2003

# Evaluation of the NCEP/NCAR Reanalysis in Terms of Synoptic Scale Phenomea: A Case Study from the Midwestern USA

Justin T. Schoof
*Southern Illinois University Carbondale*, jschoof@siu.edu

S C. Pryor
*Indiana University Bloomington*



Copyright 2003 Royal Meteorological Society. International Journal of Climatology. 23: 1725–1741 (2003). Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.969

INTERNATIONAL JOURNAL OF CLIMATOLOGY
*Int. J. Climatol.* **23**: 1725–1741 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.969

# EVALUATION OF THE NCEP–NCAR REANALYSIS IN TERMS OF SYNOPTIC-SCALE PHENOMENA: A CASE STUDY FROM THE MIDWESTERN USA

J. T. SCHOOF* and S. C. PRYOR
*Atmospheric Science Program, Department of Geography, Indiana University, Student Building 120, 701 East Kirkwood Ave., Bloomington, IN 47405-7100, USA*

*Received 26 November 2002*
*Revised 26 August 2003*
*Accepted 11 September 2003*

## ABSTRACT

We evaluate the ability of the National Centers for Environmental Prediction (NCEP)–National Center for Atmosphere Research (NCAR) reanalysis to represent the synoptic-scale climate of the Midwestern USA relative to radiosonde data. Independent, automated synoptic classifications, based on rotated principal component analysis (PCA) of 500 hPa geopotential heights, 850 hPa air temperatures, and 200 hPa wind speeds and a two-step clustering algorithm, result in a 15-type NCEP–NCAR synoptic classification and a 14-type radiosonde classification. The classifications are examined in terms of similarities and differences in the modes of variance manifest in the PCA solutions, the spatial patterns and variability of input variables within each weather type, and the temporal variability of the occurrence of each weather type. The classifications are then compared in terms of these characteristics and the degree of mutual class occupancy. Although the classifications identify a number of the same weather types (in terms of the input data, PCA solution, and mutual occupancy), the correspondence is imperfect. To assess whether the differences in the classifications are due to errant assignment of data to clusters or to differences in the fundamental modes present in the data sets as represented by the PC loadings and scores, a third targeted classification is undertaken that categorizes the NCEP–NCAR reanalysis data according to the radiosonde PCA solution. This classification exhibits a higher degree of similarity to that derived using the radiosonde data (in terms of both interpretability and mutual class occupancy), but the solutions still exhibit considerable differences. It is probable that the discrepancies are partly a function of the differing data structures and densities, but they may also reflect differences in the intensity of synoptic-scale phenomena as manifest in the data sets.


KEY WORDS: synoptic climatology; eigenvector-based synoptic classification; weather-type analysis; NCEP–NCAR reanalysis; radiosonde

## 1. INTRODUCTION

General circulation models (GCMs) represent a critical tool for quantifying likely climatic consequences of modified atmospheric composition. A key aspect of evaluation of GCMs is an assessment of their ability to represent the current climate and features of the atmospheric flow via comparison with ambient data or representations of fundamental phenomena (e.g. McKendry *et al*., 1995; Boyle, 1998; D'Andrea *et al*., 1998; Stratton, 1999). These evaluations focus not only on individual variables, but also increasingly on the ability of the models to represent synoptic-scale features that dominate the midlatitudes in what D'Andrea *et al*. (1998) refer to as 'phenomenon diagnostics'. In principle, these analyses can evaluate dynamic and thermodynamic characteristics and the sub-grid-scale features of the atmosphere responsible for a large majority of regional impacts of climate change (e.g. precipitation associated with frontal passages (Katzfey and Ryan, 2000)), and they will become increasingly frequent as GCM resolution increases. These analyses, in turn, are critically

* Correspondence to: J. T. Schoof, Atmospheric Science Program, Department of Geography, Indiana University, Student Building 120, 701 East Kirkwood Avenue, Bloomington, IN 47405-7100, USA; e-mail: jschoof@indiana.edu

dependent on the robustness of observationally based synoptic classifications of current conditions to be used in the evaluation.

To classify atmospheric conditions properly at the synoptic-scale, long-term data records are needed from stations, or grid points, throughout the region of interest. The National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis (NCEPR) project (Kalnay *et al*., 1996; Kistler *et al*., 2001) was designed to provide homogenized (gridded) records of atmospheric fields to support climate research by assimilating data from multiple sources with model short-range forecasts. The coherence, accessibility and completeness of the NCEPR data set make it attractive for climate studies on topics ranging from climate variability and synoptic climatological analyses to comparative analyses of GCM performance. Here, we evaluate the ability of the NCEP–NCAR data set to represent the synoptic climate of the Midwestern USA relative to the 'raw' radiosonde data series, which form part of the data assimilated by the NCEPR. This research is warranted because although the NCEP–NCAR data set has been evaluated extensively relative to independent data sets (e.g. Kalnay *et al*., 1996; Renfrew *et al*., 2002), several shortcomings of the data set have been documented (e.g. Hines *et al*., 2000; Marshall, 2002; Swail and Cox, 2000; Trenberth and Stepanik, 2002; Renfrew *et al*., 2002). Also, previous analysis of the data set has typically focused on individual and collocated variables rather than ensembles representing fundamental features/scales and interrelationships of the data-set variables.

## 2. DATA/STUDY AREA

The primary objective of this study is to evaluate the ability of the NCEPR to reproduce synoptic-scale climate and variability in the Midwestern USA. The study area selected for this analysis extends from 35° to 45 °N latitude and from 80° to 95 °W longitude and includes Indiana, Illinois, Ohio, Kentucky, Tennessee, Michigan, Missouri, West Virginia, Iowa, and parts of several others states (Figure 1). This study area was chosen because the relatively flat study region reduces the effects of terrain on circulation, and also due to the availability of data from a relatively large number of radiosonde stations. Within the study area, nine radiosonde stations (Figure 1) and 35 grid points (2.5° × 2.5° spacing) are available from the following two databases, respectively:

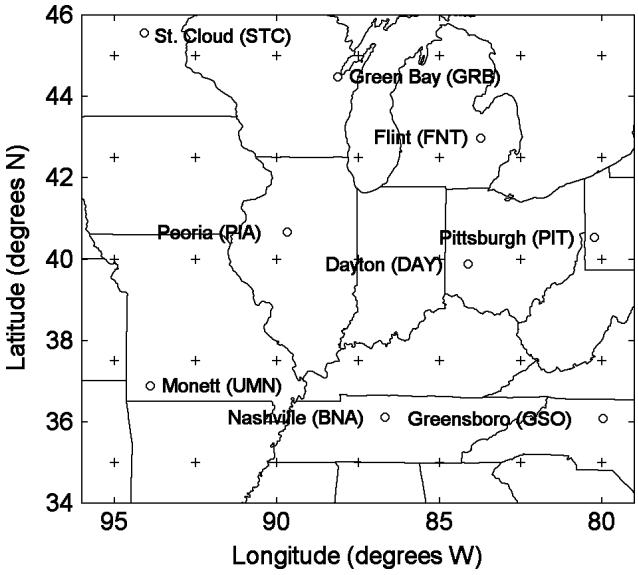


Figure 1. Map of study area showing the 35 NCEPR grid points (+) and the nine-station rawinsonde network (○, station name included in figure)

 

1. The Radiosonde Data of North America (1946–96) database issued by the Forecast Systems Laboratory and National Climatic Data Center (FSL–NCDC, 1997). This dataset is referred to as RSONDE hereafter.
2. The NCEP–NCAR 40-year reanalysis data series (Kalnay *et al*., 1996).

The former is used to derive a synoptic classification against which the latter is compared.

Note that though the radiosonde stations chosen for this study represent a broad cross-section of climates within the study region, a few areas are poorly resolved (Iowa and northern Missouri, southwest Ontario). Hence, contoured results for these areas should be treated with caution.

