INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **32**: 1792–1802 (2011)
Published online 25 July 2011 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2394

RMetS
Royal Meteorological Society

A six-step approach to developing future synoptic classifications based on GCM output

Cameron C. Lee* and Scott C. Sheridan
Department of Geography, Kent State University, Kent, OH 44242, USA

ABSTRACT: One way in which global climate model (GCM) output can be utilized to infer local impacts is through the use of synoptic climatology: creating a set of atmospheric patterns that capture the variability in the climate system, and then analyzing trends and variability in the frequency of these patterns moving into the future. In this paper, we demonstrate a new synoptic climatological technique for classifying atmospheric patterns that can be used in conjunction with GCM output data (in this case, the Community Climate System Model 3). We apply this method to 850-hPa temperature patterns over the contiguous United States to derive daily categorizations. A total of 15 clusters are created from the data set; once the mean GCM bias is removed, historical cluster frequencies in the GCM data set are not statistically different from those of the reanalysis data set. In the future, significant changes in frequency are observed across most of the transition season clusters, as they broaden in seasonality at the expense of winter clusters, some of which nearly entirely disappear. Changes are greater moving further into the future, and greater for the more carbon-intensive special report on emissions scenarios (SRES) scenarios (A1FI, A2) than the less-intensive scenario tested (B1). Diagnostics test how well the mean patterns of the reanalysis data set, GCM historical data set and the future GCM data sets resemble each other. For some clusters, mean bias between the historical and future data sets grows substantially by the end of the 21st century under the more carbon-intensive scenarios. Copyright © 2011 Royal Meteorological Society

KEY WORDS synoptic climatology; climate change; GCM; circulation patterns; principal component analysis

Received 26 January 2011; Revised 1 June 2011; Accepted 12 June 2011

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) has concluded that it is very likely that anthropogenically induced climate change due to increasing levels of greenhouse gas emissions will increase global temperatures by anywhere from 1.1 to 6.4 °C by the 2090s (IPCC, 2007a). Beyond these global averages, there is much interest in understanding the potential local manifestations of climate change, to better understand the adaptations that may be necessary. Much research has suggested that natural hazards such as heavy precipitation events (Meehl *et al*., 2005), droughts (Sheffield and Wood, 2008) and heat waves (Hayhoe *et al*., 2010) all could increase in the future, while debates continue on whether an increase in tropical cyclone frequency or intensity could result due to this warming (Anthes *et al*., 2006; Pielke *et al*., 2005, 2006; Shepherd and Knutson, 2007).

The primary means by which such prognoses are made involves the output of global climate models (GCMs). GCMs are continually being improved to take an increasing number of different factors into consideration when trying to project future climates under different scenarios. Due to their ability to reproduce observed climates of the past, scientists have used GCMs to project the impacts of climate change on fields such as agriculture, human health, water resources and a number of other critical areas (IPCC, 2007b). In their fourth and most recent assessment report (AR4; IPCC, 2007b), the IPCC went into considerable detail in outlining the variables that GCMs have shown a marked improvement in projecting since their third assessment report 6 years before. Among such variables are large-scale distributions of temperature and pressure, and the general atmospheric circulation.

Nevertheless, due to the coarse spatial resolution of contemporary GCMs and their difficulty in accurately projecting certain variables such as low-level moisture and precipitation, most local- to regional-based impact research requires that some type of downscaling estimate be made from GCM output (Wilby *et al*., 2004). Dynamic downscaling techniques, including Regional Climate Models (RCMs; Christensen *et al*., 2007) have become more widely used in recent decades, although they are still computationally expensive and contain many inherent uncertainties, such as the choice of domain size and the ability of RCMs to produce similar features to the global models (Castro *et al*., 2005). Numerous statistical downscaling methods, which employ one of many techniques that use existing information about the relationship between broader-scale circulation and local climate impact, have also been employed to downscale GCM output to the local scale. Synoptic climatological techniques

* Correspondence to: Cameron C. Lee, Department of Geography, Kent State University, Kent, OH 44242, USA. E-mail: cclee@kent.edu

are one such means of statistical downscaling (Vrac *et al*., 2007).

Synoptic methods utilize a holistic approach to the climate system, taking the large-scale and long-term climatology of the atmosphere into consideration to create daily classifications of atmospheric circulation or weather types, and associating them with a variety of smaller scale surface events for analysis purposes (Yarnal, 1993). A variety of different synoptic classifications and classification methods have been applied and evaluated over the years; many of them are summarized by Yarnal *et al*. (2001) and Huth *et al*. (2008). More recently, with a focus on Europe through the COST733 Action (Philipp *et al*., 2010), an attempt has been made to systematically catalogue and evaluate a number of them. Huth (2010) evaluates 23 of these different methods in terms of their ability to stratify surface temperatures based on the historic record – finding that the number of clusters, seasonality and spatial domain play a large role in determining a classification's synoptic-climatological applicability.

Since synoptic methods are based largely on statistical associations to past environments, they are also useful in climate change impact research because they are able to take advantage of the variables GCMs are best at replicating in order to predict those that GCMs project less well. Because GCMs have become increasingly accurate at replicating *historic* atmospheric temperature and flow patterns; researchers have begun to project *future* patterns as well, and then apply these classifications to project impacts on weather events such as future freezing rain events (Cheng *et al*., 2007c), heavy precipitation events (Cheng *et al*., 2010), monthly precipitation (Ghosh and Mujumdar, 2006), air pollution (Cheng *et al*., 2007a, 2007b) and heat waves (Hayhoe *et al*., 2010). Often in the case of precipitation, the results derived from utilizing synoptic methods to project future precipitation are more accurate than the precipitation projected by the GCM itself (Wetterhall *et al*., 2009). Because the local-scale setting is conditional upon the surrounding synoptic environment, utilizing synoptic climatological methods in conjunction with certain well-projected variables from GCM output can serve a wide variety of local-scale impact applications.

A range of different synoptic studies have examined GCM ability to replicate observed historical patterns (Sheridan and Lee, 2010). The methods range from more traditional methods such as Kirchhofer (Saunders and Byrne, 1996; Schoof and Pryor, 2006) and automated Lamb weather types (Demuzere *et al*., 2009), to hybrid air mass classification schemes involving threshold determination (Schwartz, 1996) and more complex 'fuzzy' clustering techniques (Ghosh and Mujumdar, 2006; Wetterhall *et al*., 2009). Additionally, some studies use the still-emerging method of self-organizing maps (SOMs; Cassano *et al*., 2006; Lynch *et al*., 2006; Hewitson and Crane, 2006), while Huth (2000) demonstrates the effectiveness of T-mode principal components analysis (PCA) in classifying 500-hPa heights. Sheridan and Lee (2010; 2011) present a more exhaustive list and a brief overview of such studies.

