

Identification and intermodel comparison of seasonal circulation patterns over North America

M. Vrac,^{a,*} K. Hayhoe^{b,c} and M. Stein^d

^a Center for Integrating Statistical and Environmental Science, The University of Chicago, Chicago, IL 60637

^b Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801

^c Department of Geosciences, Texas Tech University, Lubbock, TX 79409

^d Department of Statistics, The University of Chicago, Chicago, IL, 60637

Abstract:

Shifts in the frequency of atmospheric circulation patterns throughout the year can characterize seasonality and provide a means to evaluate the ability of atmosphere-ocean general circulation models (AOGCMs) to reproduce key dynamical structures influencing regional climate. However, the identified characteristics of these patterns can depend on the clustering method used, the physical properties of the model generating the circulation fields, and even their spatial scale. Using two statistical clustering methods, a mixture model and a hierarchical method, we show that these factors can result in distinctly different atmospheric patterns characterizing seasonal circulation over eastern North America based on NCEP and ERA-40 geopotential height and sea level pressure (SLP) fields. Consistency is improved through constraining the number of clusters at each level to 4 or 5, and using the mixture model method. In general, this method tends to produce more consistent results across the various datasets and is more sensitive to day-to-day variations in pattern frequencies than the traditional hierarchical method. Applying the mixture method to PCM and CCSM3 simulations reveals some clear differences relative to reanalysis-based patterns. In particular, the AOGCMs do not reproduce the observed duration and/or location of summer, with seasonal shifts (PCM) and summer extensions (CCSM) that are strongest at lower levels. However, taking into account the uncertainty introduced by the different factors, these AOGCMs do successfully capture many of the observed large-scale drivers of seasonality. These results lend support to circulation-based downscaling, but also highlight some systematic model biases, and hence the ongoing potential for improvements in model parameterization and dynamics. Copyright © 2006 Royal Meteorological Society

KEY WORDS circulation patterns; seasonal shifts; clustering methods

Received 31 January 2006; Revised 14 June 2006; Accepted 31 July 2006

INTRODUCTION

Large-scale atmospheric circulation patterns correspond to shapes or structures that tend to repeat over time. These patterns are the primary drivers of day-to-day and inter-annual variations in surface climate, bringing with them local to regional-scale changes in temperature, humidity, precipitation, and other features of surface weather that are characteristic of that given pattern (Salinger and Mullan, 1999; Aizen *et al.*, 2001; Slonosky *et al.*, 2001; Burnett *et al.*, 2004; etc.). Seasonal circulation patterns are also one of the primary intermediaries connecting solar forcing with surface-level climate to produce the familiar characteristics of the four seasons over higher latitudes such as the North American continent. For example, the polar front jet stream is a pattern of high-altitude winds that separates colder, drier air to the north from warmer, moister air to the south. It grows stronger and moves southward during Northern Hemisphere winter,

then weakens and shifts northward during summer. The position and intensity of the jet affects the dominant circulation patterns and the associated seasonality over North America.

Over larger spatial scales (continental to hemispheric), atmospheric circulation is viewed as one of the more robust outputs from atmosphere-ocean general circulation models (AOGCMs), after mean temperature (Covey *et al.*, 2003). For this reason, atmospheric circulation patterns are often used in downscaling approaches such as 'weather typing' which derive projected changes in surface climate from AOGCM-simulated changes in upper-air characteristics. To what degree, however, are AOGCM-based patterns consistent with reanalysis-based circulation patterns? Defining these structures, understanding their connection to surface climate, and assessing the ability of AOGCMs to reproduce them are essential steps to effectively applying statistical relationships to AOGCM output to resolve future change at the local to regional scale.

Large-scale patterns in atmospheric circulation can be identified subjectively or objectively, through clustering and other pattern recognition methods. The most famous

* Correspondence to: M. Vrac, Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL) Laboratoire CEA/CNRS/UVSQ Centre d'Etudes de Saclay, Orme des Merisiers, Bat. 701 91191 Gif-sur-Yvette, France. E-mail: mathieu.vrac@cea.fr

example of the subjective approach is the well-known Lamb Weather Types for the British Isles identified by Lamb (1972). Approaches such as this imply prior knowledge regarding the location and extent of the patterns. In contrast, an objective approach determines patterns from gridded data using clustering methodologies. Some recent examples include the hierarchical descending clustering method used to identify three weather states in the Columbia River Basin for the winter season (Zorita *et al.*, 1993); a hierarchical agglomerative clustering method used to derive a climatology of severe storms in Virginia (Davis *et al.*, 1993) or to define climate regions in the northern plains (Bunkers *et al.*, 1996); a “*K*-means” clustering procedure used to define circulation types over Europe (Huth, 2001); a mixture of copula functions applied to determine air mass types from vertical atmospheric profiles (Vrac *et al.*, 2005); and finally, a fuzzy rule approach to modelling monthly patterns of precipitation in Hungary (Pongracz *et al.*, 2001).

Objective clustering approaches are based on a number of assumptions, the usefulness and validity of which we examine here. First, in most of these applications, patterns are defined from ‘consistent’ data separated into seasons, or at least into separate winters and summers. In this way, the resulting clusters (and the studies following the clustering steps) are not influenced by data that could falsify season-specific modelling. On the other hand, if the first step is to remove the seasonality of the data, it is then impossible to describe this seasonality and to assess the degree to which independent datasets are able to produce a consistent seasonality in circulation patterns. For this reason, here we use daily datasets covering the entire year. We first investigate whether it is possible to produce a comprehensive overview of annual climatological atmospheric structures that characterize the dominant seasons (winter and summer) and describe the transitions between those seasons (spring and fall) – first from reanalysis fields, as produced by the National Centers for Environmental Prediction (NCEP) and the European Centre for Medium-Range Weather Forecasts (ECMWF ERA-40), and then from historical total (natural + anthropogenic) forcing simulations from two of the National Center for Atmospheric Research (NCAR) AOGCMs, the Community Climate System Model version 3 (CCSM3) and the Parallel Climate Model (PCM) version 1.

We next explore the question of how to define a ‘good’ way to identify these structures. Using a traditional hierarchical agglomerative clustering method (HAC) and a more sophisticated mixture model approach using the expectation maximization (EM) technique, we evaluate the degree to which the structures we find are dependent on the clustering algorithm used. We then assess the implications of statistical clustering and model resolution for the validity and the meaning of these patterns.

Building on our comparison of clustering methods, we examine three secondary uncertainties that affect the patterns identified. The first is the sensitivity of the clusters and their frequencies to model resolution, using two

ERA-40 datasets at 1.125° and 2.5° resolution. The second is the degree to which dominant seasonal circulation patterns vary at four different levels in the atmosphere (500, 700, and 850 mb, and SLP). Seasonality that is more evident at some levels than others has implications for relating surface characteristics for a given season to upper-air forcing. Third, we ask to what degree model parameterization affects the characteristics and seasonal frequency of the patterns identified. Through comparing principal component patterns, densities, and loading values as well as seasonal frequencies of the patterns derived from NCEP and ERA-40, we examine the uncertainty in these observationally constrained patterns and their seasonality to quantify the magnitude of “quasi-observational error”. This is consistent with Jung (2005), who demonstrated that reanalysis output fields still contain systematic errors because of the modelling techniques involved.

Finally, building on the previous analysis which constrained the magnitude of uncertainty based on reanalysis output fields, we examine the patterns and seasonal frequencies resulting from the two AOGCMs to see whether they are successful at reproducing reanalysis-based climatological atmospheric circulation patterns. Although several studies have assessed the ability of AOGCMs to reproduce the seasonal temperature cycle (Covey *et al.*, 2000) and surface climate characteristics (Kunkel and Liang, 2005), this analysis provides a more fundamental check of model ability to simulate the patterns driving surface change.

DATA, MODELS, AND METHODS

In this study, we examine the dominant atmospheric circulation patterns over North America based on two reanalysis datasets and historical simulations from two AOGCMs, combined with two well-known clustering algorithms. The reanalysis data used here are the NCAR/NCEP reanalysis at 2.5° × 2.5° spatial resolution (Kalnay *et al.*, 1996), and the ECMWF/ERA-40 data, for which we have datasets at spatial resolutions of 1.125° × 1.125° and 2.5° × 2.5°, respectively (Källberg *et al.*, 2004). We also identify atmospheric circulation patterns as simulated by historical total-forcing runs from the NCAR/DOE PCM (Washington *et al.*, 2000) and the NCAR CCSM3 (Hurrell *et al.*, 1998; Collins *et al.*, 2006; Hurrell *et al.*, 2006). PCM and CCSM have spectral triangular horizontal resolutions of T42 and T85, which are approximately equivalent to a global average resolution of 2.8° × 2.8° and 1.4° × 1.4°, respectively. For this study, we focus on the region bounded by [26°N, 65°N] × [110°W, 67°W], which covers the eastern half of North America from about the Rockies to the East Coast and from about the Mexican border to the middle of Canada. Our analysis is based on daily data for a 20-year time period from the recent past, ranging from 1 January 1980 through 31 December 1999.

Daily fields of geopotential heights (*Zg*) are among the variables most used in the literature to define atmospheric

patterns. Here, we cluster daily fields of Zg at multiple pressure levels (850, 700, and 500 mb) and daily SLP fields separately to determine the atmospheric structures dominant in each. Weather states can then be compared between pressure levels, between individual datasets of the same type, and between reanalysis data and simulated model output.

To define the atmospheric patterns, we first employ a mixture of statistical distributions (Pearson, 1894). Although generally considered as a very effective clustering method, the mixture approach has rarely been applied to atmospheric studies. In this approach, we estimate g , the probability density function (pdf) of one of the variables (e.g. Zg850), as a weighted sum or *mixture* of K parametric pdfs g_i ($i = 1, \dots, K$) with parameters α_i . As described in Appendix A, an Estimation-Maximization (EM) type algorithm (Dempster *et al.*, 1977; for review, see McLachlan and Peel, 2000) is used to estimate the parameters of the mixture model; hence, we refer to this method as the EM approach.

The mixture model using the EM algorithm is used to cluster the 4 data sets (NCEP, ERA40, PCM, CCSM) on each variable (Zg850, Zg700, Zg500, and SLP) separately. Because the region over North America described above corresponds to a large number of grid cells (from 240 up to 1368, depending on the spatial resolution of each data set), which would require estimation of an unwieldy number of dimensions for each pdf, an unrotated principal components analysis (PCA) is first applied to each variable and data set to reduce the dimension of the problem while keeping the main part of the variance of the data.

The data were first standardized such that the mean for each statistical variable was set to 0 and its variance to 1. In this way, we avoid the chance that any of the statistical variables might artificially assume a greater variance than the others. For each principal component, the spatial map of the PCA loading values has been checked to ensure that they produced plausible structures, with plots available online for the interested reader (see website – Vrac and VanDorn, 2006).

For each applied PCA, we retain the first C components, where C is chosen such that these components contain at least 99 percent of the cumulative variance but the first $C - 1$ components do not. Next, we run the EM algorithm and select the number of clusters K using the Bayesian Information Criterion (BIC). In general, for the four variables Zg500, Zg700, Zg850, and SLP, C corresponds respectively to the first 22, 17, 15 and 31 components for the reanalysis data sets (NCEP and ERA-40); slightly fewer components are required for the modelled data sets (CCSM and PCM).