For this comparative study, it is necessary that the variables of interest are available from both data sets. The variables chosen for inclusion in the analysis are twice-daily (0 GMT, 12 GMT) observations of 500 hPa heights, 850 hPa air temperatures, and 200 hPa wind speeds. The variables are chosen broadly to represent circulation patterns, thermal advection, and jet stream strength respectively. The classification presented here is for the summers of 1971–93, where summer is defined following Schoof and Pryor (2001) as year-day 100 (10 April) to year-day 300 (27 October). The summer period is chosen due to the importance of agriculture within the study area. To ensure that any differences in the classifications are due to differences in the data sets, observations are only included in the analysis when coincident data are available from both data sets. In addition, time periods with large numbers of radiosonde geopotential height and temperature errors (as determined by NCEP–NCAR quality control; Collins, 2001) are removed from the analysis to minimize differences caused by radiosonde errors that are flagged by NCEP–NCAR complex quality control before inclusion in the reanalysis. Information regarding the rejection of upper-level winds by the NCEPR was not available. This analysis assumes that the missing data are randomly distributed, that the 23 year period is sufficient to represent the climate of this region and that the climate record for this period exhibits stationarity. Examination of the input data from each of the data sets broadly supports the last assumption. Although the input variables exhibit small trends (calculated from percentiles to avoid missing data issues), these trends are not generally statistically significant and trends between the RSONDE and NCEPR data sets are in agreement. The mean input fields show excellent agreement, though there are small differences in the standard deviations of the input variables (Figure 2).

## 3. METHODS

There are several methods available for the comparison of weather-type classifications (Huth, 2000). First, classifications can be applied to two data sets independently, allowing the dominant weather types from both data sets to be identified. The second method of comparing classifications requires that the classes are predetermined, which is often difficult in climatological studies. Although this method eases comparison between classifications, it does not require that either classification characterize the underlying data structure. The third method of comparison requires classification of the first data set, followed by projection of the results onto the second data set. This method classifies the second data set according to the structure found in the first, providing further insight into the physical basis of both classifications. Because this study is primarily concerned with the reproduction of synoptic-scale climate in the NCEPR, a multi-method approach is adopted. This approach allows examination of independent classifications of the input fields and an examination of the improvement achieved by projecting the NCEPR data onto the clusters defined using the observations (RSONDE).

### *3.1. Development of the synoptic classifications*

To meet the stated research objective, automated (objective) synoptic classifications (Yarnal, 1993) are performed using rotated principal components analysis (PCA) (Richman, 1986) and a two-step clustering algorithm (Davis and Kalkstein, 1990). In this analysis, modified S-mode PCA (Richman, 1986) is applied primarily as a data reduction technique (nine stations × two observations per day × three variables = 54 input variables for RSONDE; 35 grid points × two observations per day × three variables = 210 input variables

 

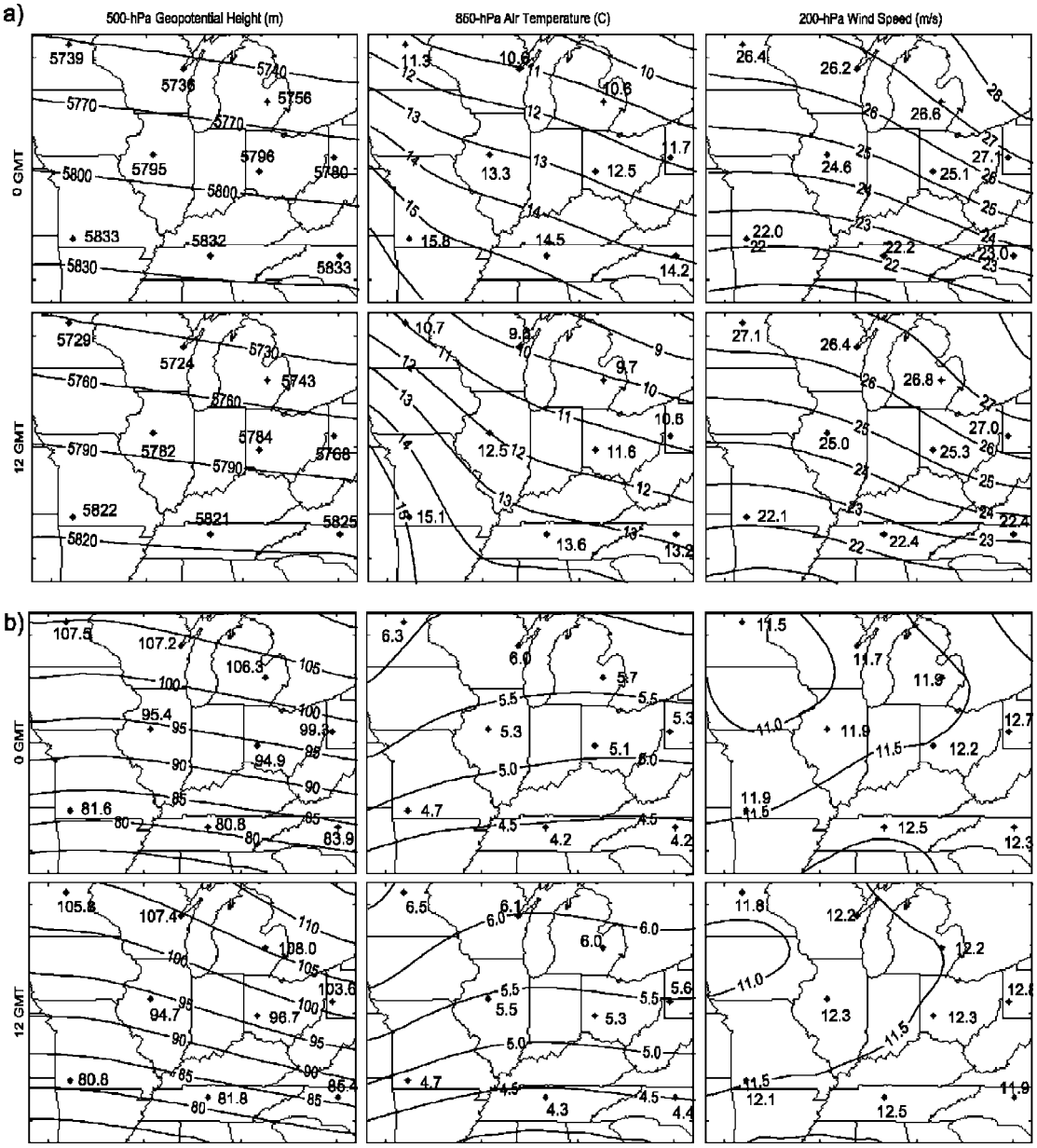


Figure 2. (a) Means and (b) standard deviations of RSONDE and NCEPR input data at 0 GMT and 12 GMT. In each plot, the contours represent NCEPR values. RSONDE values are shown in text adjacent to each station

for NCEPR) and to eliminate correlation between input variables for subsequent cluster analysis. However, the PC loadings matrices (representing the relationships between each variable and each PC) and PC scores matrices (representing the relationships between each observation and each PC) are also valuable tools in the interpretation of the classification results (see Section 4.1).

Correlation matrices are used as the dispersion matrices for the PCA to weight the input variables evenly, and the resulting PCs are rotated using (orthogonal) Varimax rotation (Richman, 1986; Bonell and Sumner, 1992; Brinkmann, 1999a). The number of PCs to retain for rotation and interpretation is evaluated using a number of methods, including scree plots (Catell, 1966), logarithmic eigenvalue plots (Craddock and Flood,

 

1969) and statistical tests, such as the *N* rule (Overland and Preisendorfer, 1982). The resulting PC scores are then clustered to produce synoptic weather types using a hierarchical, agglomerative algorithm followed by a non-hierarchical 'reassigning' algorithm (Davis and Kalkstein, 1990).

Based on previous success in climatological analyses (Kalkstein *et al*., 1987; Schoof and Pryor, 2001), the hierarchical average linkage clustering method (Sokal and Michener, 1958) is applied to the RSONDE and NCEPR PCA solutions. This method is based on the comparison of squared Euclidean distances between individual observations in two clusters. The similarity is taken as the average of these squared distances over all observations within the two clusters. One of the most difficult problems in the application of cluster analysis is the determination of the proper number of clusters (Kalkstein *et al*., 1987; Fovell and Fovell, 1993). In this analysis, several clustering 'cut-off rules' including the pseudo-*F* statistic (Calinski and Harabasz, 1974) and simple plots of cluster merging level versus number of clusters are applied. A graphical method, based on the increase in within-cluster standard deviation at each clustering stage is also employed. Because hierarchical clustering methods do not allow reclassification of an observation, a second, non-hierarchical clustering (convergent *k*-means method) is then applied, using the means of the PC scores associated with each of the hierarchically derived clusters as seeds for the new clusters. This step allows observations that were clustered early to be reassigned if they are closer to the centroid of a different cluster. Each of these resulting clusters represents a commonly occurring weather pattern.