Among studies using more traditional methods, McKendry *et al*. (2006) apply PCA and k-means clustering to sea-level pressure (SLP) data in the Pacific Northwest region using a Synoptic Typer application developed by the Australian Bureau of Meteorology. While the researchers find that the GCM can replicate the pattern shapes, some of these patterns are over- or under-represented in frequency in the 20th century – making future applications difficult. Cheng *et al*. (2007a, 2007b) use four weather types created from a previous study using PCA, average-linkage clustering, and discriminant function analysis (DFA) to develop a model for future air pollution in south central Canada, finding that high-pollution days in the area could increase by 180% by the latter part of the 21st century under some scenarios. In a separate study, Cheng *et al*. (2007c) define weather types in a similar manner to examine the potential risk of freezing rain events in south central Canada under future emissions scenarios. The authors suggest that synoptic weather typing is an effective technique in trying to evaluate the frequency of freezing rain events; concluding that they could increase in frequency by as much as 135% by the 2080s. Each of these studies highlight the utility of synoptic methods in analyzing certain aspects of the future climate – especially extreme events and precipitation – that are not as well projected by contemporary GCMs.

In this paper, we demonstrate a new synoptic climatological technique for classifying atmospheric patterns that can be used in conjunction specifically with future GCM output data. While the main goal is to describe the method, we demonstrate the utility of the method by applying it to 850-hPa temperature patterns over the contiguous United States, to derive daily categorizations into one of several clusters. Besides the goal of describing and demonstrating the method, the aim of our research process is to obtain realistic and discriminating clusters. Although any variable that is well replicated historically by a GCM can be used, we utilize 850-hPa temperature as it is one of the more commonly used variables in synoptic climatology, is a good indicator of surface temperature, and can be used in a wide variety of applications. The following sections detail the use of the method in classifying historical and future clusters, the resulting changes in the frequency and seasonality of these clusters in future decades, and a discussion of the benefits of and the limitations to this method.

2. Data and methods

In order to get a uniform grid of data to derive accurate clusters of historical atmospheric patterns that are easily comparable to GCM output data, the National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR; together, NCEP/NCAR) reanalysis data set (hereafter referred to as

the NNR20c data set) is used in this study (Kalnay *et al*., 1996). The 850-hPa temperature data are selected from the NNR20c data set at a once-daily temporal resolution (at 1200 UTC) over a 45-year period (from 1 September 1957 through 31 August 2002 – hereafter referred to more simply as 1957–2002) at a 5° by 5° spatial resolution spanning the United States from 56 °N to 21 °N latitude and 108 °W to 68 °W longitude. The domain clearly needs to be representative of the region for which projections are desired, which can vary geographically as well as in terms of the variable being used (Saunders and Byrne, 1999; Enke *et al*., 2005). Our grid spacing choice of 5° is a compromise intended to assure that the broader-scale features are paramount in determining cluster membership, as smaller grid cells have been shown to confound certain classifications (Demuzere *et al*., 2009). This resolution represents a sufficient scale at which to capture typical synoptic-scale circulation patterns and has been used in a number of other studies (Demuzere *et al*., 2009). We utilize 850-hPa temperature data rather than surface temperatures because the data at this level are naturally more smoothly spatially distributed than their surface counterparts, matching up well with the coarseness of the spatial resolution (Schwartz, 1996), although it is recognized that in the higher terrain of the Rocky Mountains this level may be difficult to interpret due to its proximity to the surface.

The GCM used in this research is the Community Climate System Model 3 (CCSM3); a GCM that combines four different models that separately project the atmosphere, the land surface, sea ice and the oceans; which are then integrated through a fifth component, a coupler, which coordinates the entire system (Collins *et al*., 2006).

The temperature data obtained from the GCM are at the same spatial domain and resolution as the data from the NNR20c data set. The CCSM3 is available for a historical period for verification purposes to check its ability to replicate the historical climate, and for different future emissions scenarios as well. The GCM scenarios used herein span the realm of IPCC emissions futures – A1FI represents a 'business as usual' fossil-fuel reliant, high emissions scenario; B1 represents a more environmental-friendly, low emissions scenario; and A2 is an intermediate emissions future. These scenarios were chosen in order to obtain the broadest range of possible future climates in order to best demonstrate the method. Output data from the CCSM3 in the historical run spanned the same 1957–2002 time period of interest, and from 2000 to 2099 in the three future runs. Data from the GCM's 20th century run are hereafter referred to as GCM20c data, while data that come from the GCM's projection of the future under one of the SRES scenarios are collectively referred to as GCM-Future data.

Due to the fact that many studies have found that climate change impacts on weather could manifest themselves as changes in seasonality (Sheridan and Lee, 2010), raw data, that is, data that have not been standardized to account for seasonality, are often preferred for analysis in impact studies. In terms of using GCM data, our initial attempt to cluster GCM20c data resulted in a significant mismatch of cluster frequency between the GCM20c and NNR20c datasets, suggesting that the GCM is not adequately reproducing lower tropospheric thermal patterns. An examination of the climatology of the GCM20c revealed a systematic bias, with 850-hPa temperature values overestimated in the winter (relative to the NNR20c) in the northern part of the study region, and underestimated in the south; while in summer, there is a warm bias in the centre of the domain with a cold bias farther north. To rectify this bias, the mean difference over the historical period between the NNR20c and GCM20c datasets has been calculated at each grid point on a monthly basis (e.g. all Januaries in 1957–2002). This difference is then subtracted from each GCM20c data value (as in Demuzere *et al*., 2009). With this GCM bias removed (or *debiased*), the methods below produce substantially better matches (lower mean biases) between GCM20c and NNR20c clusters. These debiased data sets are analyzed in this research. To account for this systematic bias in the future projections, the same monthly mean model bias is then also removed at each grid point in each of the GCM-Future data sets (A1FI, A2 and B1 scenarios) as well.

After debiasing the GCM data, synoptic categorizations are created using a six-part process (Figure 1). This six-part process is repeated for each of the three sets of merged NNR20c, GCM20c and GCM-Future data involved in the investigation (one for each of the three future scenarios). The six-part method outlined below represents *one iteration* of this process (e.g. NNR20c, GCM20c and GCM-Future A1FI). All analyses are run with SPSS statistical software.

The **first** part in the process is to perform an unrotated PCA on only the NNR20c data. The 850-hPa temperature values at each of the 72 grid points are entered into the PCA as the variables. The resulting principal component scores (PCs) with eigenvalues greater than one are retained and saved as variables – as this threshold signifies that the PC is accounting for at least as much variability as the variable it is replacing, and is a standard practice in synoptic research. Cuell and Bonsal (2009) note, however, that in some cases varying the threshold for PC retention can significantly change the outcome of a cluster analysis that produces synoptic types. In this research, several different permutations were tested, although the standard method is used here as no other permutations of PC retention resulted in a markedly more discriminating final classification. Ultimately, the number of PCs chosen is the decision of each investigator, with the ability of the classification to discriminate between different patterns and/or appropriately partition surface events being of substantial importance. Following the necessary data reduction, the uncorrelated PCs are then used in the cluster analysis in the following step.