In addition to the EM clustering method based on mixture models, we also apply an alternative clustering method to each data set, the more commonly-used Hierarchical agglomerative clustering (HAC) method. The HAC approach does not assume any (explicit) distribution for the variables. Starting with the items to be clustered (i.e. daily fields) as groups with only one

element (singletons), at each step the HAC regroups two groups according to a given criterion, until only one group remains, comprising all the items. The criterion that we employ here to bring two groups together is the Ward criterion (Ward, 1963) that maximizes the interclass variance (i.e. minimizing the intraclass variance) at each step. For the HAC method, other criteria could have been used instead of Ward's. However, we justify its exclusive use here for several reasons. First, this criterion procedure is one of the most (if not *the* most) commonly-employed criterion in HAC methods. Moreover, it is a statistically-based criterion. This is not clearly the case for the other criteria classically used in HAC, such as minimum, complete or average linkage. Finally, because we also want to look at how a number of other factors affect the analysis, we do not want to use too many clustering methods here, and comparison of the model-based approach with the Ward's criterion procedure is sufficient to illustrate differences between seasonal patterns provided by EM or HAC methods.

The patterns obtained from this alternative approach provide valuable information regarding the sensitivity of the clustering results to the method used. As for the EM approach, we apply the HAC method with Ward criterion to the same C principal components (PC). For HAC, the optimal number of clusters can be estimated through a visual 'elbow' criterion rather than a BIC criterion as for the EM algorithm. For instance, plotting the evolution of the intraclass variance according to the number of patterns produces an 'elbow' that arises when the next merging step (i.e. the next lower number of patterns) would result in a sharp rise of the intraclass variance. That is, for a higher number of clusters, the criterion does not change substantially between clusters and the larger number of individual clusters will not contain as much independent information as would a smaller number of clusters.

IDENTIFYING THE DOMINANT SEASONAL ATMOSPHERIC PATTERNS

Independent clustering for each data set and vertical level using the EM algorithm with the BIC, or the HAC with the elbow criterion, results in an optimal number of clusters for each individual reanalysis/vertical level/clustering algorithm combination that ranges from 2 to 5. Table I lists the optimal number of clusters determined by the EM algorithm according to BIC and by the HAC algorithm using the elbow criterion, for each reanalysis data set and each available level (Zg500, Zg700, Zg850, and SLP).

Although the number of clusters differs from one data set to another, the relative number of clusters for each level is consistent for a given clustering method. Specifically, in Table I we see that the BIC criterion seems to prefer relatively fewer patterns for the lower levels (SLP and Zg850) and more patterns for the higher levels (Zg700 and Zg500). At lower levels in the atmosphere, dynamics are more strongly affected by turbulence at

Table I. Optimal number of clusters from the EM algorithm as determined by the BIC and from the HAC determined by the visual elbow criterion, for each reanalysis data set and each available variable, where ‘–’ corresponds to a level at which the data was unavailable.

Method	Data set	SLP	Zg850	Zg700	Zg500
EM (BIC)	NCEP 2.5	3	5	5	5
	ERA40 1.125	4	4	5	4
	ERA40 2.5	–	4	5	5
	NCEP 2.5	5	3	3	2–3
HAC (elbow)	ERA40 1.125	4	4	3–4	3–4
	ERA40 2.5	–	4	2	2–4

the planetary boundary layer and topographic features. Hence, variability at SLP and 850 mb is much stronger than at higher levels, which tend to be dominated by more steady zonal flows. This is illustrated in Figure 1 by the first (mainly winter) and last (mainly summer)

patterns from the 4 SLP patterns and the 5 Zg500 patterns as determined by the EM method. Figure 1 (a) and (b) present these patterns for NCEP and (c) and (d) for ERA-40.

At first glance, the smaller number of patterns for lower levels found by the EM approach seems counter-intuitive. If variability is larger, should there not be more patterns? For these lower levels, if we want to capture *all* of the variability by means of our clusters, we would require a very large number of patterns, which in turn implies a large number of parameters in the mixture model. However, the BIC is actively trying to *prevent* the use of too many parameters. Hence, the large number of clusters that would be needed to capture the whole variability at lower levels is not reached. Instead, the BIC gives us a relatively small number of patterns that explains only the main features of SLP, regrouping several common characteristics into more general patterns. This also explains why the SLP patterns tend to vary much more from one data source to another as compared with

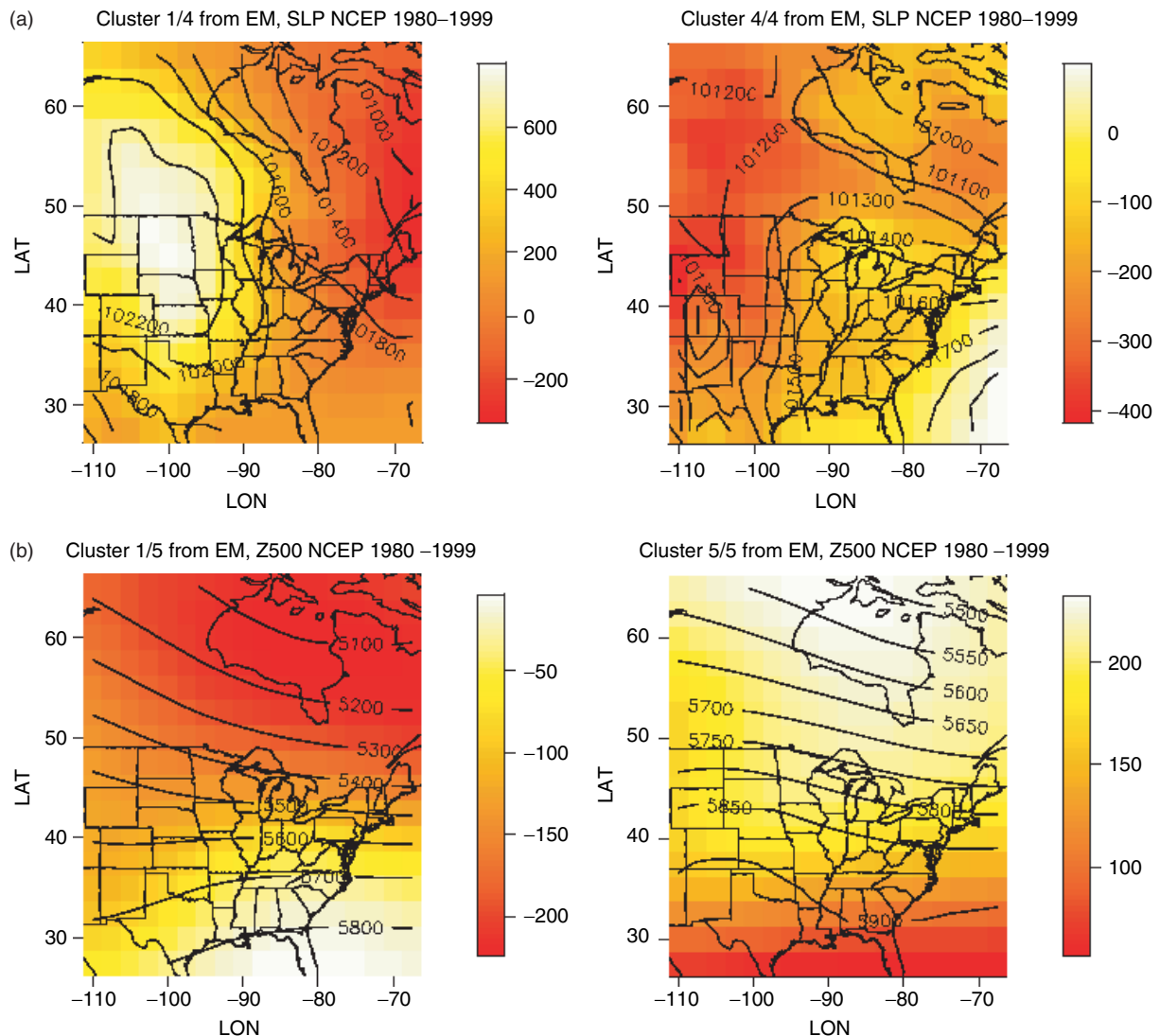


Figure 1. First (winter) and last (summer) patterns for NCEP SLP (a), NCEP Zg500 (b), ERA40 1.125° SLP (c) and ERA40 1.125° Zg500 (d) from EM.

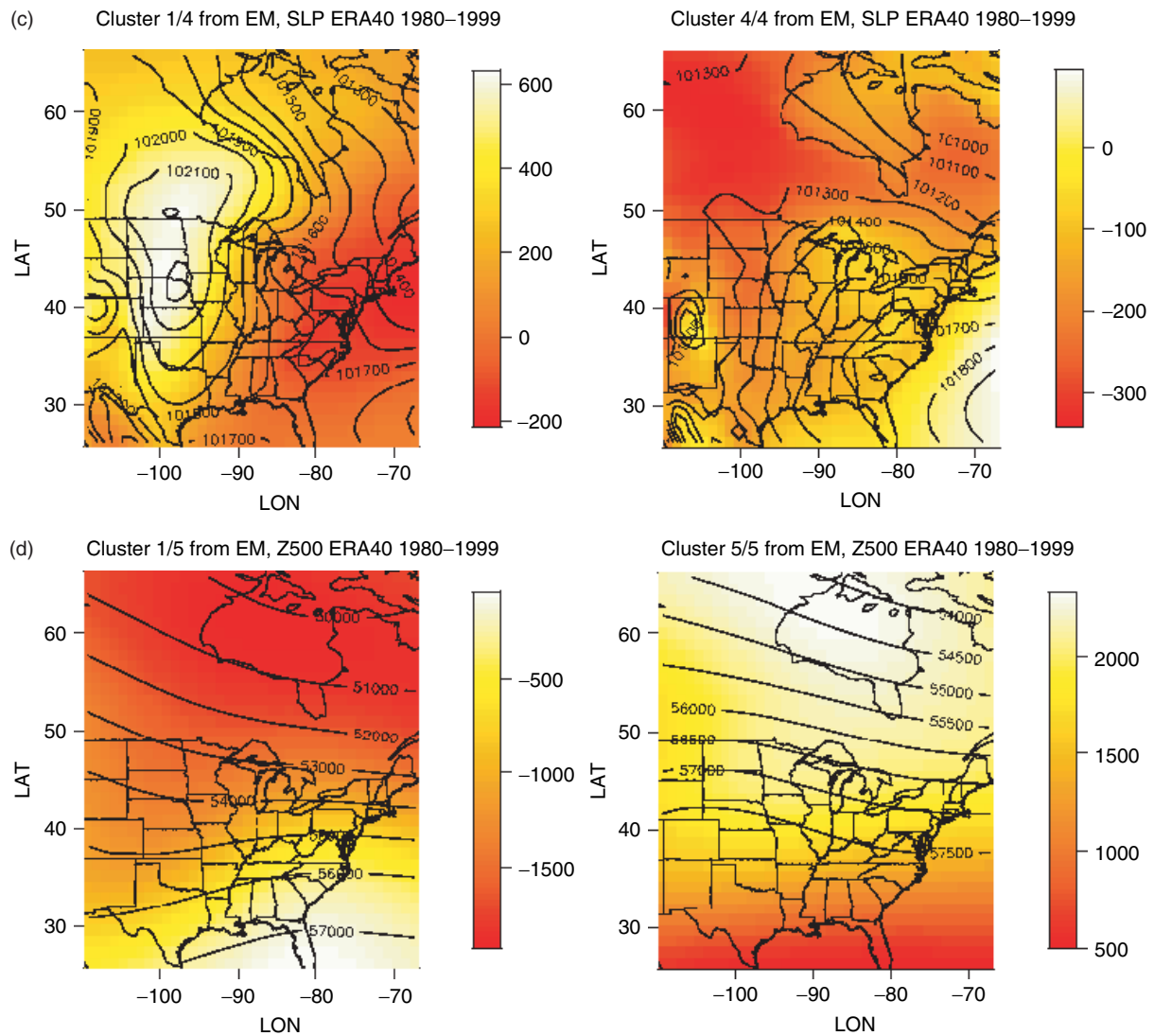


Figure 1. (Continued).

the upper-air patterns. Rather than representing a single consistent pattern, because of its high variability the SLP patterns are actually conglomerates of several different patterns. For a given dataset, the smaller-scale temporal patterns are likely combined in different, inconsistent ways to produce 3 to 5 primary clusters for one dataset that do not appear to bear much (if any) resemblance to those found for another dataset (Figure 1, compare the differences between a and c with the similarities between b and d).