The results of the synoptic classifications are analysed in terms of the PCA solutions (both PC loadings and PC scores), the spatial patterns and variability of the input data within each cluster, and the temporal variability of occurrence of weather type. The classifications are then compared in terms of these characteristics and the degree of mutual class occupancy.

A key consideration in the interpretation of the synoptic classifications is the role of observational error within the radiosonde record. Although observational errors cannot be quantified directly (Collins, 2001), published accuracy estimates (OFCM, 1997) allow examination of the effects of possible errors in radiosonde observations on the PCA results. We used a Monte Carlo method to examine these effects using random perturbations within the individual variable accuracies (0.5 °C for 850 hPa air temperature, 15 m for 500 hPa geopotential height, and 1.5 m $s^{-1}$ for 200 hPa wind speed).

### *3.2. Projection of RSONDE classification*

In addition to development of independent synoptic classifications from the two data sets, a second analysis is performed in which the PC scores from the NCEPR are projected onto the RSONDE classification using the characteristics of the RSONDE clusters. This analysis is undertaken to assess whether the differences in the classifications are due to errant assignment of data to clusters or due to differences in the fundamental 'modes' present in the data sets as represented by the PC loadings and scores. This process involves defining the seeds for the *k*-means clustering of NCEPR data using the results of the RSONDE classification. Rather than using the cluster centroids from the NCEPR classification, new centroids are defined by computing the mean NCEPR PC scores for each of the RSONDE clusters. These new means are then used as the seeds for the *k*-means cluster analysis. The resulting clusters are then compared with the independently derived classifications.

## 4. RESULTS

### *4.1. Analysis of the independent classifications*

*4.1.1. Modes of variability as manifest in the results of principal components analysis.*. On the basis of the truncation methods described in Section 3.1, six and nine PCs are retained for the RSONDE and NCEPR classifications respectively. These PCs explain 82.6% and 91.0% of the variance of the original RSONDE and NCEPR data sets respectively, and in each case the first three PCs explain 72.0% and 75.0% of the original variance respectively. The truncated, unrotated solutions are then subjected to a Varimax orthogonal rotation. To interpret the PCA solution, the matrices of PC loadings and PC scores are examined; to compare the two

 

PCA solutions, Pearson product-moment correlation coefficients $r$ between RSONDE PCs and NCEPR PCs are computed.

As shown in Table I, each of the PCs from the NCEPR solution shows a high degree of correspondence ($|r|$ of PC scores that is greater than 0.65 and significant at the 99% confidence level) with at least one PC from the RSONDE data set. Because autocorrelation in the PC time series can result in increased significance, we compute an effective sample size using the method of Dawdy and Matalas (1964). The results presented in Table I imply that the fundamental modes of variability (location and characteristics of synoptic phenomena) in the two data sets are similar. For example, in each data set the first PC exhibits strong loadings on 500 hPa geopotential heights, with high loadings on 850 hPa air temperatures over the western half of the study area ($r = 0.97$ for PC1 from the two data sets). PC2 in the RSONDE solution is most highly correlated with PC4 of the NCEPR solution ($r = 0.81$), which exhibits strong loadings on 200 hPa wind speeds in the northwest corner of the study area. Examination of RSONDE PC2 and NCEPR PC4 suggests that this mode is associated with the migration of the polar jet stream in and out of the study area. When the PC scores are positive, the jet stream is located over the southern part of the study area. Negative scores are associated with warmer conditions when the jet stream has been displaced north of the study area. The solutions also exhibit strong similarity with respect to PC3 loadings — strong negative loadings on 850 hPa air temperatures and 500 hPa geopotential heights in the eastern part of the domain. The positive (negative) mode of this PC is thus associated with low (high) pressure to the east of the study area and subsequent cold (warm) air advection. As might be expected from the PC3 loadings, the PC3 scores exhibit seasonal behaviour opposite to that of PC1. PC4 in the RSONDE solution corresponds to NCEPR PC2 ($r = 0.81$), with strong loadings on 200 hPa wind speeds in the southwest part of the study area. RSONDE PC5 is correlated with both NCEPR PC5 ($r = -0.61$) and NCEPR PC8 ($r = -0.66$). Whereas RSONDE PC5 shows strong links to wind speed in the eastern part of the study area, NCEPR PC5 loads strongly on wind speeds in the northeastern part of the study area and NCEPR PC8 loads strongly on wind speeds in the southeastern part of the study area. RSONDE PC6 is strongly correlation with both NCEPR PC6 ($r = -0.75$) and NCEPR PC7 ($r = 0.49$). RSONDE PC6 exhibits strong loadings on 850 hPa air temperatures in the western half of the study area, and NCEPR PC7 and PC9 exhibit strong loadings on 850 hPa air temperatures in the southwest and northwest regions of the domain respectively.

To assess the importance of errors in the radiosonde data, the data were randomly perturbed within individual variable accuracies (Section 3) to produce 100 'synthetic' data sets which were then subjected to PCA. In each case this analysis produced six significant PCs exhibiting similar spatial and temporal characteristics to those derived from the RSONDE data, suggesting that the instrumental errors do not obscure the modes of variability in the RSONDE data.

*4.1.2. Synoptic types resulting from the cluster analysis.* PC scores represent the relationship between each PC and the individual observations. Therefore, clustering the PC scores will yield groups that contain similar

Table I. Pearson product-moment correlations between the NCEPR and RSONDE PC scores

| | RSONDE1 | RSONDE2 | RSONDE3 | RSONDE4 | RSONDE5 | RSONDE6 |
|---|---|---|---|---|---|---|
| NCEPR1 | 0.974** | −0.015 | 0.006 | 0.092** | −0.031 | 0.004 |
| NCEPR2 | −0.065* | 0.168** | 0.017 | 0.808** | 0.113** | −0.068* |
| NCEPR3 | −0.035 | −0.072* | 0.962** | 0.085** | 0.033 | −0.047 |
| NCEPR4 | −0.073* | 0.805** | −0.015 | 0.045 | −0.167** | 0.064* |
| NCEPR5 | −0.016 | −0.202** | −0.102** | 0.426** | −0.608** | 0.188** |
| NCEPR6 | 0.077* | 0.147** | −0.071* | 0.019 | 0.195** | −0.754** |
| NCEPR7 | 0.110** | 0.255** | 0.127** | −0.149** | 0.146** | 0.488** |
| NCEPR8 | 0.040 | 0.204** | 0.161** | −0.253** | −0.659** | −0.247** |
| NCEPR9 | −0.092** | −0.196** | −0.007 | −0.015 | −0.184** | −0.186** |

*: Significant at 95% confidence level; **: significant at 99% confidence level.

 

days with respect to the input data. As discussed in Section 3.1, hierarchical clustering methods initially treat each observation as a cluster and merge observations until only one cluster remains. Because hierarchical methods do not allow reclassification of cluster members, *k*-means cluster analysis is also applied at each clustering level. In both classifications, the application of *k*-means clustering results in decreased within-cluster variability.

The optimal solution for terminating clustering is chosen such that the clustering is stopped just before the merger of two distant clusters. In this analysis, the primary tool used to determine the optimal solution was to plot the increase in within-cluster standard deviation of the first few PC scores as a function of cluster number (Figure 3). These plots suggest a 14-cluster solution for the RSONDE data and a 15-cluster solution for the NCEPR data. In both classifications, the solution chosen corresponds to large increases in the variance of the first few PCs.

The 14-class RSONDE and 15-class NCEPR solutions are described in detail in Tables II and III, respectively in terms of weather-type size, and a characterization of each class in terms of the dominant relationships between the class and the PC solution and between the class and the input data. These tables also summarize the data in terms of the seasonality of occurrence of each weather type, where the summer is arbitrarily divided into four equal temporal bins according to year-day (100–150, 151–200, 201–250, and 251–300).

Tables II and III highlight some of the characteristics of the individual classifications, but it is the similarities and differences between the classifications that are of primary interest in this study. Table IV shows the relationships between cluster members in the two classifications (i.e. mutual class occupancy) and suggests that there are similarities between the classifications, including several cases in which nearly identical clusters are extracted from both the RSONDE and NCEPR PCA solutions (e.g. RSONDE 14 and NCEPR 8; RSONDE 10 and NCEPR 3).