The **second** part of this six-part process is to perform a two-step clustering (TSC; SPSS, 2001) on the retained PCs of the NNR20c data set. The goal of this step

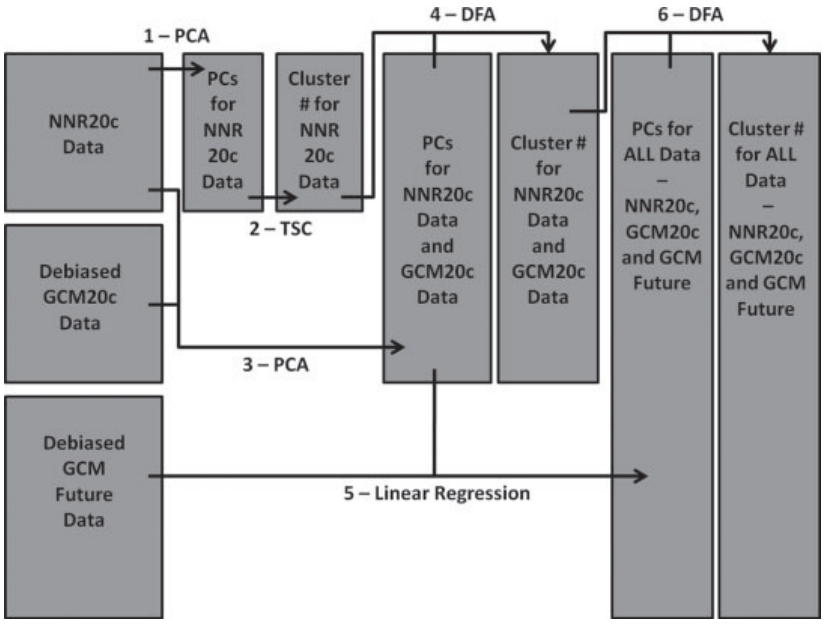

Figure 1. Flow chart depicting the data sets and statistical analyses used in each part of the six-step approach.

is to categorize every day in the NNR20c data set into one of 15 clusters – in essence, defining the initial spatial temperature patterns that all GCM data (both GCM20c and GCM-Future) will be based upon. The TSC is set to initially 'pre-cluster' the observations based on a distance criterion, followed by an agglomerative hierarchical clustering of the pre-clusters in the second step (SPSS Inc., 2001). The number of clusters is set at 15 after examining a number of different permutations, as the decision of the number of clusters to use can also substantially affect the final clusters produced. Among others, Spekat *et al*. (2010) discuss limitations in deciding on the number of clusters. Thus, while a total of 15 clusters are able to most accurately depict the typical modes of temperature variability over the seasonal cycle in this demonstration of the method, for other research this may not be optimal. The ultimate goal of the project must be considered when deciding on the number of clusters to use (Yarnal, 1993), as the final number of clusters chosen is a substantial factor in determining the ability of a classification to resolve surface environments (Huth, 2010; Schiemann and Frei, 2010; Tveito, 2010).

The **third** part of this process is to perform a second PCA, this time including the temperature values in both the NNR20c data set and the GCM20c data set as variables entered into the PCA. The same settings are used as in the previous PCA, which creates uncorrelated PCs for both the NNR20c and the GCM20c data sets that can be used for the DFA in the next step. A linear regression (to create a model from the NNR20c data set to predict the PCs of the GCM20c data set – similar to what is used in part 5 below) is not used here because the variability of the GCM20c data set is thought to be an important aspect of the classification as well.

The PCs created from the previous step are then subjected to a DFA in the **fourth** part in the process. The DFA uses the cluster variable created in part two as the grouping variable and the newly created PCs (from part three) as the independent variables. The stepwise method is used and the classification of the prior probabilities is computed using the group sizes from the clusters. The predicted group membership number is saved as a variable in the data set. Following this process, each case in both the NNR20c data set and the GCM20c data set is classified into one of the 15 clusters. Initially, a separate PCA and cluster analysis of just the GCM20c data set was used in hopes of matching GCM20c patterns to NNR20c patterns. However, some GCM20c patterns did not match well with NNR20c patterns; or matched up equally well with multiple NNR20c patterns. Thus, the variability of the NNR20c data set also needs to be captured in the PCs to help the GCM replicate NNR20c patterns for the purposes of the classification. In using DFA, the patterns created using GCM20c resemble the NNR20c patterns. In performing the DFA, the NNR20c data set is reclassified in addition to the classification of the GCM20c into clusters, in order to maintain cohesiveness among the two data sets. In this process, around 11.8% of the NNR20c days are reclassified to a different cluster from their original classification.

The **fifth** part in the process involves the development of linear regression relationships between the PCs and the original values for both the GCM20c and NNR20c data sets, and utilizing these relationships to predict PC values for the GCM-Future data. The retained PCs created in the third step of this process are used as the dependent variables to be predicted, while the temperature values – each of the grid points used for the PCA in parts 1 and 3 – for all data sets are used as potential

independent variables. The stepwise method of linear regression is used (with the significance values set at $\alpha = 0.05$ for entry and $\alpha = 0.10$ for removal), and the unstandardized predicted values for each PC are saved as variables in the data set. Several different permutations of the number of potential variables to retain in the regression equations were tested, with minimal frequency changes observed (less than 0.3% in all cases). A separate linear regression is repeated for each retained PC created in step 3. The advantage of using regression here instead of PCA is that future clusters are as similar to 20th century clusters as possible, maintaining the shape of each cluster and allowing an examination of changes in frequency and seasonality. Thus, if by the 2090s, the climate is substantially warmer over the region, as is predicted by many studies (IPCC, 2007a), using PCA may result in patterns that occur mostly in the future and rarely in the past – which would significantly hinder a comparison of patterns between the two time periods. Regression allows these predicted PC values to be based on 20th century variability and thus allows the succeeding step to classify these clusters accordingly.

The final clusters are created in the **sixth** part of this process – a DFA of the PCs created from the linear regression in the previous step. The DFA cluster number variable created in part four is used as the grouping variable and the predicted group membership is saved as a variable in the data set. Additionally, the classification of the prior probabilities is computed using the group sizes from the previous DFA groups. The predicted group membership into a cluster is based on a percent likelihood of that day's inclusion into that cluster. While likelihood values are produced for each day's membership in each of the 15 clusters, for ease of interpretability, each day is assigned to the cluster for which the likelihood is greatest. This DFA results in every day's temperature grids (in the NNR20c, GCM20c and GCM-Future data sets) being classified into one of the 15 clusters. Each cluster's mean 850-hPa temperature value at each grid point can then be used to map the cluster. Comparing the clusters created using the NNR20c data set with the DFA clusters created in part four above, 8.2% of days are reclassified into a different cluster; while 8.3% of the GCM20c days that were reclassified using the DFA in the sixth step are assigned to a different group than the one assigned in the fourth step.

For the purposes of the demonstration of the method herein, once complete, this six-part process is then repeated twice more for the two other data sets representing the two other GCM-Future scenarios. If more variables, GCMs, and/or scenarios are used for future projections, then the six-part method is repeated for each permutation of variables and model-scenarios. Additionally, the method presented above has flexibility with regard to the number of principal components retained (in parts one and three), the number of clusters chosen (in part two), and the entry and removal criteria for the stepwise portions of the method (linear regression in step five and DFA in parts four and six).