In Table I, we can also see that the HAC method (with the elbow criterion) gives us a smaller “optimal” number of patterns than the EM approach constrained using the BIC. It does not display as clear a difference between higher- and lower-level numbers of clusters. As our goal is to characterize each primary season as well as the transitional seasons by at least one pattern, it is clear that 2 or 3 clusters are too few. In contrast, higher numbers of clusters tend to result in several clusters that do not contain relevant information. Multiple iterations with free and constrained values of K reveal that introduction of

more clusters produces clusters that either represent mean zonal east–west flow (as seen from plotting the clusters themselves), which do not exhibit a strong seasonality but remain fairly constant from month-to-month, or mimic previously-identified clusters and their seasonality for that same dataset/level combination (as revealed by monthly histograms).

To facilitate comparison between data sets and models, it is necessary to constrain the number of clusters identified at each level to a reasonable number that can be well-represented across all data sets. Ideally, these clusters should contain an optimal amount of information that captures the seasonality of circulation patterns over North America while remaining relatively independent. For this reason, through iteration on the reanalysis data sets and examination of the resulting cluster patterns and monthly histograms of the frequency of each pattern for each level, we have determined optimal numbers of clusters for each vertical level to be 5/5/5/4 for Zg500/Zg700/Zg850/SLP. We also verified that increasing the number of clusters at each level by

just one, to 6/6/6/5, did not contribute any additional information previously lacking to characterize seasonality over the region and period of interest (1980–1999).

For the remainder of this analysis, therefore, all comparisons of structures are based on this consistent number of patterns. Furthermore, we assume that we must retrieve equivalent structures and information from AOGCM simulations with the same number of patterns. If we detect more, fewer, or different patterns in the AOGCM data using the aforementioned optimization criterion, we can conclude that they are ‘false’ or arise from characteristics

of the model that are not verified by the observationally-constrained reanalysis fields.

When the number of clusters is constrained, comparison between NCEP data (at 2.5°) and ERA-40 (at 1.125°) shows greatly improved consistency across all levels and patterns. This is illustrated in Figure 2, which shows the five patterns of the 700-mb level for NCEP and ERA-40 1.125° . Transitional patterns (spring, fall) tend to be clearer at higher levels, while winter is clearer at lower levels and the single summer pattern is dominant throughout all levels for July and August. Although there

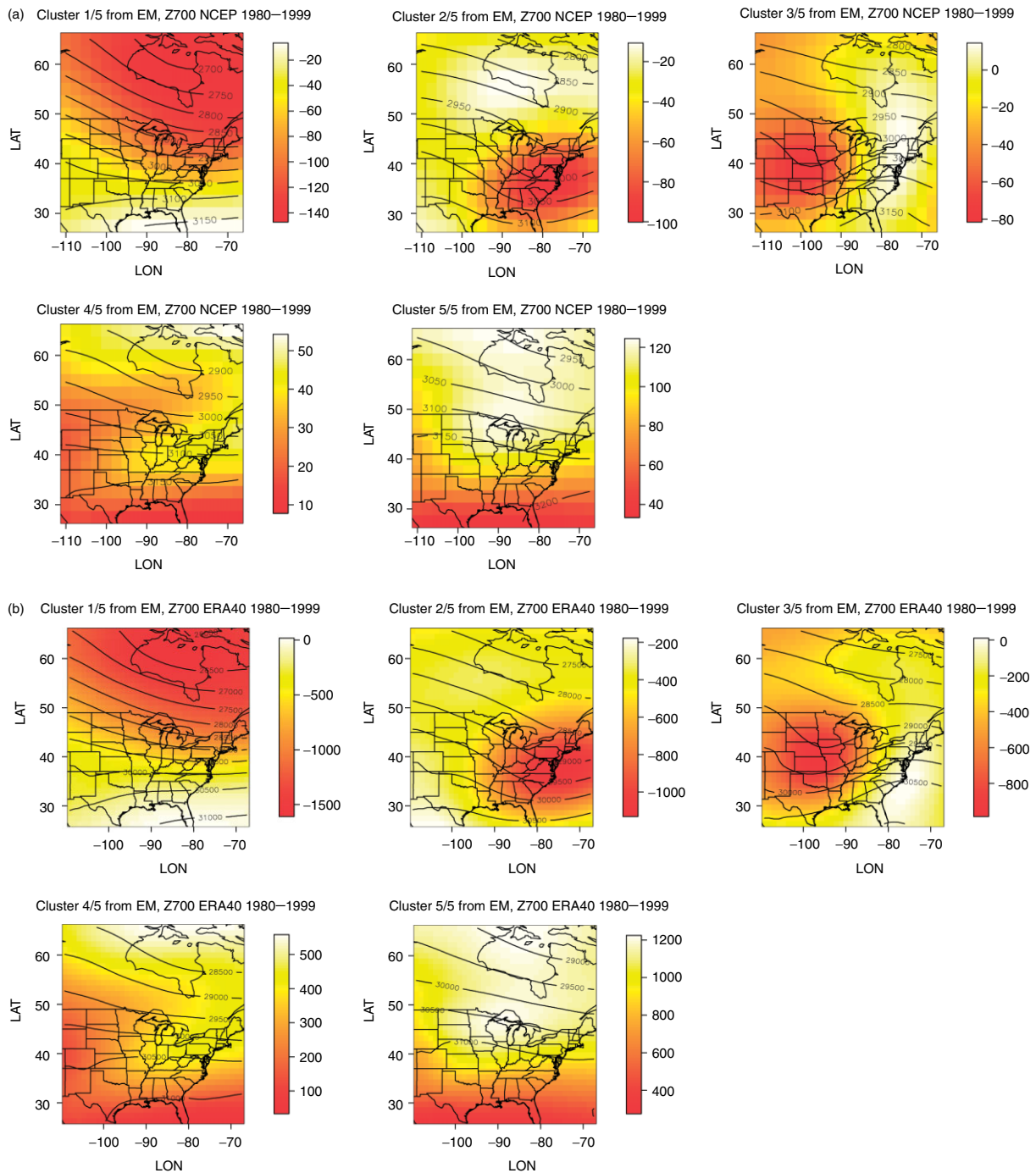


Figure 2. Five primary seasonal patterns identified using the EM algorithm for NCEP Zg700 (a) and ERA40 1.125° Zg700 fields. Coloured contours represent anomalies and black lines the mean geopotential heights for that pattern.

are certainly differences between the NCEP and ERA-40 patterns, the general appearance of each is quite similar. This is true for both the anomalies (coloured contours) and the average geopotential height for each pattern (black lines).

In summary, we see that objective clustering approaches tend to result in different numbers of optimal clusters when applied to different sets of reanalysis output. These clusters are difficult to compare; however, when the number of clusters is constrained, the resulting clusters show similar patterns and seasonal frequencies.

SENSITIVITY OF PATTERNS TO CLUSTERING METHOD

Because of the nature of these two different clustering methods, EM and HAC, *a priori* we would expect some differences in the clusters identified depending on the method used. For example, the HAC method attempts to define patterns that are centred around their means. So, each day is assigned to the pattern for which the distance between the vector of variables for that day and the vector of means for the pattern is minimized. In other words, HAC will tend to provide us with strong *average* information and a sharp distinction between the patterns as they will have significantly different mean values. In contrast, the EM method takes the variance of the data into account to define patterns according to Equation. (2). That means that a day can belong to more than one pattern at the same time, with different probabilities. This induces a 'fuzzier' distinction between patterns.

As a consequence of the distinctions between these methods, although the HAC method gives us easier patterns to understand, the EM model-based approach allows us to characterize transitions between patterns in a more realistic and probabilistic way. Hence, in the structures resulting from these two methods, we would expect to see a strong temporal signal – i.e. clear frequent patterns for each of the primary seasons (winter and summer) – with HAC, and more variability – i.e. more patterns present with low frequencies – for EM, at least for transition seasons (spring and fall) where the signal is weaker and the variance more important.

The sensitivity of the resulting patterns to the clustering method used is illustrated in Figure 3, which presents the monthly histograms of the five structures for ERA40 1.125° resolution, with empty bars for the EM method and shaded bars for HAC. The number associated with each pattern does not represent any order *per se*, but they have been ordered this way to make the temporal shifts between the patterns obvious, from the main winter patterns (on the left) to the main summer patterns (on the right), with the transition patterns in the middle. The number of structures present in each month is higher for EM as compared with HAC. For that reason, most EM patterns display smaller frequencies than HAC,

except for pattern 4. As expected, HAC shows stronger frequencies for the predominantly winter and summer patterns (1 and 5), while EM produces a more distinct, gradual, and consistent transition from 2 through 4 in the spring and back again in the fall. This characteristic can generally be seen in most data sets and for all levels, with only a few exceptions. First, for ERA-40 1.125°, EM seems to catch a stronger winter signal and a weaker summer signal than HAC does for Zg500 (see website – Vrac and VanDorn, 2006 – for additional figures). For Zg850, it is the opposite, while for SLP, EM identifies both winter and summer patterns that are stronger than HAC, with stronger fall and slightly weaker spring from EM than from HAC (see website Vrac and VanDorn, 2006).

For clusters based on NCEP reanalysis (histograms equivalent to those given in Figure 3 for ERA-40 are shown in Figure 4 for NCEP), the variability inside each month is still larger for EM than for HAC, but the temporal evolution of the monthly histograms is different than that of the ERA-40 clusters (compare Figure 4 with Figure 3). This is true for both EM (empty bars) and HAC (shaded bars). In addition, here the EM algorithm shows both stronger winter/summer frequencies at 700mb as well as the more consistent, gradual transitions as compared to HAC. For NCEP Zg500, EM displays a strong summer signal but a slightly weaker winter as compared with HAC. Similar features are present for Zg850 (see website Vrac and VanDorn, 2006, for additional figures). For SLP, EM captures stronger signals for all four seasons as compared to HAC, even if April and May are slightly weaker for EM (see website Vrac and VanDorn, 2006).