Five of 14 synoptic types defined using the RSONDE data share over half of their members (observations) with just one class from the NCEPR classification (Table IV) (RSONDE 2 and NCEPR 11; RSONDE 3 and NCEPR 12; RSONDE 4 and NCEPR 15; RSONDE 10 and NCEPR 3; and RSONDE 14 and NCEPR 8). Eight of the remaining nine RSONDE classes share over half of their members with just two NCEPR classes (RSONDE 1 with NCEPR 7 and 9; RSONDE 5 with NCEPR 5 and 14; RSONDE 6 with NCEPR 10 and 1; RSONDE 7 with NCEPR 2 and 7; RSONDE 9 with NCEPR 13 and 1; RSONDE 11 with NCEPR 12 and 6; RSONDE 12 with NCEPR 1 and 6; and RSONDE 13 with NCEPR 9 and 14). The exception to the

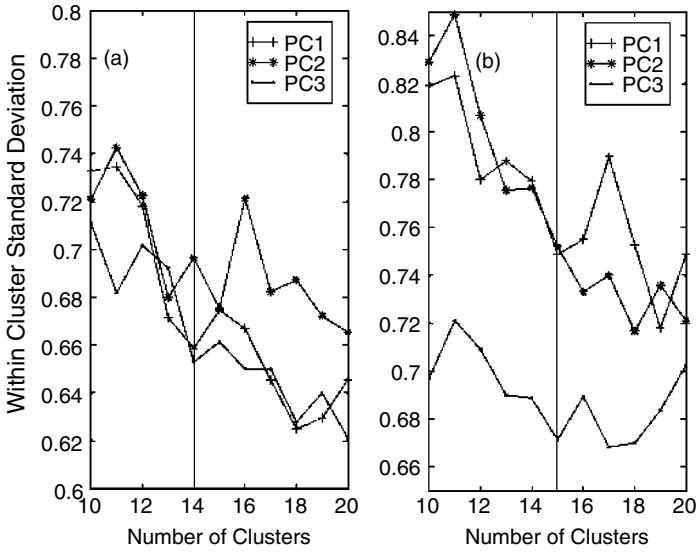


Figure 3. Clustering 'cut-off' diagrams for (a) RSONDE and (b) NCEPR. The vertical lines represent the chosen clustering level, consistent with the increases in the standard deviations of the leading PC scores

 

Table II. Description of RSONDE weather-type classification. The table shows the class size as a percentage (column 2), class size in number of member observations (column 3), relation to PCs (PC scores with magnitude greater than 0.5 are shown with the appropriate sign, and those with magnitude greater than 1.0 are shown with double symbols in columns 4–9), input data anomalies associated with each weather type for the four quadrants of the study area (high or low anomalies and regions are shown with a + or − respectively in columns 10–12), and seasonal variability (percentage of observations in the temporal bin specified by the day number at the top of columns 13–16)

| Class | Class size | | PC characteristics | | | | | | Input data characteristics | | | | | | Seasonal variability (year day) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Size (%) | Size (*n*) | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | T850 | | HT500 | | WS200 | | 100–150 | 151–200 | 201–250 | 251–300 |
| 1 | 2.9 | 43 | −− | | ++ | | + | | − | − | − | − | − | − | 88.4 | 0.0 | 0.0 | 11.6 |
| | | | | | | | | | − | − | − | − | + | + | | | | |
| 2 | 6.8 | 101 | −− | | ++ | − | ++ | − | − | + | − | + | + | − | 15.8 | 17.8 | 9.9 | 56.4 |
| | | | | | | | | | + | + | − | + | + | − | | | | |
| 3 | 4.4 | 65 | − | ++ | | | | −− | − | − | − | − | + | + | 49.2 | 26.2 | 7.7 | 16.9 |
| | | | | | | | | | − | − | − | − | + | + | | | | |
| 4 | 4.4 | 65 | + | ++ | + | + | − | | − | − | − | − | + | + | 32.3 | 1.5 | 1.5 | 64.6 |
| | | | | | | | | | − | − | − | − | + | + | | | | |
| 5 | 5.4 | 81 | | | ++ | + | + | | − | − | − | − | + | + | 19.8 | 9.9 | 24.7 | 45.7 |
| | | | | | | | | | − | − | − | − | + | + | | | | |
| 6 | 10.2 | 151 | + | | + | | | − | + | + | + | + | + | + | 7.3 | 25.8 | 47.0 | 19.9 |
| | | | | | | | | | + | + | + | + | − | − | | | | |
| 7 | 4.9 | 72 | −− | | | | − | | − | − | − | − | − | + | 69.4 | 16.7 | 0.0 | 13.9 |
| | | | | | | | | | − | − | − | − | + | + | | | | |
| 8 | 13.8 | 204 | | − | | | | | + | + | + | + | − | − | 15.7 | 40.7 | 30.4 | 13.2 |
| | | | | | | | | | + | + | + | + | − | − | | | | |
| 9 | 5.5 | 81 | | | | | −− | ++ | + | − | − | − | + | + | 29.6 | 33.3 | 24.7 | 12.4 |
| | | | | | | | | | + | + | + | − | − | + | | | | |
| 10 | 4.5 | 67 | | | | −− | − | | − | − | − | − | − | − | 74.6 | 6.0 | 0.0 | 19.4 |
| | | | | | | | | | − | − | − | − | + | + | | | | |
| 11 | 7.3 | 108 | | − | | | − | −− | − | − | + | − | + | + | 20.4 | 19.4 | 20.4 | 39.8 |
| | | | | | | | | | − | − | + | − | + | + | | | | |
| 12 | 9.9 | 146 | | − | | + | | + | + | + | + | + | + | + | 1.4 | 30.1 | 54.8 | 13.7 |
| | | | | | | | | | + | + | + | + | − | + | | | | |
| 13 | 7.1 | 104 | | + | − | + | ++ | | − | − | − | − | − | + | 15.4 | 24.0 | 41.4 | 19.2 |
| | | | | | | | | | − | − | − | − | + | + | | | | |
| 14 | 12.7 | 188 | + | | | −− | + | | + | + | + | + | − | − | 7.5 | 30.9 | 50.0 | 11.7 |
| | | | | | | | | | + | + | + | + | − | − | | | | |

Table III. Description of NCEPR weather-type classification. The table shows the class size as a percentage (column 2), class size in number of member observations (column 3), relation to PCs (PC scores with magnitude greater than 0.5 are shown with the appropriate sign in columns 4–9), input data anomalies associated with each weather type for the four quadrants of the study area (high or low anomalies and regions are shown with a + or − respectively in columns 10–12), and seasonal variability (percentage of observations in the temporal bin specified by the day number at the top of columns 13–16)

| Class | Class size | | PC characteristics | | | | | | | | | Input data characteristics | | | | | | Seasonal variability (year day) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Size (%) | Size (*n*) | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | T850 | | HT500 | | WS200 | | 100–150 | 151–200 | 201–250 | 251–300 |
| 1 | 10.8 | 159 | + | | | − | ++ | + | − | | | + | + | + | + | − | + | 12.0 | 28.9 | 33.3 | 25.8 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 2 | 6.9 | 102 | | + | | − | | + | | | | + | + | − | − | − | − | 65.7 | 22.6 | 1.0 | 10.8 |
| | | | | | | | | | | | | − | − | − | − | + | − | | | | |
| 3 | 4.7 | 69 | − | −− | | | ++ | −− | ++ | + | | − | − | − | − | − | − | 73.9 | 1.5 | 0 | 24.6 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 4 | 8.0 | 118 | + | − | | | − | | | | | + | + | + | + | + | − | 12.7 | 31.4 | 28.8 | 27.1 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 5 | 7.3 | 107 | − | − | | | | | − | | − | − | + | − | + | − | + | 7.5 | 24.3 | 55.1 | 13.1 |
| | | | | | | | | | | | | + | + | + | + | + | − | | | | |
| 6 | 8.1 | 119 | | | | | + | − | − | | | + | + | + | + | − | − | 1.7 | 35.3 | 48.7 | 14.3 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 7 | 4.1 | 61 | −− | | ++ | | −− | | + | −− | ++ | − | − | − | − | − | − | 80.3 | 3.3 | 0 | 16.4 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 8 | 10.8 | 159 | | + | + | | − | ++ | | | ++ | + | + | + | + | − | − | 6.3 | 32.1 | 52.8 | 8.8 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 9 | 5.0 | 74 | | | − | ++ | ++ | −− | ++ | −− | | − | − | − | − | − | + | 37.8 | 10.8 | 18.9 | 32.4 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 10 | 6.9 | 102 | | + | − | | | | + | − | − | + | + | + | + | + | + | 6.9 | 35.3 | 47.1 | 10.8 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 11 | 5.2 | 77 | | | ++ | | | | | − | | − | + | − | + | + | + | 22.1 | 15.6 | 7.8 | 54.6 |
| | | | | | | | | | | | | − | + | − | + | + | − | | | | |
| 12 | 5.6 | 82 | − | − | | − | | | + | ++ | | − | − | − | − | + | + | 40.2 | 18.3 | 3.7 | 37.8 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 13 | 7.8 | 115 | | | | | | −− | − | | + | + | + | + | − | + | − | 1.7 | 39.1 | 50.4 | 8.7 |
| | | | | | | | | | | | | + | + | + | + | − | + | | | | |
| 14 | 4.2 | 62 | − | − | − | + | + | + | | | | − | − | − | − | − | + | 30.7 | 17.7 | 11.3 | 40.3 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 15 | 4.7 | 70 | | | ++ | + | | | | | − | − | − | − | − | + | + | 24.3 | 2.9 | 4.3 | 68.6 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |

Table IV. Mutual class occupancy for RSONDE and NCEPR weather-type classifications. The numbers shown in the table indicate the number of coincident days in the two data sets which are classified into the two specified clusters. Hence, large numbers represent clusters sharing a large number of observations and small numbers indicate few shared observations

| NCEPR | RSONDE | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| 1 | 2 | 1 | 0 | 5 | 2 | 33 | 0 | 21 | 18 | 1 | 7 | 63 | 0 | 6 |
| 2 | 4 | 2 | 4 | 0 | 0 | 3 | 26 | 46 | 8 | 6 | 0 | 1 | 1 | 1 |
| 3 | 6 | 0 | 0 | 12 | 0 | 1 | 0 | 0 | 3 | 43 | 4 | 0 | 0 | 0 |
| 4 | 0 | 22 | 1 | 0 | 0 | 19 | 0 | 49 | 1 | 2 | 4 | 4 | 0 | 16 |
| 5 | 0 | 7 | 2 | 0 | 33 | 18 | 8 | 15 | 1 | 0 | 0 | 12 | 7 | 4 |
| 6 | 0 | 1 | 1 | 2 | 0 | 3 | 0 | 32 | 2 | 4 | 19 | 28 | 2 | 25 |
| 7 | 19 | 0 | 4 | 3 | 9 | 2 | 16 | 0 | 3 | 4 | 0 | 0 | 1 | 0 |
| 8 | 0 | 2 | 1 | 0 | 0 | 1 | 0 | 15 | 0 | 1 | 1 | 1 | 16 | 121 |
| 9 | 11 | 0 | 0 | 1 | 2 | 2 | 9 | 2 | 3 | 3 | 1 | 0 | 31 | 9 |
| 10 | 0 | 2 | 0 | 0 | 2 | 61 | 1 | 7 | 13 | 0 | 2 | 14 | 0 | 0 |
| 11 | 0 | 61 | 2 | 0 | 6 | 5 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| 12 | 0 | 2 | 38 | 0 | 1 | 0 | 1 | 0 | 3 | 0 | 35 | 0 | 2 | 0 |
| 13 | 0 | 1 | 7 | 0 | 2 | 2 | 2 | 17 | 26 | 3 | 14 | 19 | 20 | 2 |
| 14 | 1 | 0 | 5 | 0 | 18 | 0 | 8 | 0 | 0 | 0 | 3 | 1 | 24 | 2 |
| 15 | 0 | 0 | 0 | 42 | 6 | 1 | 0 | 0 | 0 | 0 | 16 | 3 | 0 | 2 |

similarity in the classifications is the largest RSONDE class (RSONDE 8), which represents positive 850 hPa air temperature and 500 hPa geopotential height anomalies and negative 200 hPa wind speed anomalies over the entire domain.

Similarly, five of the 15 synoptic types defined using the NCEPR data share over half their members (observations) with just one class from the RSONDE analysis (NCEPR 3 and RSONDE 10; NCEPR 8 and RSONDE 14; NCEPR 10 and RSONDE 6; NCEPR 11 and RSONDE 2; and NCEPR 15 and RSONDE 4). A further eight share more than half of their members with only two classes as defined using the RSONDE data (NCEPR 1 with RSONDE 12 and 6; NCEPR 2 with RSONDE 8 and 7; NCEPR 4 with RSONDE 8 and 2; NCEPR 6 with RSONDE 8 and 12; NCEPR 7 with RSONDE 1 and 7; NCEPR 9 with RSONDE 13 and 1; NCEPR 12 with RSONDE 3 and 11; and NCEPR 14 with RSONDE 13 and 5). NCEPR 5 and NCEPR 13 feature wind speed anomalies (negative in the northwest and southeast corners of the domain and positive elsewhere for NCEPR 5 and the opposite pattern from NCEPR 13) that are not captured in the classification derived from the RSONDE network.

As anticipated, the greatest degree of coherence between the classifications is found for the types exhibiting either strong high pressure (e.g. RSONDE 14, NCEPR 8; see Tables II and III and Figure 4) or strong low pressure (RSONDE 10, NCEPR 3), suggesting that both classifications are most able to resolve these features and their respective temporal variability within the summer (see Tables II and III). Four distinct low-pressure types with different wind speed anomalies are captured by both classifications. RSONDE 10 and NCEPR 3 exhibit negative 500 hPa height anomalies with negative 200 hPa wind speed anomalies in the northern part of the domain and positive 200 hPa wind speed anomalies in the southern part of the domain; RSONDE 1 and NCEPR 7 exhibit the same 200 hPa wind speed anomalies as RSONDE 10 and NCEPR 3, but with much colder 850 hPa air temperatures; RSONDE 3 and NCEPR 12 have positive 200 hPa wind speed anomalies over the entire domain; RSONDE 4 and NCEPR 15 also have positive 200 hPa wind speed anomalies over the entire domain, but again with much colder 850 hPa air temperatures.

The largest discrepancies between the RSONDE and NCEPR most likely result from differences in the resolution of the RSONDE and NCEPR input data. Poorly resolved RSONDE weather-type observations (e.g. RSONDE 8) are assigned to several types derived from NCEPR and, as discussed above, poorly resolved NCEPR types include information that is not fully resolved by the coarser radiosonde network.

 