Frequency differences between the NNR20c and GCM20c, and between the GCM20c and future scenarios, are tested by the G-test for goodness of fit. The G-test is a likelihood ratio test similar to the χ^2 goodness-of-fit test, except it is defined as:

$$G = 2\Sigma[O_i \times \ln(O_i/E_i)] \quad (1)$$

where O_i are the observed values (e.g. GCM20c cluster frequencies in Step 2) and E_i are the expected values (e.g. NNR20c cluster frequencies in Step 2).

3. Results

3.1. The clusters

The 15 clusters created are shown in Figure 2, mean annual frequencies for each cluster are found in Table I, and mean seasonal frequencies in Figure 3. The clusters are reordered to progress generally from summer-dominant to winter-dominant. The summer is dominated by two clusters, A and B. **Cluster A** is the most commonly occurring cluster, and accounts for over half of all days in both July and August. The warmest of all clusters, a strong thermal ridge extends over most of the western United States, with areas below 10 °C confined to the extreme northwest and northeast. **Cluster B**, the second-most common cluster overall, features a much weaker thermal gradient, with warm conditions over much of the country. In contrast to A, this cluster is more broadly warm-season dominant, with frequencies from 28 to 38% of days between June and October.

Clusters C–F dominate the transition seasons. **Cluster C** is a late-spring-dominant cluster, occurring on over half of all days in May. A thermal ridge dominates the western United States, with a baroclinic zone incorporating much

Table I. Mean annual present frequency of each cluster.

Cluster	1957–2002		2050–2059			2090–2099		
	NNR20c	GCM20c	A1FI	A2	B1	A1FI	A2	B1
A	18.3	15.9	16.3	16.3	15.9	14.3	14.6	16.1
B	15.5	16.1	14.3	17.2	15.3	16.0	16.8	15.6
C	8.3	8.3	**19.8**	**15.7**	**14.1**	**27.3**	**22.7**	**12.9**
D	6.4	7.8	**14.3**	**11.8**	10.7	**18.9**	**17.4**	10.6
E	5.9	5.4	4.5	6.7	5.8	3.6	5.2	6.3
F	4.4	5.1	**2.4**	3.6	3.8	**0.5**	**1.7**	4.3
G	5.2	3.9	**1.2**	**1.0**	**1.7**	**0.4**	**0.5**	**1.4**
H	4.1	5.4	3.3	3.8	5.5	**1.4**	**1.9**	5.0
I	5.3	4.1	**1.0**	**1.6**	2.7	**0.2**	**0.7**	2.7
J	5.7	5.7	5.1	4.6	7.2	**1.6**	**1.9**	5.9
K	4.6	3.9	5.7	5.4	4.7	**7.7**	**7.1**	4.6
L	4.8	6.2	4.6	3.7	4.2	**2.0**	**3.3**	4.4
M	2.7	2.8	1.4	1.3	1.8	**1.0**	**1.0**	1.7
N	3.1	3.6	1.9	2.6	2.3	2.6	2.1	2.7
O	5.6	5.9	4.1	4.8	4.2	**2.6**	**3.2**	6.0

Statistically significant differences ($p < 0.05$) between GCM-Future frequencies and GCM20c frequencies are bolded.

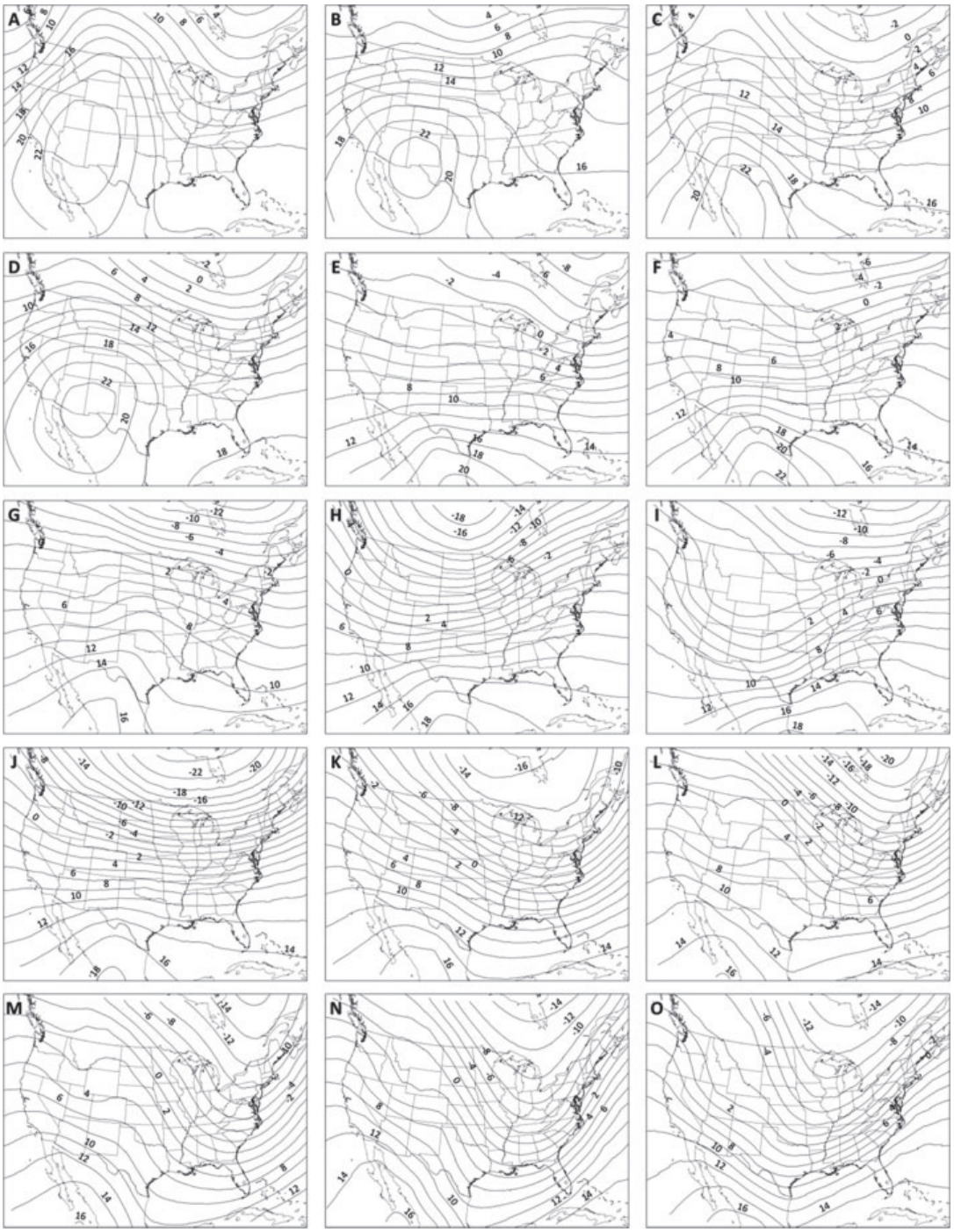

Figure 2. The 850-hPa temperature patterns for each cluster created with the six-step method. Contours are in intervals of 2 °C. Maps represent mean conditions from the GCM20c data set.

of the south central and southeastern United States. **Cluster D** is mostly a late-summer and early-autumn cluster that is similar to cluster C, although warmer throughout, with the main thermal gradient shifted northwards to the Great Lakes and northeastern United States. **Cluster E** occurs throughout the winter, and then peaks in March and April. Zonal isotherms across much of the country are the main features of this cluster. **Cluster F** features a broad zonal thermal pattern, with a baroclinic zone across the north central United States. It occurs most often in the spring, peaking in April at nearly 30% of all days with over 10% frequency in both March and May.