In general, winter and summer patterns and seasonal shifts are well-resolved by both the EM and HAC approaches. However, it appears that the EM method exhibits several improvements over the hierarchical method that is commonly used in many atmospheric applications (Davis *et al.*, 1993; Zorita *et al.*, 1993; Bunkers *et al.*, 1996; Fovell, 1997; Unal *et al.*, 2003; etc.), of which we use HAC as an example. First, more patterns must be prescribed for HAC than would be found using a subjective elbow criterion to capture the full range of seasonal patterns. Furthermore, it appears that although EM may produce slightly lower frequencies for each pattern as it distributes the variability across more patterns in any given month, EM also appears to model the transition between the seasons more smoothly than HAC. In addition, the EM method still captures the distinct dominant seasonal patterns for winter and summer, and its lesser sensitivity to strong mean signals means that it produces more 'realistic' frequencies of patterns. These advantages, combined with the lower sensitivity of the EM method to spatial resolution (see Appendix B) appears sufficient to substantiate its use as the primary clustering method for the AOGCM analysis that follows.

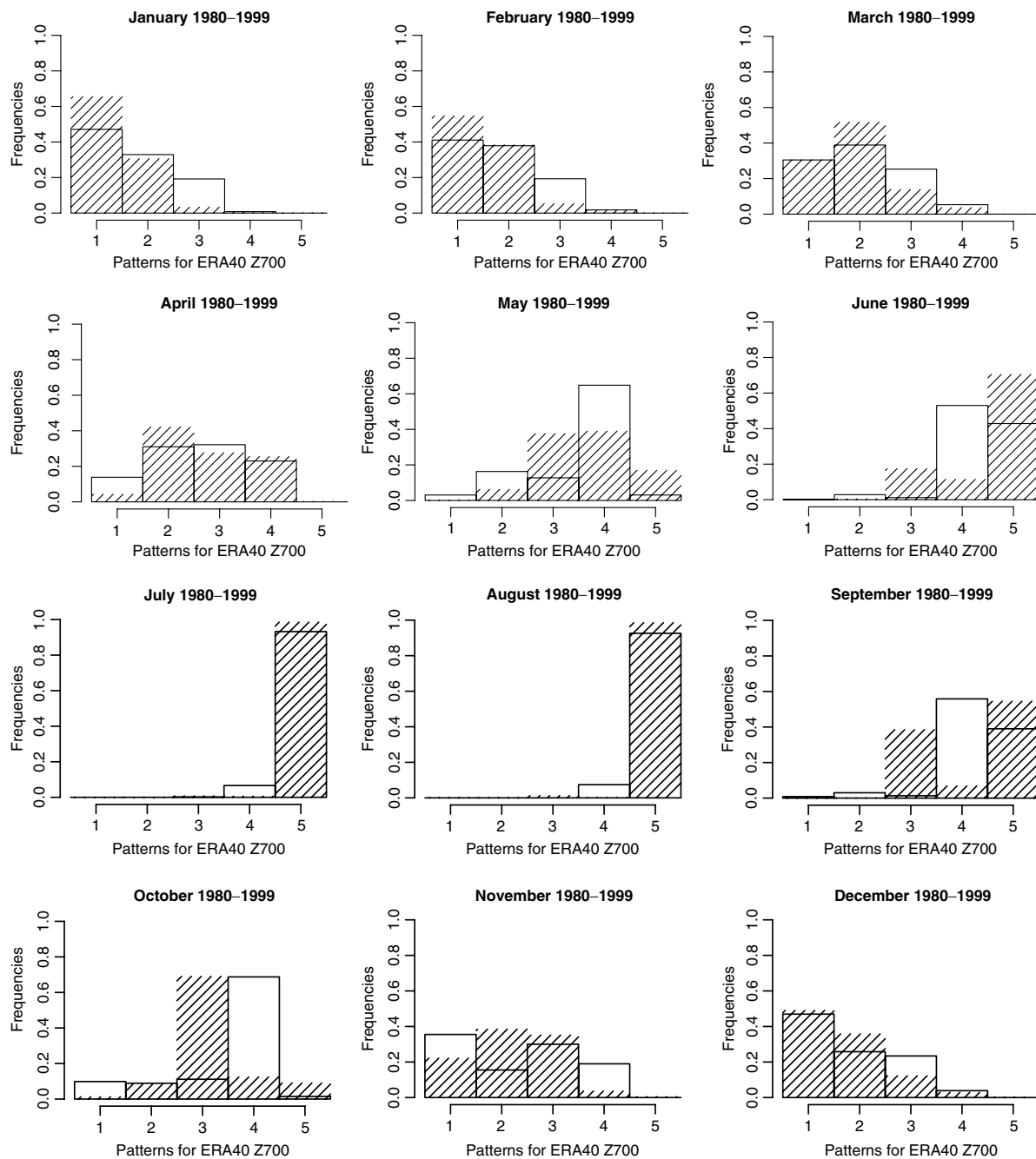


Figure 3. Monthly histograms of the 5 ERA40 Zg700 patterns from EM (empty bars) and from HAC (shaded bars).

ADDITIONAL UNCERTAINTIES DUE TO SPATIAL RESOLUTION AND MODEL FORMULATION

Additional factors may be responsible for the differences in patterns and monthly frequencies observed between the ERA-40 and NCEP data, even when the same clustering method is used. For that reason, we next examine whether part of this discrepancy between the ERA-40 and NCEP patterns (for a given set of patterns based on either EM or HAC) could arise from their different spatial resolution, which is 1.125° for ERA40 and 2.5° for NCEP, or from the way in which the individual reanalysis models assimilate and simulate atmospheric fields.

We first derive new patterns based on ERA-40 data re-gridded to 2.5° spatial resolution, identical to that of NCEP. Comparing the ERA-40 patterns based on the higher and lower resolution data with the NCEP patterns (see Appendix B for details and figures), we see that data resolution does appear to play some role in differentiating between NCEP and ERA-40, particularly at higher levels (using EM) and for the transition and winter seasons (HAC). In general, HAC appears more sensitive to resolution issues than the EM method, particularly for the lower-elevation levels that display stronger correlations with surface climate (at least for humidity variables, Huth, 2005; and temperature, Xoplaki *et al.*, 2003).

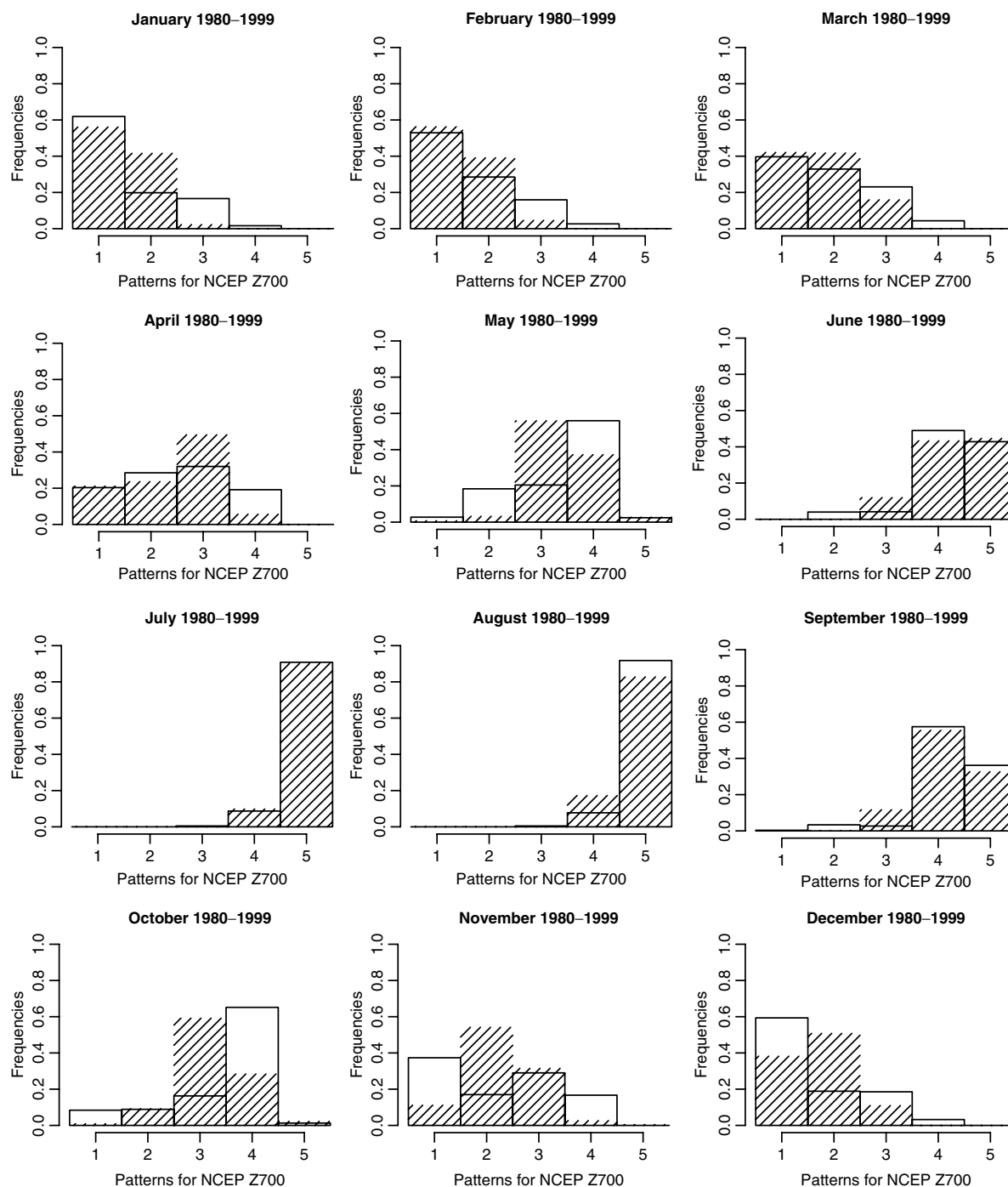


Figure 4. The same as Figure 3, but for NCEP dataset.

Technically, NCEP and ERA-40 reanalysis fields represent the closest estimate we have to the “true” state of the atmosphere on a given day. However, the truth is that these fields are model output, albeit forced by regular surface and upper-air observations. Jung (2005) and others have shown that model formulation is capable of introducing systematic errors into the output. Hence, we next examine the degree to which the differences we have already seen in seasonal patterns and their relative frequencies may be a function of the differences between the NCEP and ERA-40 modelling frameworks.

The analyses of the PCs of the reanalysis data (see Appendix C for details and figures) shows that the variances of the PCs are similar for NCEP and ERA-40 2.5°. However, comparison of the PCs and statistical distributions reveal some significant differences between the PCs simulated by each model. Clear shifts in the mean densities of the PCs generated by each model show that the structures of the NCEP and ERA-40 data are not equivalent. Hence, even when the number of patterns is constrained, the objective clustering methods still define different patterns for NCEP and ERA-40, with slightly different geopotential and temporal properties.