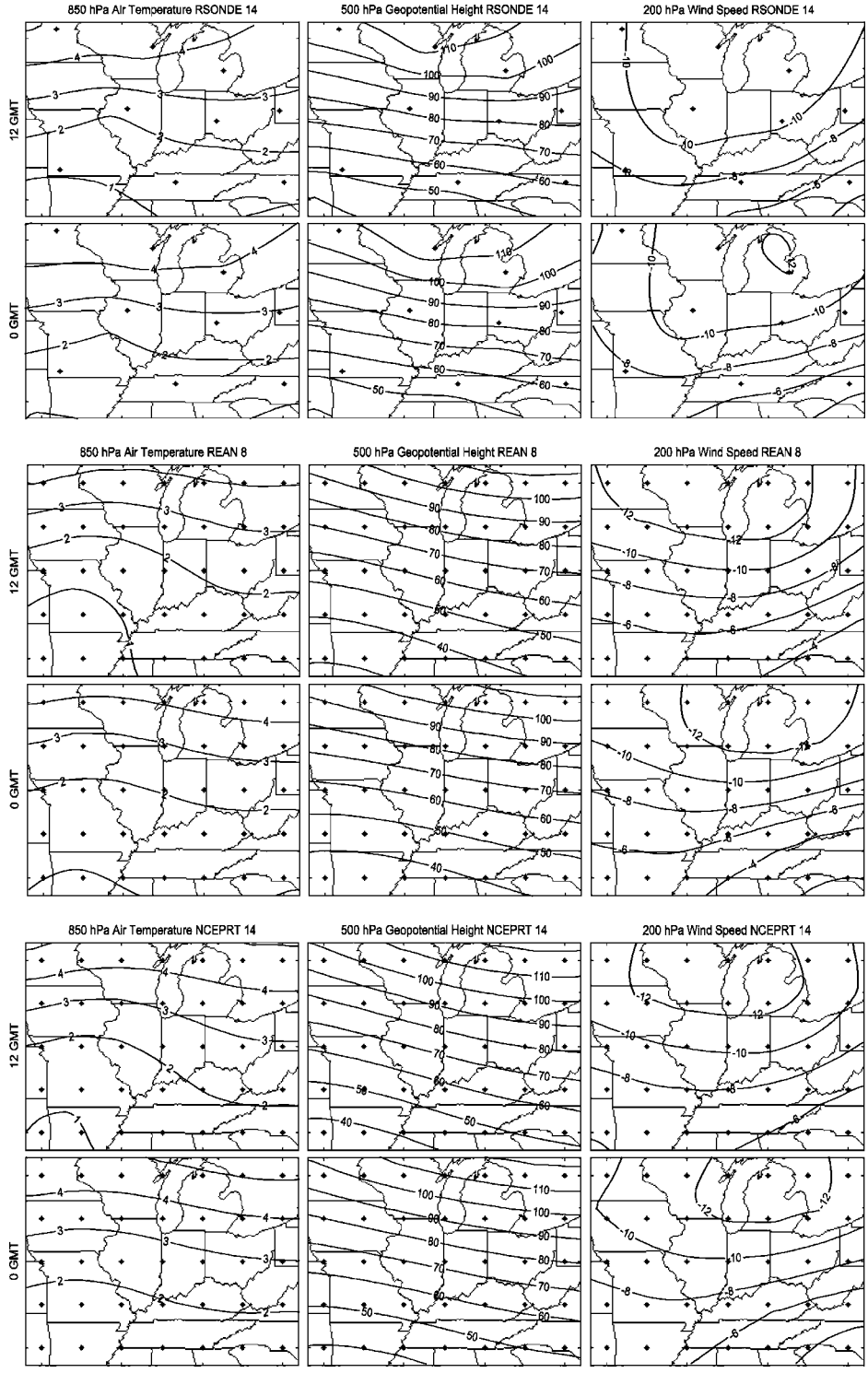


Figure 4. Input data anomalies (deviations from the mean of all observations) for RSONDE 14, NCEPR 8, and NCEPRT 14. The figures provide an example of good agreement between the classifications. Each plot shows the mean anomaly for each input variable for observations assigned to this class. The contours for RSONDE 14 should be interpreted with caution in areas that are not well resolved by the radiosonde network

 

### *4.2. Analysis of the targeted classification*

In addition to the RSONDE and NCEPR classifications discussed in Section 4.1, a third 'targeted' classification (NCEPRT) is undertaken. This analysis uses the centroids defined by computing the mean NCEPR PC scores for each of the RSONDE clusters as the seeds for the *k*-means cluster analysis. In this way, it is possible to deduce whether discrepancies in the previously discussed classifications resulted from failure of the NCEPR PCA solution to capture the most important modes of variability in the input data, or to errant assignment of individual observation days to clusters. Details of the NCEPRT 14-class solution, are shown in Table V including the percentage of input observations associated with each weather type, the total number of observations in each weather type, a characterization of each class in terms of the dominant relationships between the weather type and the PC solution and between the weather-type and the input data, and an evaluation of weather-type seasonality.

The application of this third classification slightly improves the agreement between the NCEPR classification and that derived from the RSONDE data (Table VI). Seven of the 14 RSONDE classes share over 50% of their members (observations) with just one class from the NCEPRT classification (RSONDE 1 and NCEPRT 1; RSONDE 2 and NCEPRT 2; RSONDE 3 and NCEPRT 3; RSONDE 4 and NCEPRT 4; RSONDE 10 and NCEPRT 10; RSONDE 13 and NCEPRT 13; and RSONDE 14 and NCEPRT 14). Five additional RSONDE classes share over half of their members with just two NCEPRT classes (RSONDE 5 with NCEPRT 5 and 13; RSONDE 6 with NCEPRT 6 and 9; RSONDE 9 with NCEPRT 9 and 12; RSONDE 11 with NCEPRT 3 and 11; and RSONDE 12 with NCEPRT 12 and 11). Six of the NCEPRT classes share over 50% of their members (observations) with just one class from the RSONDE classification (NCEPRT 2 and RSONDE 2; NCEPRT 4 and RSONDE 4; NCEPRT 6 and RSONDE 6; NCEPRT 10 and RSONDE 10; NCEPRT 13 and RSONDE 13; and NCEPRT 14 and RSONDE 14) and six additional NCEPRT classes share over half of their members with just two RSONDE classes (NCEPRT 1 with RSONDE 1 and 7; NCEPRT 3 with RSONDE 3 and 11; NCEPRT 5 with RSONDE 5 and 2; NCEPRT 7 with RSONDE 7 and 8; NCEPRT 8 with RSONDE 8 and 6; and NCEPRT 12 with RSONDE 12 and 6). Again, the classes with the highest correspondence represent deep cyclones and anticyclones; e.g. the strong high-pressure class (RSONDE 14 and NCEPRT 14; see Figure 4) and the four distinct low-pressure classes discussed above (Section 4.1). Although the targeted classification improves the level of agreement between the remaining weather types, the correspondence between the classifications is far from complete. For example, RSONDE 8 is not well discriminated in either the NCEPR free or targeted analyses (Figure 5), and both reanalysis-based classifications have classes whose members are not well discriminated with the RSONDE classification (e.g. NCEPR 13, NCEPRT 9). Therefore, although there is some similarity between the two solutions, a solution approaching one-to-one correspondence is not obtained. This result suggests that though the NCEPRT (and therefore NCEPR) PCA solution captures similar modes of variability to those present in the RSONDE solution, the clustering algorithms do not extract the same clusters from the two input data sets.

## 5. DISCUSSION

In Section 4 we showed that although classifications based on radiosonde data and the NCEPR data sets identify several common weather types, there is considerable dispersion between the two classifications, with several of the classes defined from the NCEPR data set exhibiting features and group membership with multiple classes in the RSONDE analysis, and *vice versa*. This indicates substantial differences in the synoptic types identified in the two data sets. Possible reasons for these discrepancies are many but include the following:

- Within-type variability as a result of both grid-point density and classification of continuous data into discrete classes (e.g. Yarnal, 1993; Brinkmann, 1999b). The NCEPR data are on a $2.5^\circ \times 2.5^\circ$ grid, but the radiosonde data used in this study were not gridded. Therefore, there are two fundamental differences in the input data sets: number of points (35 for NCEPR and nine for RSONDE) and distance between points (considerably larger distances for the radiosonde observations; see Figure 1). As expected (Yarnal, 1993),

 

Table V. Description of NCEPRT weather-type classification. The table shows the class size as a percentage (column 2), class size in number of member observations (column 3), relation to PCs (PC scores with magnitude greater than 0.5 are shown with the appropriate sign in columns 4–9), input data anomalies associated with each weather type for the four quadrants of the study area (high or low anomalies and regions are shown with a + or − respectively in columns 10–12), and seasonal variability (percentage of observations in the temporal bin specified by the day number at the top of columns 13–16)