Clusters G–O are all moderately frequent winter-dominant clusters, each accounting for between 2.7 and 5.7% of days in the NNR20c data set. **Cluster G** is a mild cluster, with a slight thermal ridge in the Southern

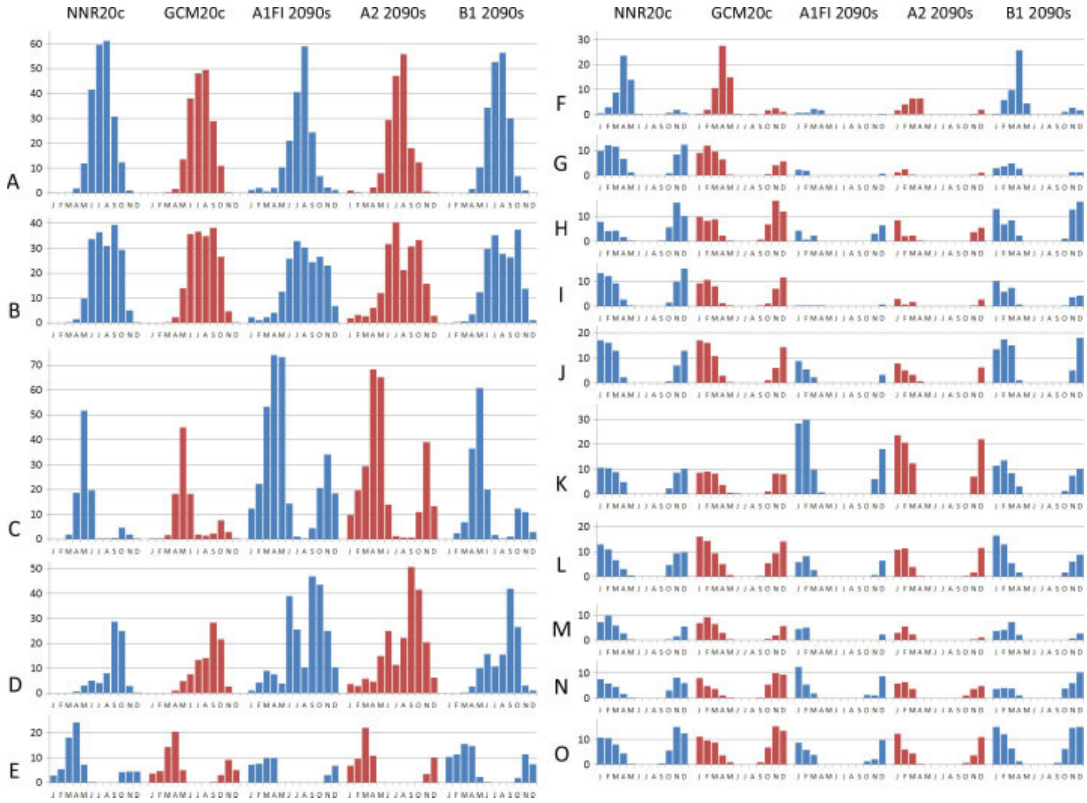

Figure 3. Mean monthly frequency for each cluster for each of the data sets listed. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

Plains and only a part of the northeastern United States below 0 °C. A strong temperature gradient with very cold air approaching the north central United States is the main feature associated with **Cluster H** – a cluster that occurs relatively infrequently and is confined to winter. **Cluster I** features a thermal trough over much of the Rockies and the south central United States, and a thermal ridge over the eastern United States. This winter-dominant cluster suggests a cyclone in the central United States, with much of the eastern United States in the warm sector. Conversely, **Cluster J** contains the strongest thermal gradient of any cluster, with over a 30 °C range in temperature across the coterminous United States. **Cluster K** is marked by a strong cold front moving into the northeast United States, dipping the mean 0 °C line far into the southeastern United States. **Clusters L and M** are generally similar to cluster K, with progressively sharper thermal troughs across the northeastern United States. **Clusters N and O** are similar as well, except that in both cases the thermal trough is shifted farther westwards, to the Great Lakes and Mississippi River Valley, respectively. Both N and O have their peak frequency towards the beginning of the cold season.

3.2. Comparison of NNR20c and GCM20c cluster frequencies and means

In comparing the frequency of the 15 clusters in the NNR20c and GCM20c data sets, substantial consistency appears in their annual (Table I), seasonal and monthly frequencies (Figure 3). The annual mean frequency differs by more than 1.5% of days in only one cluster, Cluster A, where the GCM20c data set underestimates the occurrence observed in the NNR20c data set by 2.4% of days. Several of the less-frequent clusters, such as H and L, have relatively small discrepancies that translate into much larger percent deviations due to their lower frequency, although none of these differences are statistically significant. The seasonality of each cluster is well captured, with the peak month of occurrence identical for each cluster in both of the data sets. For each cluster, there are no statistically significant differences in the seasonal distribution of frequency. Only for the aggregate frequency of meteorological summer (June–August) is the frequency difference across all clusters between the NNR20c and GCM20c data sets nearly statistically significant ($p = 0.074$); for all other months and seasons, there are no statistically significant differences in cluster frequency.

By virtue of the method by which these clusters are derived, unsurprisingly there is a very high correlation between the temperature fields for each cluster (Table II). Pearson correlation coefficients (between the mean temperature at each grid point in NNR20c and the corresponding grid in GMC20c) are at least 0.994 for all clusters except cluster C (0.991); the mean bias between NNR20c and GCM20c mean temperatures is under 1 °C