Thus, even for models constrained by real-world observations, it is clear that some uncertainty is introduced by the physical characteristics of both the variable fields (e.g. spatial resolution) as well as the model formulations. This is an important point to consider when we next compare AOGCM-based patterns NCEP and ERA-40 reanalysis-based patterns, as it implies that *a priori* we should expect the AOGCM-based patterns to show some differences as compared with the reanalysis-based patterns simply because of their different spatial resolution and physical parameterizations.

COMPARISON BETWEEN REANALYSIS AND AOGCM-SIMULATED PATTERNS

AOGCM-based clustering patterns

We next apply the EM clustering algorithm to historical total forcing simulations by CCSM3 and PCM, covering the same region and time period as the reanalysis data sets. The HAC method was also applied, but as we have

already shown it to be less stable than the EM method and as the primary results obtained using the HAC approach were not qualitatively different although less striking than those obtained using the EM method, these results are not shown here or discussed further.

As for the reanalysis data, the clustering algorithms are applied separately to daily geopotential heights for the same levels (Zg500, Zg700, and SLP; Zg850 not available) to define a pre-set number of patterns K which, based on reanalysis data, are 5/5/4 for Zg500/Zg700/SLP. In Figure 5, we see the first (winter) and fifth (summer) patterns for CCSM and PCM. Although the patterns are similar, they are clearly not identical. CCSM has stronger negative anomalies over the northern part of the domain than does PCM for pattern #1, while the opposite is true for pattern #5 – there is a larger positive anomaly for PCM over the northern part of the domain relative to CCSM. However, the general spatial patterns in terms of both anomalies (coloured contours) and average geopotential heights (black lines) are similar between the models.

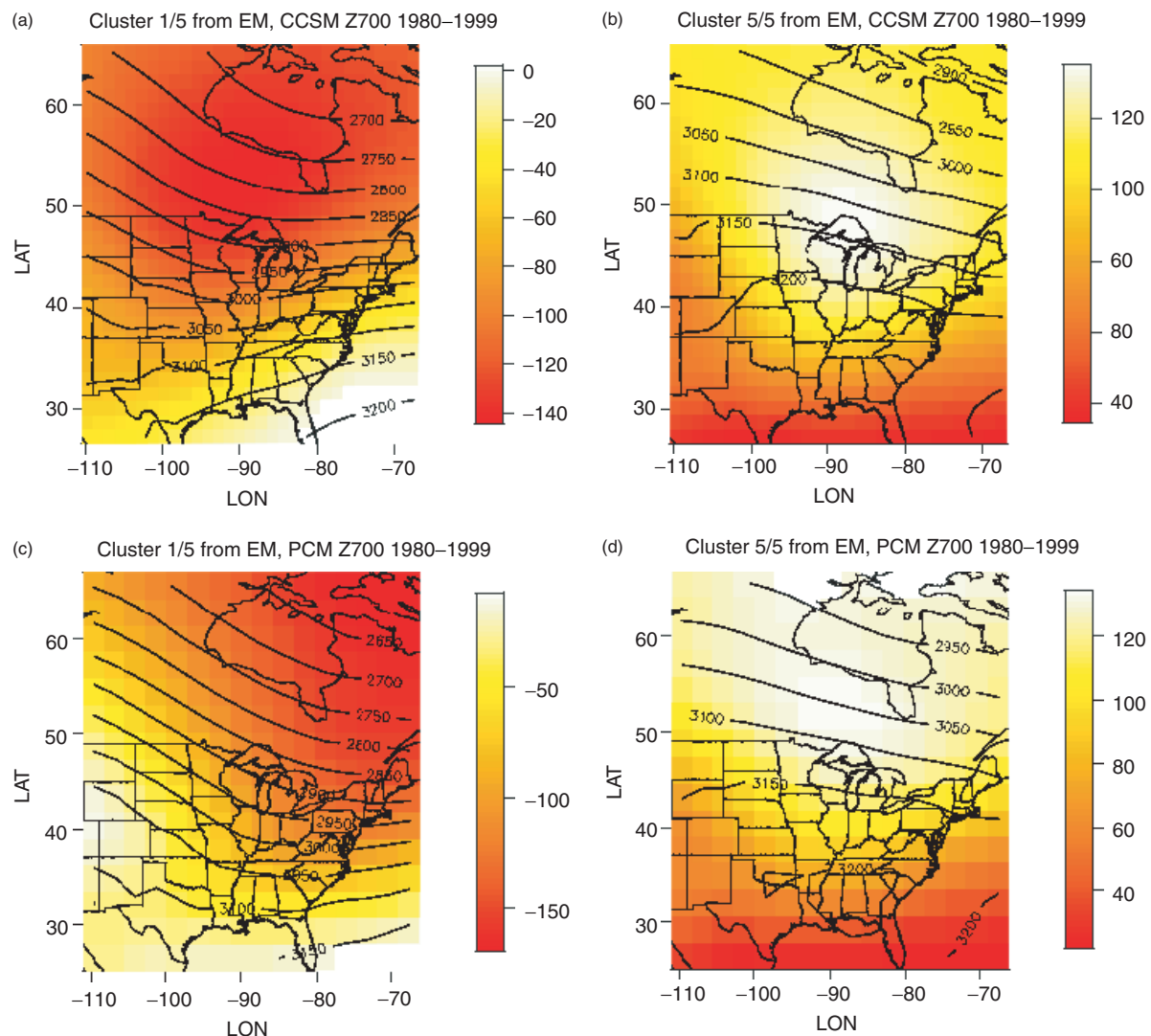


Figure 5. The first (mainly winter) and fifth (mainly summer) patterns resulting from EM-based clustering for CCSM Zg700 (a) and (b) and for PCM Zg700 (c) and (d).

A qualitative comparison between the AOGCM-based patterns and the reanalysis-based patterns shown in Figure 2 for NCEP and ERA-40 reveals some strong similarities. Overall, it is clear that the AOGCMs are successfully capturing the major characteristics of mean summer and winter circulation patterns over the eastern half of North America. However, there are also some clear differences. For pattern #1 (winter), PCM appears to do better than CCSM at capturing the transition from strongly negative to less negative anomalies that, in the reanalysis and PCM plots, runs in a northwest direction from the Midwest up to Alberta. However,

PCM overestimates the gradient in the southwest. For pattern #5 (summer), PCM again appears to be slightly more successful than CCSM at capturing the northward extension in the large positive anomaly that, in the reanalysis patterns, reaches upward into Hudson Bay.

AOGCM-based seasonal circulation frequencies

Further information is obtained from monthly histograms showing the AOGCM-simulated frequency of each pattern (Figure 6) that can then be compared to the same histograms generated from the reanalysis data (Figure 3 and 4). Comparing Figure 3 and 4 with

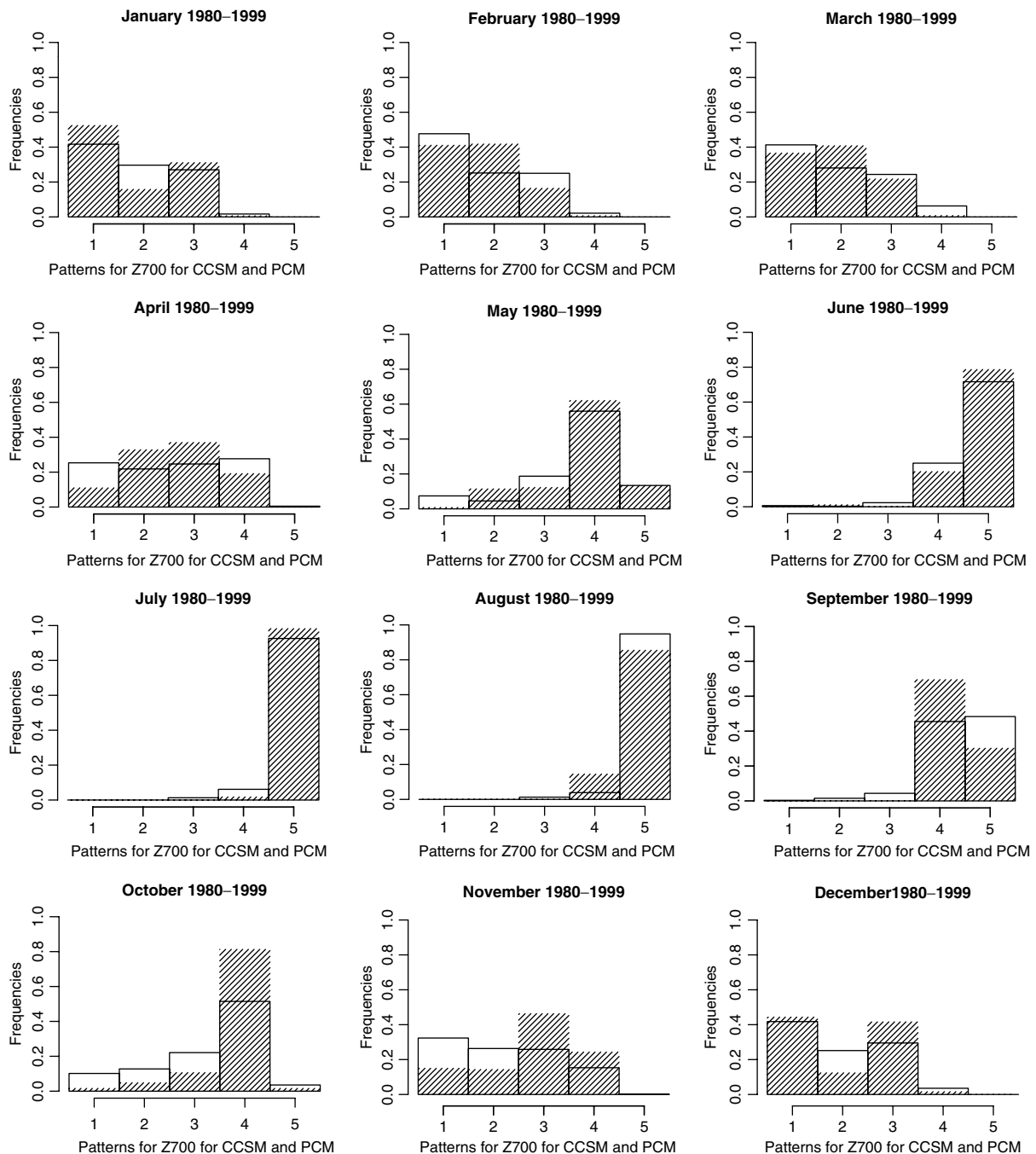


Figure 6. Monthly histograms for EM-generated patterns for the 5 CCSM Zg700 patterns (empty bars) and for the 5 PCM Zg700 patterns (shaded bars).

Figure 6, we first see that the AOGCMs clearly show the shift from the 'winter' pattern (#1) through the transition patterns (#2-4) to the summer pattern (#5), and back again in the autumn. So there is no question but that the AOGCMs are capable of successfully capturing the primary seasonal patterns and transitions over eastern North America. As with the patterns compared above, however, there are some significant differences between the individual models themselves and between the AOGCM-simulated frequencies and reanalysis-based ones.