| Class | Class size | | PC characteristics | | | | | | | | | Input data characteristics | | | | | | Seasonal variability (year day) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Size (%) | Size (*n*) | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | T850 | | HT500 | | WS200 | | 100–150 | 151–200 | 201–250 | 251–300 |
| 1 | 3.3 | 49 | −− | | ++ | − | − | − | | | | − | − | − | − | − | − | 81.6 | 2.0 | 0 | 16.3 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 2 | 6.5 | 96 | | ++ | − | + | | | + | −− | − | − | + | − | + | + | − | 18.8 | 21.9 | 7.3 | 52.1 |
| | | | | | | | | | | | | + | + | − | + | + | − | | | | |
| 3 | 5.7 | 84 | − | | | ++ | | + | | + | | − | − | − | − | + | + | 40.5 | 21.4 | 1.2 | 36.9 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 4 | 5.2 | 77 | | | ++ | | | − | | | | − | − | − | − | + | + | 23.4 | 5.2 | 3.9 | 67.5 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 5 | 5.8 | 85 | | | | + | ++ | −− | + | | | − | + | − | + | − | + | 18.8 | 20.0 | 40.0 | 21.2 |
| | | | | | | | | | | | | − | + | − | + | + | − | | | | |
| 6 | 7.7 | 114 | | + | | − | + | | | − | | + | + | + | + | + | + | 9.7 | 33.3 | 45.6 | 11.4 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 7 | 5.0 | 74 | −− | | | | | | | | | − | − | − | − | − | − | 64.9 | 16.2 | 0 | 18.9 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 8 | 10.0 | 147 | | | − | | − | + | | | | + | + | + | + | + | − | 23.1 | 27.2 | 31.3 | 18.4 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 9 | 8.7 | 129 | | | | | + | + | −− | + | + | − | − | + | − | + | − | 7.8 | 36.4 | 45.0 | 10.9 |
| | | | | | | | | | | | | + | + | + | + | + | + | | | | |
| 10 | 4.8 | 71 | − | | ++ | | − | ++ | | | + | − | − | − | − | − | − | 77.5 | 1.4 | 0 | 21.1 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 11 | 7.5 | 110 | + | + | + | + | | | | + | | + | + | + | + | + | − | 3.6 | 30.0 | 51.8 | 14.6 |
| | | | | | | | | | | | | + | − | + | + | − | − | | | | |
| 12 | 10.8 | 160 | + | | | | + | | | | | + | + | + | + | − | + | 10.0 | 31.9 | 34.4 | 23.8 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |
| 13 | 7.1 | 104 | | − | | | | | | ++ | | − | − | − | − | − | + | 32.7 | 11.5 | 18.3 | 37.5 |
| | | | | | | | | | | | | − | − | − | − | + | + | | | | |
| 14 | 11.9 | 176 | + | − | | | − | | | | | + | + | + | + | − | − | 3.4 | 35.2 | 54.6 | 6.8 |
| | | | | | | | | | | | | + | + | + | + | − | − | | | | |

Table VI. Mutual class occupancy table for RSONDE and NCEPRT weather-type classifications. The numbers shown in the table indicate the number of coincident days in the two data sets which are classified into the two specified clusters. Hence, large numbers represent clusters sharing a large number of observations and small numbers indicate few shared observations

| NCEPRT | RSONDE | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| 1 | 24 | 0 | 0 | 3 | 9 | 0 | 11 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 2 | 0 | 70 | 3 | 0 | 3 | 4 | 1 | 9 | 0 | 0 | 4 | 0 | 0 | 2 |
| 3 | 0 | 2 | 38 | 0 | 3 | 0 | 4 | 0 | 2 | 0 | 30 | 1 | 4 | 0 |
| 4 | 0 | 0 | 0 | 44 | 7 | 1 | 0 | 0 | 0 | 0 | 20 | 3 | 0 | 2 |
| 5 | 0 | 18 | 2 | 0 | 33 | 3 | 9 | 10 | 0 | 0 | 2 | 3 | 4 | 1 |
| 6 | 0 | 4 | 0 | 0 | 4 | 69 | 4 | 7 | 14 | 0 | 2 | 10 | 0 | 0 |
| 7 | 3 | 4 | 10 | 0 | 1 | 0 | 21 | 20 | 6 | 6 | 0 | 0 | 2 | 1 |
| 8 | 1 | 0 | 2 | 0 | 0 | 29 | 4 | 73 | 5 | 2 | 3 | 14 | 1 | 13 |
| 9 | 0 | 1 | 8 | 0 | 7 | 4 | 5 | 17 | 27 | 3 | 12 | 20 | 22 | 3 |
| 10 | 6 | 0 | 0 | 11 | 0 | 0 | 0 | 2 | 4 | 44 | 4 | 0 | 0 | 0 |
| 11 | 0 | 1 | 2 | 3 | 0 | 0 | 0 | 25 | 1 | 7 | 27 | 24 | 2 | 18 |
| 12 | 1 | 0 | 0 | 3 | 0 | 38 | 1 | 21 | 19 | 1 | 2 | 68 | 1 | 5 |
| 13 | 8 | 0 | 0 | 1 | 14 | 0 | 12 | 0 | 2 | 3 | 1 | 0 | 55 | 8 |
| 14 | 0 | 1 | 0 | 0 | 0 | 3 | 0 | 20 | 0 | 0 | 1 | 3 | 13 | 135 |

the larger number of grid points results in a greater number of significant PCs and more weather types in the NCEPR classification.

- Sampling scale. In principle, the classification derived from the NCEPR may be able to resolve smaller scale circulation features, due to the higher spatial resolution. Each grid-point value in the NCEPR data is representative of a $2.5^\circ \times 2.5^\circ$ box, whereas radiosonde measurements are representative of a small column of the atmosphere and may, therefore, be influenced by small-scale disturbances. With input data from many sources, the output grid resolution is only one factor in the spatial resolution of the NCEPR. Whereas examination of maps associated with individual days suggests comparable features between the two networks, the clustering results suggest sampling scale, and perhaps more resolved regions in the NCEPR, may play a key role in the classification differences.
- Differences in the quality control systems used on the two data sets. By excluding time periods with large numbers of radiosonde errors and by testing the sensitivity of radiosonde measurements to instrument accuracies, we have attempted to minimize the use of observations that may have been replaced or altered for inclusion in the NCEPR.
- Quality of observations for individual variables. Both data sets used in this comparison have strengths and weaknesses. With the availability of aircraft observations in the mid-1980s, the NCEPR data may provide a better representation of 200 hPa wind speeds (although at a different sampling scale) than the radiosondes. However, poorly resolved land surfaces in the NCEPR suggest that radiosonde-based 850 hPa air temperatures may be more realistic. These are important considerations for the interpretation of the solutions and their similarities/differences.
- Published work on the rejection of radiosonde data by NCEP–NCAR quality control is limited to radiosonde temperatures and geopotential heights. Tables II, III, and V suggest that many of the important differences in the RSONDE and NCEPR classifications are related to differences in the 200 hPa wind speed fields.
- Differential ratio of signal to noise in the two data sets. Recent work (e.g. Paatero and Hopke, in press) suggests that differentially noisy variables in the individual input data sets might be manifest as variations in the classifications. If the level of noise between corresponding input variables is not consistent between the two input data sets, then this could also impact the comparison of the resulting classifications.
- Systematic differences in terms of the synoptic-scale phenomena as manifest in the two data sets. The comparison of the PCs derived from the two data sets implies a relatively high degree of correspondence in

 