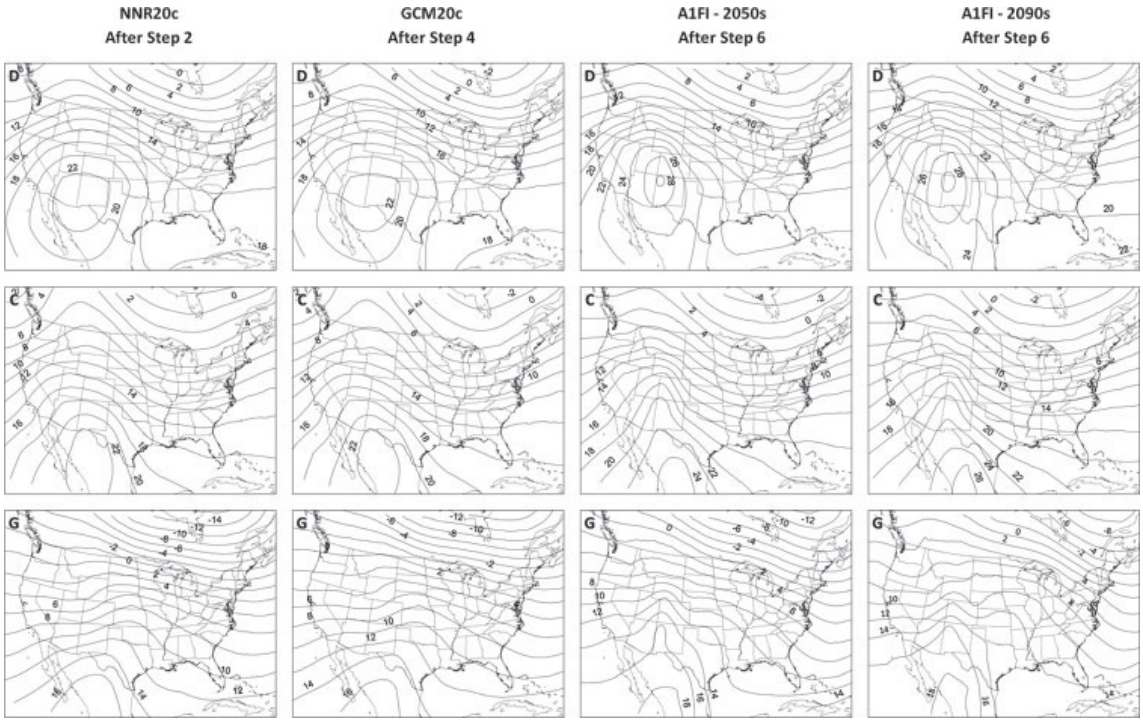

Figure 4. Qualitative diagnostic maps showing the *range* of progressions of the pattern shapes from the different steps of the six-part method. The top pattern (D) represents the pattern with the highest correlation between the GCM20c and the NNR20c data sets, the middle pattern (C) has the lowest correlation between the GCM20c and the NNR20c data sets and the bottom pattern (G) has the lowest correlation between the GCM20c and the A1FI 2090s.

Table II. Pearson correlation coefficient (r) and mean bias (°C) in comparing mean-temperature fields for each cluster between listed data sets.

Cluster	NNR20c GCM20c		GCM 20c GCM 2050s A1FI		GCM 20c GCM 2090s A1FI	
	r	Bias	r	Bias	r	Bias
A	0.998	−0.07	0.983	3.37	0.970	5.57
B	0.997	0.10	0.989	2.64	0.988	4.06
C	0.991	−0.08	0.993	0.38	0.991	1.52
D	0.999	0.87	0.995	2.62	0.994	3.65
E	0.996	−0.24	0.991	1.18	0.992	2.93
F	0.999	−0.02	0.988	0.36	0.967	2.63
G	0.997	0.48	0.981	2.11	0.966	3.78
H	0.997	−0.23	0.997	1.55	0.994	3.04
I	0.999	0.20	0.994	2.05	0.986	3.91
J	0.999	−0.04	0.997	2.76	0.994	4.70
K	0.999	−0.12	0.995	2.23	0.989	3.78
L	0.994	0.05	0.995	2.10	0.987	4.35
M	0.997	0.24	0.988	2.26	0.973	3.49
N	0.998	0.24	0.992	2.03	0.977	2.68
O	0.998	−0.14	0.996	1.42	0.991	2.57

for all clusters. These results strongly suggest that the clusters produced within the CCSM3 over the historical period match those that appear in the reanalysis data set.

The ability of the classification to categorize GCM20c patterns that are similar in shape to NNR20c patterns can be evaluated qualitatively in the maps provided in Figure 4. These maps provide a diagnostic tool to evaluate the progression of the classification through each of the six-parts of the method. Cluster D has the highest correlation between the NNR20c and the GCM20c data sets, while cluster C has the lowest. While cluster D is very similar throughout North America between the first two left-hand columns in Figure 4, despite having a comparatively low correlation, cluster C only differs slightly in the depth of the eastern thermal trough around James Bay, but the shape and magnitude of the isotherms in the western ridge remain largely unchanged.

3.3. Comparison of GCM20c and GCM-Future cluster frequencies and means

As is typically observed in climate projection studies, there are greater changes to cluster frequencies across the more carbon-intensive scenarios (here, A1FI and A2) than the less-intensive scenarios (B1), and generally greater changes as the century progresses (Figure 3). However, in the B1 scenario, cluster frequencies in the 2090s are actually more similar to those observed in the 20th century than those of the 2050s are.

Through the 2050s, the majority of the 15 clusters do not change significantly in annual frequency from the GCM20c; only 5, 4 and 2 clusters show statistically significant changes in the A1FI, A2, and B1 scenarios, respectively. Only three clusters increase in frequency across all three scenarios: C, D and K; with the largest

increases by far in cluster C, which increases from 8.3% of days in GCM20c to between 14.1% (B1) and 19.8% (A1FI) by the 2050s. These increases largely come at the expense of cluster F and most winter-dominant clusters; the summertime clusters, A and B, do not change much in overall frequency. In the 2090s, these same trends appear with a greater magnitude in A1FI and A2 (in both, 11 clusters have statistically significant frequency changes), and with similar magnitude in B1 (still with only two statistically significant changes), thus substantially increasing the range across the scenarios. C and D are the most common clusters in A1FI and A2, together accounting for more than 46% of all days in 2090s A1FI, compared with around 16% in GCM20c. In the A1FI scenario, by the 2090s three clusters (F, G and I) nearly entirely disappear, with mean frequencies less than 0.5%; most other winter clusters are less than half as frequent as they are in GCM20c.

Changes in seasonality by the 2090s are shown in Figure 3. Again, in most cases the 2090s B1 scenario most closely resembles GCM20c, and the A1FI and A2 scenarios present the largest shifts. Although not changing in overall frequency, summer clusters A and B change substantially in seasonality. Cluster A becomes more concentrated in the core of summer, with a distinct peak in July and August, with decreased occurrences in May, June, September and October. Cluster B complements A by occurring with greater frequency throughout the year, occurring somewhat less in August (due to the increase in A) but substantially more into the autumn. Clusters C and D, the clusters with the largest overall increases, show large increases in their seasonal extent. In the 20th century, C occurs mostly in spring and D in autumn; by the 2090s both occur significantly more often in both spring and autumn, and D also has a significant presence in the summer. The remaining clusters do not shift substantially in their seasonality; their increases (cluster K) or decreases (all others) generally reflect more or less equal changes throughout the portion of the year in which they occur. Seasonal changes in the 2050s (not shown) are generally midway in between GCM20c and the 2090s for A1FI and A2, and roughly the same as the 2090s in B1.

In comparing the cluster mean-temperature fields within each scenario, as with all other results, greater discrepancies are observed further into the future, and with the more carbon-intensive scenarios. For A1FI (Table II), the mean Pearson correlation coefficient between the GCM20c and 2050s temperature fields across all 15 clusters is 0.992, falling to 0.984 by the 2090s. Average mean bias increases, to 1.94 °C in the 2050s and 3.51 °C in the 2090s. The other two scenarios (not shown) contain better fits. Indeed, for the B1 scenario, the mean Pearson correlation coefficient in comparing cluster mean maps between GCM20c and the 2090s (0.997) is identical to the mean generated in comparing NNR20c to GCM20c, although the mean bias across the clusters (1.24 °C) is somewhat larger.