Although CCSM and PCM are both NCAR-based models, they have different development histories and contain distinct representations of many key atmospheric and oceanic processes. In addition, the CCSM has higher resolution than the PCM. Comparing the model-simulated temporal distributions as displayed in histograms (Figure 6) with reanalysis (Figure 3 and 4), one of the first things we see is that the temporal frequency of modelled patterns derived from the newer and higher-resolution model (CCSM) are not necessarily closer to the distribution of the higher-resolution reanalysis data (ERA-40 1.125°). For level Zg500, where resolution was seen to have the strongest impact in the reanalysis data, the (1.4° resolution) CCSM histograms are closer to the NCEP (2.5°) and ERA-40 2.5° than to ERA-40 1.125°. This suggests that, at this level, model formulation and parameterization of dynamics plays a more important role in determining the degree to which seasonal circulation patterns are reproduced than the actual resolution of the model. This would not likely hold true for extremely coarse models. For PCM (2.8°), at level Zg500 we could have expected it to be closer to NCEP and ERA-40 2.5° than ERA-40 1.125°. However, we see that for Zg500, PCM does not have a clear 'closest' reanalysis partner although its patterns do resemble those from NCEP and ERA. For Zg700, the PCM Zg700 histograms are slightly closer to ERA40 2.5° for winter. It is more difficult to say something about proximity for the SLP histograms from CCSM or PCM, the reanalysis SLP histograms for NCEP being very close to the ERA40 SLP histograms for the EM method.

We next turn to the issue of seasonality, and the question of whether the AOGCMs are able to reproduce the shifts in patterns that occur by season that are apparent in the reanalysis-based patterns. The first thing we notice from Figure 6 (where patterns were identified using the EM algorithm) is the extension of the summer patterns and shifts in modelled seasonality relative to reanalysis. In general, the closer to the ground level (i.e. the higher the pressure level), the clearer the summer extension for CCSM and the more dominant the transition pattern #4 for PCM. The frequencies of the transition patterns are then reduced or shifted relative to reanalysis (Figure 3 and 4) as a consequence.

For Zg500, CCSM simulates summer conditions too early, reducing the frequencies of transitional patterns and anticipating summer characteristics by 1 month (see website Vrac and VanDorn, 2006). Indeed, the main

CCSM Zg500 summer pattern (#5) is already the most frequent pattern in June, which is clearly not true for the reanalysis data sets. In contrast, for PCM (the lower-resolution model), the appearance and disappearance of the main summer and winter patterns correspond well with the timing of the same patterns for the reanalysis data.

For Zg700, the CCSM summer pattern extends into both the shoulder seasons. Again, the main summer pattern (#5) is more frequent in September/October and May/June than it is for the reanalysis patterns. In other words, CCSM tends to simulate summer-like conditions at the Zg700 level before they really start and after they have ended, according to reanalysis. CCSM winter patterns are more consistent with reanalysis, exhibiting similar monthly frequencies. For PCM, we also have an early appearance of the main Zg700 summer pattern (#5) in May that is not as noticeable in the reanalysis data sets. This PCM summer pattern has the highest frequency in June – again indicating too early summer conditions – that is not true in reanalysis. In contrast to CCSM, which extends summer conditions into the fall, with the PCM we see early appearances of a transition pattern in August that is slightly too frequent as compared to reanalysis. At the opposite end of the annual cycle, PCM Zg700 winter conditions are very weak at the beginning of the cold season. The main PCM Zg700 winter pattern (#1) appears in October/November, i.e. at the right time but with a too low frequency relative to reanalysis.

For SLP, although the CCSM histograms for July/August are close to those for reanalysis, CCSM clearly induces modelled summer conditions too early before summer and too strong in early fall. Furthermore, the CCSM winter signal is very weak as the transitional patterns are all present in winter with about the same frequencies. That is not true in the reanalysis patterns, which clearly display a dominant winter pattern. For PCM, the frequencies of the main SLP summer pattern (#4) are not high enough for June and July – representing weak summer conditions for the first half of the summer – and too high (i.e. patterns too frequent) in the first half of fall in September and October – corresponding to very late summer conditions. This summer shift implies late fall conditions (strong in November) are still leading the histogram of December. In contrast, the reanalysis histograms already present strong winter conditions for this month. Although showing less variability than the reanalysis, PCM SLP histograms for January/February are clear enough. However, the winter conditions (and its main pattern) last too long, through March and April.

The primary conclusions we can draw from these results above are that: (1) the AOGCM models are capable of reproducing similar though not identical patterns to those produced by the reanalysis data, and (2) these patterns also show coherent seasonal signals for dominant winter and summer patterns, and transitional patterns evident in spring and fall. At the same time, however, the models have difficulties in identifying

the shifts between seasons, with summer and/or winter patterns being extended into the shoulder seasons and some transitional patterns remaining too strong into the primary seasons, a bias not seen in any of the reanalysis-based patterns regardless of the physical model used or the spatial resolution of the data.

Although sometimes shifted or reduced – with some degree related to the elevation – the most important transition characteristics between seasons do seem to be captured by the models in terms of atmospheric circulation patterns, with the general exception of a too-early commencement for the summer season, seen across most models and levels.

To verify that these results are not due to some instability of the model-based clustering method, analysis of additional CCSM and PCM simulations is already underway. However, a first sign of the stability of the method is given by looking at the patterns that we obtain from different sampling time periods.

Sensitivity of the patterns to the sampling period

A preliminary examination of the clusters and their seasonal frequencies resulting from four different 10-year data sets (1980–1989, 1985–1994, 1990–1999, and 1980–1998 for every second year) reveals two main conclusions. First, samples that takes years that are evenly distributed throughout the period exhibit similar patterns and monthly frequencies to the 20-year period, suggesting that 20 years of daily data provides a sufficiently large sample size that can be used to characterize climatological circulation patterns over eastern North America. Second, however, patterns generated from consecutive years earlier in the time period (e.g. 1980–1989) with those generated later (e.g. 1990–1999), although continuing to identify similar patterns, also show some consistent evolution in the seasonal patterns over this time. Specifically, fall patterns are extending into winter, spring patterns are beginning earlier in the year, and winter patterns are less distinct in the later time periods. These shifts can certainly be at least partially attributed to the smaller sample size. However, this shift is (1) only really evident for winter patterns and transition patterns close to winter, and (2) is present not only in ERA and NCEP, but also in PCM and CCSM, which have no reason to have the same ‘natural variability’ fluctuations as ERA and NCEP. This could suggest a common driver for the changes among all four of these models. Given that the CCSM and PCM total-forcing simulations are not being run with prescribed sea surface temperatures or any other type of observational constraint, the natural suspect for a common driver would be anthropogenic greenhouse gas emissions and their influence on climate. This suggests the possibility that the differences in the patterns that we observe between the 1980’s and the 1990’s may be due to climate change. This hypothesis will be investigated in future work that includes longer time periods.

Although this analysis was confined to the period 1980–1999, we hypothesize that some of this over-extension of the summer season has the potential to be

because of over-estimation of the effects of anthropogenic climate change on the circulation patterns, since one of the primary features of climate change over North America is to extend the summer season. This hypothesis will also be tested in future studies by comparing AOGCM-derived patterns from earlier in the century with reanalysis, as well as characterizing the shifts in the frequencies of these patterns that are expected over the coming century under a range of climate change scenarios.

CONCLUSIONS AND FURTHER WORK

In this study, we first examined the ability of clustering algorithms to identify the primary seasonal patterns in atmospheric circulation over North America, based on reanalysis data for the period 1980–1999. We then compared the resulting patterns and their monthly frequencies with those simulated by AOGCMs to assess the degree to which models are able to capture the seasonal characteristics seen in reanalysis. Two different clustering methods, EM and HAC, were used to define large-scale atmospheric patterns for reanalysis (NCEP, ERA40 1.125°, and ERA 40 2.5°), while the EM method only was used to identify the dominant atmospheric circulation patterns in AOGCM historical total-forcing simulations (PCM and CCSM). These two methods were applied separately to each of the four levels – Zg500, Zg700, Zg850, and SLP – to capture and compare the seasonality characteristics present in the five data sets, allowing us to draw a number of conclusions.

First, we saw that the clustering method used to determine the atmospheric structures influences the definition of these patterns. In general, the EM approach appears to be generally more reliable than the HAC approach. Specifically, more variability and stronger transition patterns are detected using the EM approach, reflecting the ability to more realistically simulate observed weather patterns even within seasons such as summer or winter that tend to be dominated by a single pattern. Furthermore, we determined that NCEP and ERA-40 do not contain the same exact seasonal features and that the differences in spatial resolution between the data sets do not entirely explain the discrepancies, particularly at lower levels.

Given that the EM method produced clusters that appear to be more stable and less sensitive to data issues such as resolution, we then applied it to clustering of AOGCM-based patterns. Although seasonal signals seem to be correctly simulated by CCSM and PCM according to both EM and HAC patterns (with some differences according to the level), we note some seasonal shifts or even extensions of the primary seasons are clearly visible (mainly for summer patterns) and are altitude-dependent. This suggests that the AOGCMs, while clearly capable of reproducing the main seasonal circulation patterns, also demonstrate significant biases in the timing and the mean shape of the patterns. These patterns appear to be

a function of the physics of the model that determines dynamics at various levels (with higher levels being closer to reanalysis than the lower levels, as would be expected since the flow at those levels is more zonal and surface-level pressure is affected by more regional-scale processes that the models may have difficulty in simulating) more than of the resolution, since higher-resolution CCSM did not perform noticeably better than PCM on a consistent basis.

Despite the demonstrated dependence of circulation pattern characteristics on the clustering method used, we believe this comparison of AOGCM-based patterns with reanalysis to be robust, for two reasons. First, additional analysis of AOGCM simulations using the HAC approach (not discussed here, see website Vrac and VanDorn, 2006, for additional figures) produced similar shifts in summer patterns. Perhaps more importantly, however, these conclusions are consistent with those of independent studies that do not involve clustering. For example, Hayhoe *et al.* (2006) find that present-day PCM and CCSM simulations have a summer northward shift in the jet stream over eastern North America that occurs approximately 1–2 months too early in the year, which would also be expected to result in premature simulation of summer conditions over the region.

It is clear that seasonal shifts present in AOGCM simulations and captured in large-scale patterns can have important repercussions if the patterns are used to project future climate change. However, before applying this method to assess the implication of future shifts in seasonal circulation patterns under a climate change scenario, we instead propose a future intermediary step to examine whether differences have occurred between a past historical and the present-day time period, not only within reanalysis data but also within the AOGCM simulations. This follow-on study of past and present seasonal evolution will be an important step to further evaluate the capability of AOGCMs to reproduce the types of shifts in circulation patterns that are likely to be occurring due to anthropogenic change, and which determine the net effect of climate change on surface conditions at the regional to local scales.

ACKNOWLEDGEMENTS

The authors thank Julie Arblaster (NCAR/ABM) for her insightful comments and suggestions.

Although the research described in this article has been funded wholly or in part by the United States Environmental Protection Agency through STAR Cooperative Agreement # R-82940201 to the University of Chicago, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

We acknowledge the PCM and CCSM3 modeling groups for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model output,

the JSC/CLIVAR Working Group on Coupled Modeling (WGCM) and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model output analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, US Department of Energy.

