition of El Niño". In: Bulletin of the American Meteorological Society 78.12, pp. 2771–2778.

Virtanen, Pauli et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: Nature Methods 17, pp. 261–272. doi: 10.1038/s41592-019-0686-2.

Ward J, H (1963). "Hierarchical grouping to optimize an objective function". In: Journal of the American Statistical Association 58.301, pp. 236–244.

Williams, Jonny (Aug. 2021). weather-types. doi: 10.5281/zenodo.5339123.
url: https://github.com/jonnyhtw/weather-types.

Williams, Keith D and MJ Webb (2009). "A quantitative performance assessment of cloud regimes in climate models". In: Climate Dynamics 33.1, pp. 141–157.