The increase in mean bias, with correlation coefficients that are still high, suggest that the clusters are likely being affected by days with a similar spatial temperature pattern, but with all temperatures above historical levels, creating the larger error. Again, Figure 4 shows the range of progression of all the patterns (represented by the extremes in patterns D, C and G) through to the future, displaying the corresponding clusters created in the 2050s and 2090s of the A1FI emissions scenario after the sixth step of the process. Despite an unsurprising warming trend, qualitatively, the classification method used is able to categorize future patterns of 850-hPa temperatures that are quite similar in shape to historic patterns, although with each of the three patterns depicting a stronger thermal ridge in the western United States.

4. Discussion and conclusions

A critical component of utilizing synoptic climatological classifications as a downscaling tool in projecting future impacts of climate change is being able to adequately replicate the observed atmospheric patterns with a GCM. The method demonstrated herein successfully reproduced NNR20c classifications of 850-hPa temperatures with the CCSM3 GCM in the historical period – in shape, frequency and seasonality – a finding that has been achieved with mixed results in many studies. For example, in using Lamb weather types in analyzing SLP patterns over Europe, Demuzere *et al*. (2009) finds that although the GCM is capable of replicating the patterns for most of the cold season, significant differences are found in the GCM circulation types over the warm season, preventing further analysis of this season. Further, in studying Arctic SLP patterns with the SOM method, Cassano *et al*. (2006) found only moderate summer coherence between reanalysis and simulated map patterns when examining a number of different GCMs.

While the results in this study also show that the model is more adept at replicating cluster frequency in the winter and transition months *versus* the summer, this is not always the case. Schoof and Pryor (2006) use Kirchhofer classifications of 500-hPa geopotential height fields from the NNR20c data set to evaluate the performance of two GCMs in replicating synoptic patterns. Results indicate that in one GCM, the most commonly occurring pattern is overestimated at the expense of the next two most commonly occurring patterns, while in the other GCM, the most common pattern is produced too seldom during the winter and spring months. And, in evaluating synoptic SLP patterns in the Pacific Northwest, McKendry *et al*. (2006) found that, out of the 13 patterns defined, 3 cold types are underpredicted, and 3 warm/wet types are overpredicted, with the trend most noticeable in the winter.

In the present study, clusters containing a large thermal ridge in the west and relatively warmer temperatures throughout are generally projected to increase in frequency and broaden in seasonality. Additionally, the frequency of a winter cluster featuring a strong cold front approaching the Atlantic coast (cluster K) is also projected to increase into the future – possibly indicating

that this area may be affected by colder winter extremes in the 21st century. Both of these trends are amplified under the higher emissions scenarios (A1FI and A2), more subdued under the most environmental-friendly scenario (B1), and are increasing over time into the 2090s – responding to future climate scenarios in a manner to be expected from previous literature. Although looking at an entirely different region, Hope (2006) notes that in examining many GCMs, each in two different scenarios, that trends in mean SLP patterns favour an increase in high pressure systems and fewer troughs in southwest Western Australia, especially in the A2 scenario *versus* B1. While exploring potential future heat-related mortality in Chicago, Hayhoe *et al*. (2010) find that oppressively warm weather days over the area may double under the B1 scenario, while possibly quadrupling in the A1FI scenario by the end of the century. Although neither of these studies looked specifically at 850-hPa temperature classifications, both imply a trend towards warmer synoptic types under higher emissions scenarios further into the future.

An interpretation of the results must incorporate a set of assumptions inherent in any synoptic climatological research. Daily classifications of any weather variable must assume that the atmosphere can be appropriately partitioned into discrete intervals represented by a once-daily classification. Additionally, in this particular case, interpretation of temperature at the 850-hPa level must be viewed with caution across the Rocky Mountains, where the 850-hPa level intersects with the earth's surface in places – hence the reason that 1200 UTC is chosen, to minimize the impact of daytime heating. Further, a certain level of within-cluster variability must be accepted, as no single cluster can possibly account for all the variability of every day assigned to it. Another goal of synoptic climatology is to develop a classification which maximizes between-cluster variability. These metrics can be evaluated in a number of ways, including using an application of a surface environment variable such as temperature or precipitation and examining how well these variables are partitioned across the clusters (Beck and Philipp, 2010), or by using a skill score with a dichotomous variable such as the days of occurrence of a certain event (Schieneman and Frei, 2010). Each of these performance examinations is highly dependent on a number of different flexible parameters inherent in all synoptic classifications – such as the number of clusters chosen – and is dependent on each unique application.

In assessing the validity of future projections, the results of the diagnostic tests done on the clusters suggest that in some cases in the future, the synoptic patterns that result are not expected to diverge significantly from those that presently occur. That mean bias increases and correlation coefficients decrease substantially in the more extreme scenarios in the more distant future (i.e. A1FI and A2 in the 2090s), however, suggests that the breadth of atmospheric patterns that comprise the set of clusters may not be sufficient to describe these cases in the future. The increase in mean bias also complicates the assessment of the importance of changes in cluster frequency compared with changes in the mean pattern for each cluster. A possible solution would be to include the GCM-Future data into the initial PCA and cluster analysis – this, however, may lead to clusters that occur only in the future and rarely/never in the 20th century, making it more difficult to examine trends in frequencies of a particular event of interest. Thus, within the theoretical framework of synoptic climatology, it would be difficult to assess the efficacy of creating additional clusters for future weather conditions that have not occurred in the historical record.

Many published studies utilizing a synoptic methodology with future projections do not fully report how well historical patterns created by the GCM resemble those of a reanalysis data set; moreover, most studies do not evaluate how similar the days classified to a certain cluster in the future resemble the days that comprise the historical cluster. Thus, while the current method is largely successful for the variable and GCMs described above, it is important to note that mixed results in other research indicates that the study area, variables and GCMs chosen all play a role in model performance, making confirmation of model ability necessary for every applied study. Along that chord, to test the robustness of the six-part method presented herein, two additional variables have also been examined (500- and 700-hPa geopotential heights) along with a second GCM (the third generation Coupled Global Climate Model from the Canadian Centre for Climate Modelling and Analysis; CGCM3), with similar trends observed (results not shown due to space considerations). The consistency in these initial results with this six-part procedure across multiple permutations helps to further validate the technique, and lends support to further use with GCM output in applied work and future impacts studies.

References

Anthes RA, Corell RW, Holland G, Hurrell JW, McCracken MC, Trenberth KE. 2006. Hurricanes and global warming – potential linkages and consequences. *Bulletin of the American Meteorological Society* **87**: 623–628.

Beck C, Philipp A. 2010. Evaluation and comparison of circulation type classifications of the European domain. *Physics and Chemistry of the Earth* **35**: 374–387.

Cassano JJ, Uotila P, Lynch A. 2006. Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part 1: arctic. *International Journal of Climatology* **26**: 1027–1049.

Castro CL, Pielke RA, Leoncini G. 2005. Dynamical downscaling: assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). *Journal of Geophysical Research* **110**: D05108.

Cheng CS, Auld H, Li G, Klaasen J, Li Q. 2007a. Possible impacts of climate change on freezing rain in south-central Canada using downscaled future climate scenarios. *Natural Hazards and Earth System Sciences* **7**: 71–87.