APPENDIX A – DESCRIPTION OF EM METHOD

To define the atmospheric patterns, we employ a mixture of statistical distributions (Pearson, 1894). In this approach, we estimate g , the pdf of one of the variables (e.g. Zg850), as a weighted sum or *mixture* of K parametric pdfs g_i ($i = 1, \dots, K$), with parameters α_i :

$$g(\mathbf{x}) = \sum_{i=1}^K \pi_i g_i(\mathbf{x}; \alpha_i). \quad (1)$$

Here π_i is called the ‘mixture ratio’ and corresponds to the posterior probability of belonging to component i , or the i^{th} pattern. In this formulation, the i^{th} pattern, say P_i , is associated with and is actually defined by the i^{th} pdf g_i . In this work, we consider that the g_i pdfs are Gaussian pdfs. Hence, we deal with a mixture of Gaussians where $\alpha_i = (\mu_i, \Sigma_i)$ with μ_i a vector of means and Σ_i the variance-covariance matrix of g_i .

To estimate the parameters $\alpha_1, \dots, \alpha_K$ and π_1, \dots, π_K of such a mixture, we use an EM type algorithm (Dempster *et al.*, 1997; for review, see McLachlan and Peel, 2000) consisting of two successive and iterative steps of expectation and maximization of the so-called complete log-likelihood. The estimation process is performed through the R function *Mclust* developed by Fraley and Raftery (2006) in the *mclust* R package. A constraint is imposed on the variances in the maximization process (M step) to avoid singularities of the likelihood. These singularities occur when a vector of the means of one component, say μ_i , is set equal to any observed multidimensional data and the variances tend to zero.

The number of components K is generally either given *a priori* according to some prior knowledge, or adjusted by an expert after comparisons of results for several values of K . In this work, several K are tested from $K = 1$ to $K = 10$ and the one minimizing the Bayesian Information Criterion (BIC, Raftery, 1986a, 1986b) is kept. More precisely, each EM run tries to fit the model given in Equation (1) with the data for a given value of K but also for several given structures of the variance matrix of the K components. Here, the variance matrix can be spherical, diagonal, or ellipsoidal and with equal or varying volumes. The BIC is then calculated for each variance structure and each K and the couple (K , variance model) minimizing the BIC is kept.

For the chosen variance model and the chosen K , patterns are obtained by applying the principle of posterior maximum. For pattern P_i ,

$$P_i = \{\mathbf{x}; \pi_i g_i(\mathbf{x}, \alpha_i) \geq \pi_j g_j(\mathbf{x}, \alpha_j), \forall j = 1, \dots, K\}. \quad (2)$$

In other words, each x is allocated to the cluster for which the associated model maximizes the posterior probability that x belongs to this cluster.

APPENDIX B – SENSITIVITY OF CIRCULATION PATTERNS TO SPATIAL RESOLUTION

In addition to differences in model formulation, we hypothesize that part of the discrepancy between the ERA-40 and NCEP patterns (for a given set of patterns based on either EM or HAC) may arise from their different spatial resolution, which is 1.125° for ERA40 and 2.5° for NCEP. For that reason, we derived new patterns based on ERA-40 data at 2.5° spatial resolution, identical to that of NCEP.

For the EM-based patterns (see website Vrac and Van-Dorn, 2006), the influence of spatial resolution appears to grow with elevation – i.e. the 2.5° NCEP and ERA-40 patterns become more and more similar to each other higher level. In fact, for Zg500, the histograms of the five patterns from NCEP are almost identical to those from the ERA-40 2.5° data. In contrast, for Zg850, the lowest level at which ERA-40 2.5° was available, the histograms of the five patterns from ERA-40 1.125° are equivalent to those from ERA-40 2.5° and distinctly different from NCEP. This level shows no influence of the spatial resolution of the data set. For Zg850 with the EM method, therefore, it seems that the differences between NCEP and ERA-40 arise from the models themselves, whereas for the Zg500 level, the primary differences between reanalysis-derived patterns is due to the resolution of

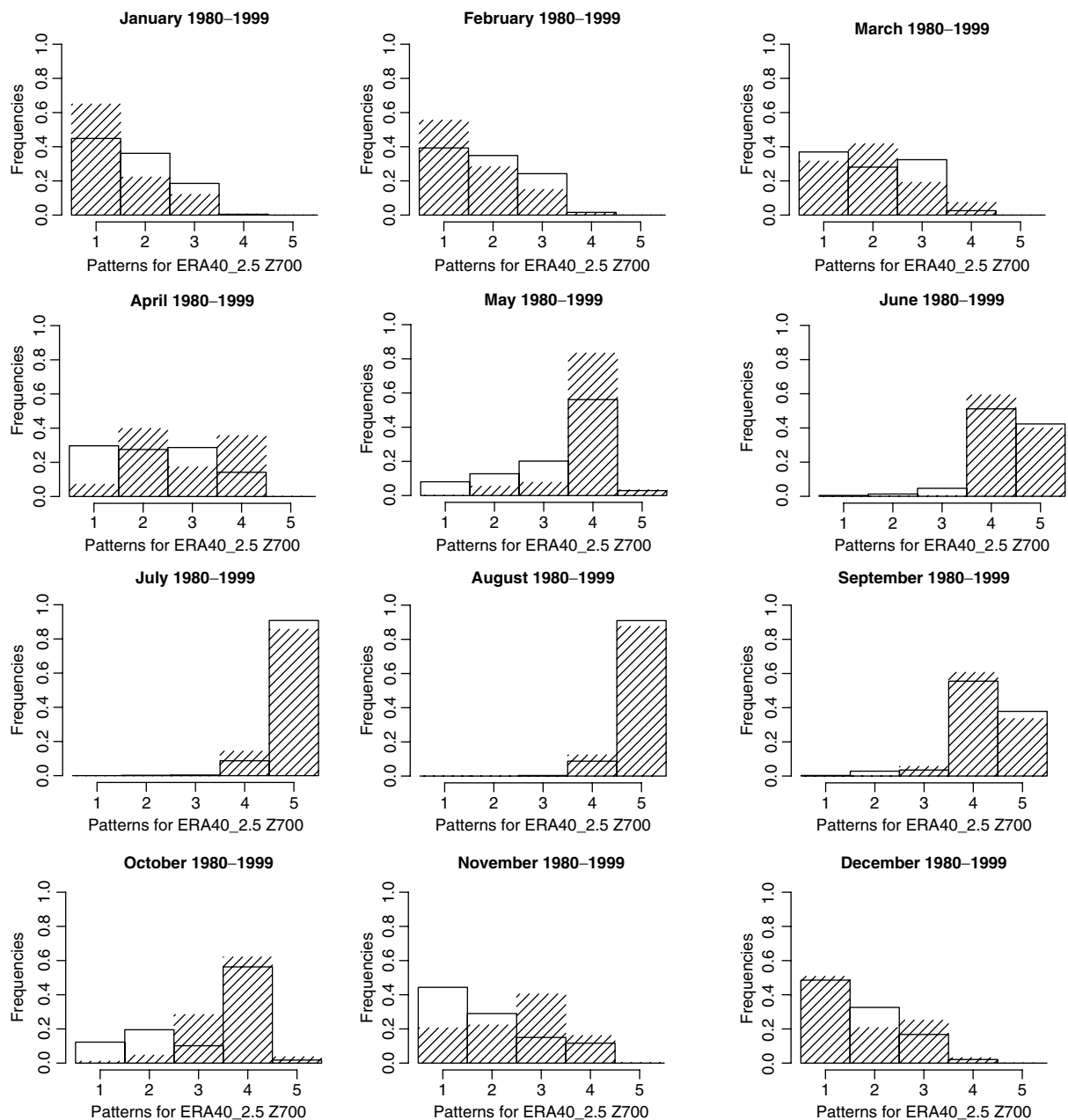


Figure 7. The same as Figure 3, but for ERA40 2.5° dataset.

the original dataset used to identify the clusters, not the model used to generate the geopotential fields.

Results for Zg700, an intermediate level, confirm this hypothesis. In Figure 7, we see influences from both factors - dataset resolution and the model used - according to season. To begin with, the histograms for the summer months (JJAS) are equivalent among all three reanalysis data sets (NCEP, ERA-40 1.125°, and 2.5° - see Figures 3 and 4), so no model or resolution differences are seen there. In the winter, NCEP patterns seem to be clearer than ERA-40 patterns (for either 1.125° or 2.5°). In addition, the ERA40 1.125° histograms for the winter months (DJF) are very similar to ERA40 2.5° winter, but not to NCEP. So, for winter, the differences seem to be model-based and the resolution does not appear to affect either the definition or the appearance of the seasonal patterns. For the transition seasons, however, April/May and October/November patterns from ERA-40 1.125° are surprisingly closer to NCEP than to ERA-40 2.5°.

For clusters determined using the HAC algorithm, there is no consistent elevation gradient. In general, relative to EM, HAC-based clusters produce more differences arising from the models themselves (NCEP, ERA-40) than from the resolution, but at the same seasons for each level. For Zg500, the histograms of summer months (JJA) are the same for ERA40 2.5° and 1.125°, with one pattern taking 100% of July and August. For the other months, the frequencies of the ERA40 2.5° patterns are clearly between the ERA-40 1.125° and NCEP frequencies, reflecting the influence of both the resolution and the actual model. For Zg700, summer month (JAS) frequencies for ERA40 2.5° are close to those for NCEP, but winter structures (1 and 2 in Figure 3) are about the same for ERA40 1.125° and 2.5°. We can see that transition seasons are different for all three of the different reanalysis data sets, suggesting that even if the resolution influences the definition and the temporal frequencies of the patterns, the discrepancy between the data sets themselves is at least as important as the resolution. This last remark is true for all four seasons for Zg850, where we

clearly see some features being different due to the models and others due to the resolution.

APPENDIX C – SENSITIVITY OF CIRCULATION PATTERNS TO MODEL PARAMETERIZATION

Since the circulation patterns identified from NCEP and ERA-40 fields at the same resolution and using the same clustering method are still not identical (see Appendix B, above), we next assess the degree to which model formulation is capable of introducing systematic errors into the output through evaluating the characteristics of the first PCs of the geopotential height fields to which our clustering methods were applied. To make our comparisons between NCEP and ERA-40 2.5° in the same factor space (i.e. space of the PC), NCEP reanalyses are projected onto the ERA-40 2.5° factor space. In other words, the rotation matrix applied to ERA-40 to define the ERA-40 PCs has, here, also been applied to the NCEP data. In this section, we only present results for Zg700, but similar results can be obtained for the other levels examined previously. Let us define $PC_i(\text{ERA-40})$ to be the linear combination given by the i^{th} PC for ERA-40 2.5°. In Figure 8, we represent ERA-40 2.5° PC i vs NCEP values of $PC_i(\text{ERA-40})$ (i.e. NCEP data projected onto the i^{th} factor axis), for $i = 1, 2$.

As expected, the comparison of the ERA-40 2.5° PCs with the NCEP values of $PC_i(\text{ERA-40})$ reveal some significant discrepancies between the PCs simulated by each model. If NCEP and ERA-40 provided the same exact data, Figure 8 would show the $y = x$ line. Here, although being close to a straight line, these graphics do not represent $y = x$ but rather a shift in this line. This shift is also evident when looking at the statistical distributions. In Figure 9, we see the densities of the PCs for ERA-40 2.5° Zg700 (red lines) as compared with the densities of the NCEP values for $PC_j(\text{ERA-40})$ (black lines), for the first two PCs of the variance of each field. Whereas the variances and correlations of the PCs are the same for NCEP and ERA-40 2.5° in the ERA-40 2.5°

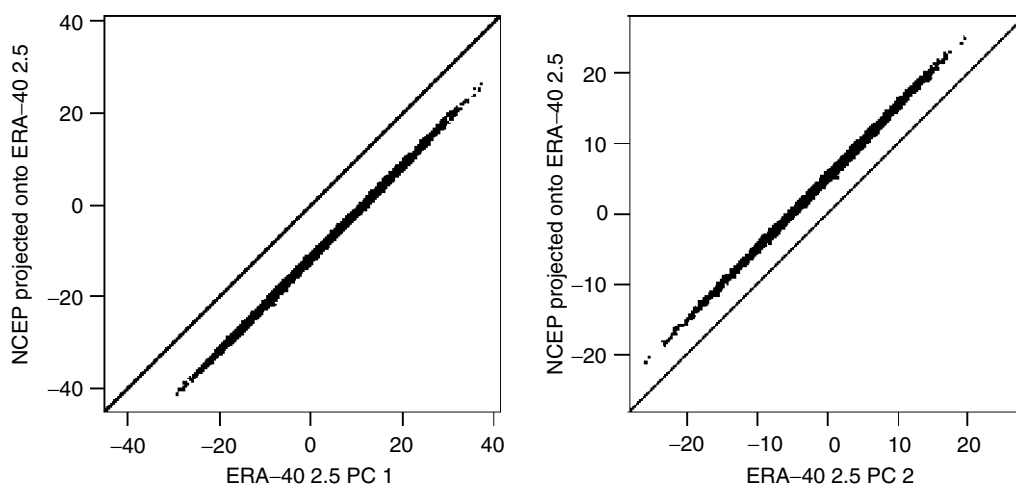


Figure 8. ERA-40 2.5° Zg700 PC i versus NCEP values of $PC_i(\text{ERA-40})$, for the first two principal components. The black line represents the $y = x$ line.

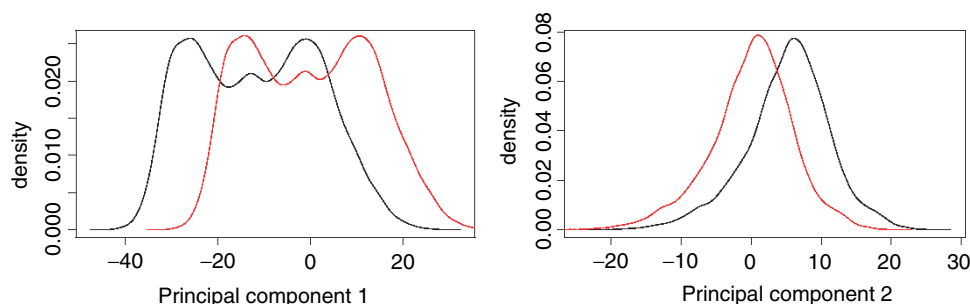


Figure 9. Densities of ERA-40 2.5° Zg700 PC i (red lines) and densities of the NCEP values of $PC_i(\text{ERA-40})$ (black lines), for the first two principal components.

factor space – as is generally not the case for raw data, i.e. not transformed via PCA (see for instance Gleisner *et al.*, 2005) – the means (and so the densities) are shifted for most of the NCEP PCs relative to ERA-40. Moreover, for PC 1, which explains the largest part of the variance, the densities are bimodal and we can remark that the modes are shifted for NCEP relative to ERA-40, leading to opposing phases (i.e. local maxima of the densities). That means that, for PC 1, an NCEP component with high density corresponds to a value of an ERA-40 component with lower density and conversely. These clear shifts of the densities of PCs seen in Figure 9 illustrate that the structures of NCEP and ERA-40 2.5° reanalyses are not equivalent and so provide an additional reason to explain why the clustering algorithms, even when constrained by the number of patterns, continue to capture different structures and thereby induce different patterns for NCEP relative to ERA-40 2.5°.

As we saw previously for the patterns themselves, it appears that the differences between the NCEP PCs and the ERA-40 2.5° PCs depend on the elevation. The higher the level of the atmosphere, the more influential the spatial resolution and so the closer NCEP PCs to the ERA-40 2.5° PCs. Moreover, the shifts between ERA-40 PCs and the NCEP values of $PC_i(\text{ERA-40})$, seen in Figures 9 and 8, are visible when plotting on the same graph (not shown here but available on website Vrac and VanDorn, 2006) ERA-40 PC i vs ERA-40 PC $i + 1$ and the NCEP values of $PC_i(\text{ERA-40})$ vs the NCEP values of $PC_{(i+1)}(\text{ERA-40})$. As expected, the scattering is similar, but when NCEP reanalyses are classified in ERA-40 2.5° clusters, the gap between NCEP and ERA-40 makes the distinction between the classes maladaptive to NCEP structure.

REFERENCES

- Aizen E, Aizen V, Melack J, Nakamura T, Ohta T. 2001. Precipitation and atmospheric circulation patterns at mid-latitudes of asia. *International Journal of Climatology* **21**: 535–556.
- Bunkers MJ, Miller JR, DeGaetano AT. 1996. Definition of climate regions in the northern plains using an objective cluster modification technique. *Journal of Climate* **9**: 130–146.
- Burnett A, Mullins H, Patterson W. 2004. Relationship between atmospheric circulation and winter precipitation $\delta^{18}\text{O}$ in central New York state. *Geophysical Research Letters* **31** DOI: 10.1029/2004GL021089.
- Collins W, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD. 2006. The community climate system model: CCSM3. *Journal of Climate* **19**(11): 2122–2143.
- Covey C, AchutaRao K, Cusbasch U, Jones P, Lambert S, Mann M, Phillips T, Taylor K. 2003. An overview of results from the coupled model intercomparison project (cmip). *Global and Planetary Change* **37**: 103–133.
- Covey C, Abe-Ouchi A, Boer G, Boville B, Cubasch U, Fairhead L, Flato G, Gordon H, Guilyardi E, Jiang X, Johns TC, LeTreut H, Madec G, Meehl G, Miller R, Noda A, Power S, Roeckner E, Russell G, Schneider E, Stouffer R, Terray L, von Storch J-S. 2000. The seasonal cycle in coupled ocean-atmosphere general circulation models. *Climate Dynamics* **16**: 775–787.
- Davis R, Dolan R, Demme G. 1993. Synoptic climatology of atlantic coast northeasters. *International Journal of Climatology* **13**: 171–189.
- Dempster A, Laird N, Rubin D. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B* **39**: 1–38.
- Fovell R. 1997. Consensus clustering of U.S. temperature and precipitation data. *Journal of Climate* **10**: 1405–1427.
- Fraley C, Raftery A. 2006. The mclust package. R package manual, The Comprehensive R Archive Network.
- Gleisner H, Thejll P, Stendel M, Kaas E, Machenhauer B. 2005. Solar signals in tropospheric re-analysis data: Comparing NCEP/NCAR and ERA40. *Journal of Atmospheric and Solar-Terrestrial Physics* **67**: 785–791.
- Hayhoe K, Wake C, Anderson B, Bradbury J, DeGaetano A, Liang X, Zhu J, Maurer E, Wuebbles D. 2006. Translating global change into regional trends: Climate drivers of past and future trends in the U.S. northeast. *Bulletin of the American Meteorological Society* (submitted).
- Hurrell J, Hack J, Boville B, Williamson D, Kiehl J. 1998. The dynamical simulation of the NCAR community climate model version 3 (CCSM3). *Journal of Climate* **11**: 1207–1236.
- Hurrell J, Hack J, Phillips A, Caron J, Yin J. 2006. The dynamical simulation of the community atmospheric model version 3 (CAM3). *Journal of Climate* **19**(11): 2162–2183.
- Huth R. 2001. Disaggregating climatic trends by classification of circulation patterns. *International Journal of Climatology* **21**: 135–153.
- Huth R. 2005. Downscaling of humidity variables: a search for suitable predictors and predictands. *International Journal of Climatology* **25**: 243–250.
- Jung T. 2005. Systematic errors of the atmospheric circulation in the ecmwf forecasting system. *Quarterly Journal of the Royal Meteorological Society* **131**: 1045–1073.
- Källberg P, Simmons A, Uppala S, Fuentes M. 2004. The ERA-40 archive. ERA-40 Project Report Series No. 17. ECMWF, Reading: UK.
- 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, Mho K, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.
- Kunkel K, Liang X. 2005. GCM simulations of the climate in the central united states. *Journal of Climate* **18**: 1016–1031.
- Lamb H. 1972. *British Isles Weather Types and a Register of Daily Sequence of Circulation Patterns*, Geophysical Memoir 116. HMSO: London.

- McLachlan G, Peel D. 2000. *Finite Mixture Model*. Wiley: New York.
- Pearson K. 1894. Contributions to the theory of mathematical evolution. *Philosophical Transactions of the Royal Society of London A* **185**: 71–110.
- Pongracz R, Bartholy J, Bogardi I. 2001. Fuzzy rule-based prediction of monthly precipitation. *Physics and Chemistry of the Earth* **9**: 663–667.
- Raftery A. 1986a. Choosing models for cross-classification. *American Sociological Review* **51**: 145–146.
- Raftery A. 1986b. A note on bayes factors for log-linear contingency table models with vague prior information. *Journal of the Royal Statistical Society Series B* **48**: 249–250.
- Salinger M, Mullan A. 1999. New Zealand climate: temperature and precipitation variations and their links with atmospheric circulation 1930–1994. *International Journal of Climatology* **19**: 1049–1071.
- Slonosky V, Jones P, Davies T. 2001. Atmospheric circulation and surface temperature in europe from the 18th century to 1995. *International Journal of Climatology* **21**: 63–75.
- Unal Y, Kindap T, Karaca M. 2003. Redefining the climate zones of Turkey using cluster analysis. *International Journal of Climatology* **23**: 1045–1055.
- Vrac M, VanDorn J. 2006. <http://temagami.tosm.ttu.edu/~data/patterns/>. Website for additional figures.
- Vrac M, Chédin A, Diday E. 2005. Clustering a global field of atmospheric profiles by mixture decomposition of copulas. *Journal of Atmospheric and Oceanic Technology* **22**: 1445–1459.
- Ward J. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**: 236–244.
- Washington W, Weatherly JW, Meehl GA, Semtner AJ Jr, Bettge TW, Craig AP, Strand WG Jr, Arblaster JM, Wayland VB, James R, Zhang Y. 2000. Parallel climate model (PCM) control and transient simulations. *Climate Dynamics* **16**: 755–774.
- Xoplaki E, Gonzalez-Rouco J, Gyalistras D, Luterbacher J, Rickli R, Wanner H. 2003. Interannual summer air temperature variability over Greece and its connection to the large-scale atmospheric circulation and mediterranean SSTs 1950–1999. *Climate Dynamics* **20**: 537–554.
- Zorita E, Hughes J, Lettenmaier D, von Storch H. 1993. Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. Report 109. Max-Planck-Institut für Meteorologie: Hamburg.