Cheng CS, Campbell M, Li Q, Li G, Auld H, Day N, Pengelly D, Gingrich S, Yap D. 2007b. A synoptic climatological approach to assess climatic impact on air quality in south-central Canada. Part I: historical analysis. *Water, Air, and Soil Pollution* **182**: 131–148.

Cheng CS, Campbell M, Li Q, Li G, Auld H, Day N, Pengelly D, Gingrich S, Yap D. 2007c. A synoptic climatological approach to assess climatic impact on air quality in south-central Canada. Part II: future estimates. *Water, Air, and Soil Pollution* **182**: 117–130.

Cheng CS, Li G, Li Q. 2010. A synoptic weather typing approach to simulate daily rainfall and extremes in Ontario, Canada: potential for climate change projections. *Journal of Applied Meteorology and Climatology* **49**: 845–866.

Christensen JH, Carter TR, Rummukainen M, Amanatidis G. 2007. Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Climatic Change* **81**: 1–6.

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JA, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD. 2006. The Community Climate System Model Version 3 (CCSM3). *Journal of Climate* **19**: 2122–2143.

Cuell C, Bonsall B. 2009. An assessment of climatological synoptic typing by principal component analysis and k-means clustering. *Theoretical and Applied Climatology* **98**: 361–373.

Demuzere M, Werner M, van Lipzig NPM, Roeckner E. 2009. An analysis of present and future ECHAM5 pressure fields using a classification of circulation patterns. *International Journal of Climatology* **29**: 1796–1810.

Enke W, Schneider F, Deutschländer T. 2005. A novel scheme to derive optimized circulation pattern classifications for downscaling and forecast purposes. *Theoretical and Applied Climatology* **82**: 51–63.

Ghosh S, Mujumdar PP. 2006. Future rainfall scenario over Orissa with GCM projections by statistical downscaling. *Current Science* **90**: 396–404.

Hayhoe K, Sheridan SC, Kalkstein LS, Greene JS. 2010. Climate change, heat waves, and mortality projections for Chicago. *Journal of Great Lakes Research* **36**: 65–73.

Hewitson BC, Crane RG. 2006. Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *International Journal of Climatology* **26**: 1315–1337.

Hope PK. 2006. Projected future changes in synoptic systems influencing southwest Western Australia. *Climate Dynamics* **26**: 765–780.

Huth R. 2000. A circulation classification scheme applicable in GCM studies. *Theoretical and Applied Climatology* **67**: 1–18.

Huth R. 2010. Synoptic-climatological applicability of circulation classifications from the COST733 collection: first results. *Physics and Chemistry of the Earth* **35**: 388–394.

Huth R, Beck C, Philipp A, Demuzere M, Ustrnul Z, Cahynová M, Kyselý J, Tveito OE. 2008. Classifications of atmospheric circulation patterns: recent advances and applications. *Annals of the New York Academy of Sciences* **1146**: 105–152.

IPCC. 2007a. Summary for policymakers.In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge, New York.

IPCC. 2007b. In *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge, New York.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.

Lynch A, Uotila P, Cassano JJ. 2006. Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part II: Antarctic. *International Journal of Climatology* **26**: 1181–1199.

McKendry IG, Stahl K, Moore RD. 2006. Synoptic sea-level pressure patterns generated by a general circulation model: comparison with the types derived from NCEP/NCAR re-analysis and implications for downscaling. *International Journal of Climatology* **26**: 1727–1736.

Meehl GA, Arblaster JM, Tebaldi C. 2005. Understanding future patterns of increased precipitation intensity in climate model simulations. *Geophysical Research Letters* **32**: L18719.

Philipp A, Bartholy J, Beck C, Erpicum M, Esteban P, Fettweis X, Huth R, James P, Jourdain S, Kreienkamp F, Krennert T, Lykoudis S, Michalides SC, Pianko-Kluczynska K, Post P, Alvarez DR, Schiemann R, Spekat A, Tymvios FS. 2010. Cost733cat – a database of weather and circulation type classifications. *Physics and Chemistry of the Earth* **35**: 360–373.

Pielke RA, Landsea C, Mayfield M, Laver J, Pasch R. 2005. Hurricanes and global warming. *Bulletin of the American Meteorological Society* **86**: 1571–1575.

Pielke RA, Landsea C, Mayfield M, Laver J, Pasch R. 2006. Reply to "Hurricanes and global warming – potential linkages and consequences.". *Bulletin of the American Meteorological Society* **87**: 628–631.

Saunders IR, Byrne JM. 1996. Generating regional precipitation from observed and GCM synoptic-scale pressure fields, southern Alberta, Canada. *Climate Research* **6**: 237–249.

Saunders IR, Byrne JM. 1999. Using synoptic surface and geopotential height fields for generating grid-scale precipitation. *International Journal of Climatology* **19**: 1165–1176.

Schoof JT, Pryor SC. 2006. An evaluation of two GCMs: simulation of North American teleconnection indices and synoptic phenomena. *International Journal of Climatology* **26**: 267–282.

Sheffield J, Wood EF. 2008. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Climate Dynamics* **31**: 79–105.

Shepherd JM, Knutson T. 2007. The current debate on the linkage between global warming and hurricanes. *Geography Compass* **1**: 1–24.

Sheridan SC, Lee CC. 2010. Synoptic climatology and the general circulation model. *Progress in Physical Geography* **34**: 101–109.

Sheridan SC, Lee CC. 2011. The self-organizing map in synoptic climatological research. *Progress in Physical Geography* **35**: 109–119.

Schieneman R, Frei C. 2010. How to quantify the resolution of surface climate by circulation types: an example for Alpine precipitation. *Physics and Chemistry of the Earth* **35**: 403–410.

Schwartz MD. 1996. An air mass-based approach to regional GCM validation. *Climate Research* **6**: 227–235.

Spekat A, Kreienkamp F, Enke W. 2010. An impact-oriented classification method for atmospheric patterns. *Physics and Chemistry of the Earth* **35**: 352–359.

SPSS Inc. 2001. The SPSS Two-Step cluster component: a scalable component to segment your customers more effectively. White Paper, Technical Report, SPSS, Chicago, IL.

Tveito OE. 2010. An assessment of circulation type classifications for precipitation distribution in Norway. *Physics and Chemistry of the Earth* **35**: 395–402.

Vrac M, Hayhoe K, Stein M. 2007. Identification of intermodal comparison of seasonal circulation patterns over North America. *International Journal of Climatology* **27**: 603–620.

Wetterhall F, Bárdossy A, Chen D, Halldin S, Xu C. 2009. Statistical downscaling of daily precipitation over Sweden using GCM output. *Theoretical and Applied Climatology* **96**: 95–103.

Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO. 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. IPCC task group on data and scenario support for impact and climate analysis (TGICA). Available at http://www.ipcc-data.org/guidelines/dgm_no2_v1_09_2004.pdf (Accessed 27 June 2011)

Yarnal B, Comrie AC, Frakes B, Brown DP. 2001. Developments and prospects in synoptic climatology. *International Journal of Climatology* **21**: 1923–1950.

Yarnal B. 1993. *Synoptic Climatology in Environmental Analysis: A primer*, Belhaven Press: London.