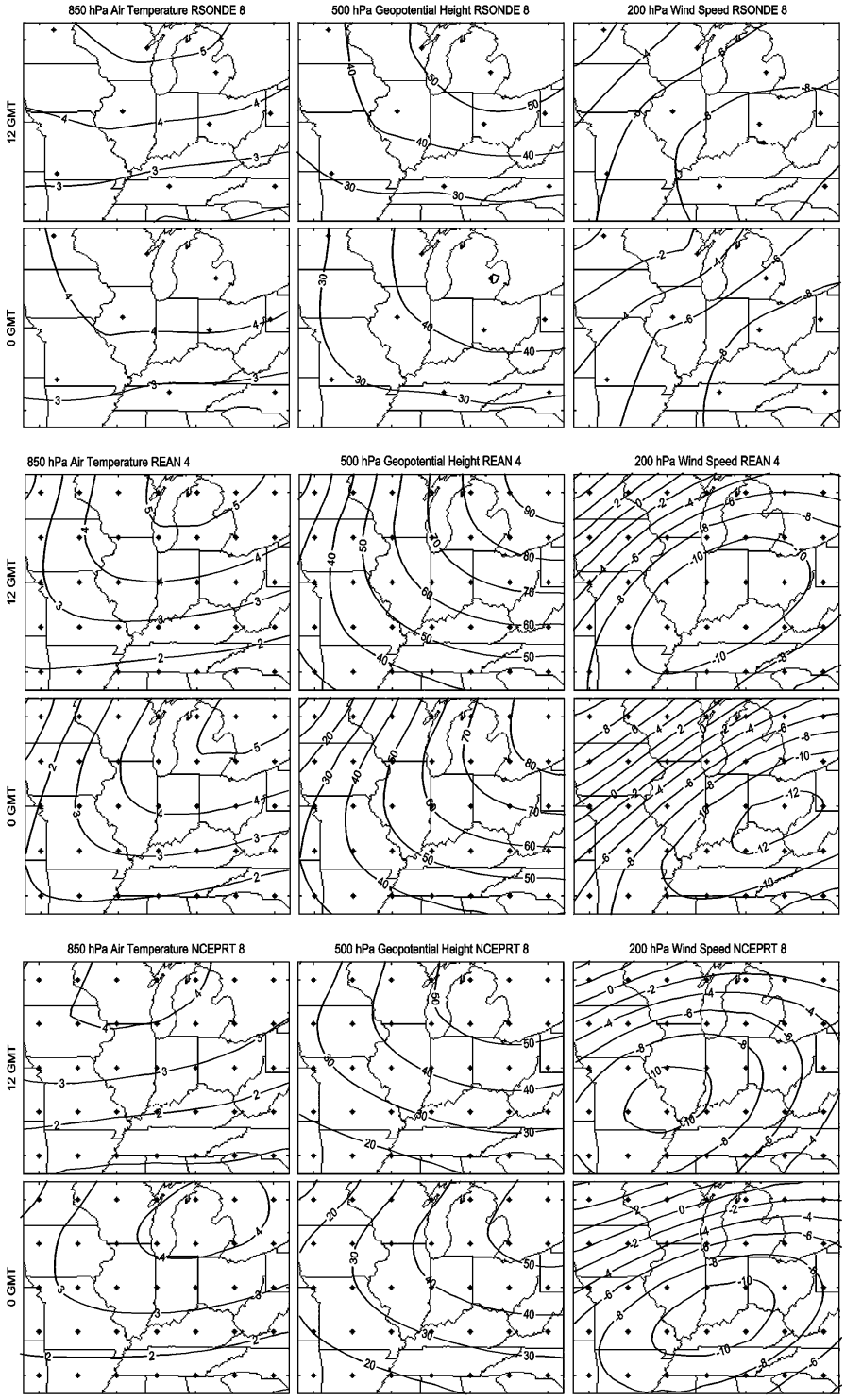


Figure 5. Input data anomalies (deviations from the mean of all observations) for RSONDE 8, NCEPR 4, and NCEPRT 8. The figures provide an example of poor agreement between the classifications. Each plot shows the mean anomaly for each input variable for observations assigned to this class. The contours for RSONDE 8 should be interpreted with caution in areas that are not well resolved by the radiosonde network

 

the modes of variability as measured using correlation of the PC scores. In addition, maps of the means and standard deviations of the input variables show general agreement (Figure 2). However, these measures do not guarantee that the individual days within the input data exhibit similar anomalies, or that the anomalies are of similar magnitude.

## 6. CONCLUSIONS

In this study, we evaluate the ability of the NCEPR to represent the synoptic climate of the Midwestern USA relative to radiosonde data. A multi-method approach is used, which allows both examination of independent classifications of the input fields and an examination of the improvement achieved by projecting the NCEPR data onto the classification derived using the RSONDE data.

Using automated synoptic classifications based on rotated PCA and a two-step clustering algorithm, a 15-class NCEPR synoptic classification and a 14-class RSONDE classification are developed. The results of the synoptic classifications are examined in terms of similarities and differences in the PCA solutions, the spatial patterns and variability of input variables within each weather type, and the temporal variability of the occurrence of each weather type. The classifications are then compared in terms of these characteristics and the degree of mutual class occupancy. Although the classifications are able to identify a number of similar weather types (in terms of the input data, PCA solution, and mutual occupancy), the classifications exhibit substantial differences and large within-type variability.

To analyse whether the differences in the classifications are due to errant assignment of data to clusters or to differences in the fundamental 'modes' present in the data sets as represented by the PC loadings and scores, a third classification categorizes the NCEPR data according to the radiosonde PCA solution using the means of the reanalysis PC scores for each of the radiosonde clusters as seeds for the NCEPR cluster analysis. The application of this third classification substantially improves the agreement between the classifications (in terms of both interpretability and mutual class occupancy), although the solutions still exhibit considerable discrepancies.

The results of this analysis suggest that the NCEPR data set exhibits similar modes of variability to those present in the radiosonde data but that the description of the intensity and details of synoptic-scale phenomena differ sufficiently to merit caution when using the NCEP–NCAR data to resolve synoptic-scale phenomena and climate relative to radiosondes. Several possible reasons for the differences in the classifications have been identified. The discrepancies documented here between the synoptic classifications derived from the RSONDE and NCEPR data sets have profound implications for climate-change research focused on GCM evaluation, since the degree of correspondence between the synoptic-scale climate as resolved in GCM simulations and the current synoptic-scale climate will be critically dependent on the data series used to derive the synoptic-scale conditions.

ACKNOWLEDGEMENTS

This research was supported in part by the Office of Science (BER), US Department of Energy, through the Midwestern Regional Center of the National Institute for Global Environmental Change under Cooperative Agreement No. DE-FC03-90ER61010 via a grant to Pryor, Barthelmie and Carreiro. We gratefully acknowledge the constructive criticisms of two anonymous reviewers.

REFERENCES

Bonell M, Sumner G. 1992. Autumn and winter daily precipitation areas in Wales. *International Journal of Climatology* **12**: 77–102.

Boyle JS. 1998. Evaluation of the annual cycle of precipitation over the United States in GCMs: AMIP simulations. *Journal of Climate* **11**: 1041–1055.

Brinkmann WAR. 1999a. Application of non-hierarchically clustered circulation components to surface weather conditions: Lake Superior Basin winter temperatures. *Theoretical and Applied Climatology* **63**: 41–56.

Brinkmann WAR. 1999b. Within-type variability of 700-hPa winter circulation patterns over the Lake Superior Basin. *International Journal of Climatology* **19**: 41–58.

 

Calinski RB, Harabasz J. 1974. A dendrite method for cluster analysis. *Communications in Statistics* **3**: 1–27.
Catell RB. 1966. The scree test for the number of factors. *Multivariate Behavioral Research* **1**: 245.
Collins WG. 2001. The operational complex quality control of radiosonde heights and temperatures at the National Centers for Environmental Prediction. Part 1: description of the method. *Journal of Applied Meteorology* **40**: 137–168.
Craddock JM, Flood CR. 1969. Eigenvectors for representing the 500-mb geopotential surface over the Northern Hemisphere. *Quarterly Journal of the Royal Meteorological Society* **95**: 576–593.
D'Andrea F, Tibaldi S, Blackburn M, Boer G, Deque M, Dix M, Dugas B, Ferranti L, Iwasaki T, Kitoh A, Pope V, Randall D, Roeckner E, Straus D, Stern W, van den Dool H, Williamson D. 1998. Northern Hemisphere blocking as simulated by 15 atmospheric general circulation models in the period 1979–1988. *Climate Dynamics* **14**: 385–407.
Davis RE, Kalkstein LS. 1990. Using a spatial synoptic climatological classification to assess changes in atmospheric pollution concentrations. *Physical Geography* **11**: 320–342.
Dawdy DR, Matalas NC. 1964. Statistical and probability analysis of hydrologic data, part III: analysis of variance, covariance and time series. In *Handbook of Applied Hydrology, a Compendium of Water-Resources Technology*, Chow VT (ed.). McGraw-Hill: New York; 8.68–8.90.
Fovell RG, Fovell MC. 1993. Climate zones of the United States defined using cluster analysis. *Journal of Climate* **6**: 2103–2135.
FSL–NCDC. 1997. Radiosonde data of North America, 1946–1996. http://www.ncdc.noaa.gov [28 August 2003].
Hines KM, Bromwich DH, Marshall GJ. 2000. Artificial surface pressure trends in the NCEP–NCAR reanalysis over the Southern Ocean and Artarctica. *Journal of Climate* **13**: 3940–3952.
Huth R. 2000. A circulation classification scheme applicable in GCM studies. *Theoretical and Applied Climatology* **67**: 1–18.
Kalkstein LS, Tan G, Skindlov JA. 1987. An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Climate and Applied Meteorology* **26**: 717–730.
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. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–470.
Katzfey JJ, Ryan BF. 2000. Midlatitude frontal clouds: GCM-scale modeling implications. *Journal of Climate* **13**: 2729–2745.
Kistler R, Kalnay E, Collins W, Saha S, White G, Woollen J, Chelliah M, Ebisuzaki M, Kanamitsu M, Kousky V, van den Dool H, Jenne R, Fiorino M. 2001. The NCEP–NCAR 50year reanalysis: monthly mean CD-ROM and documentation. *Bulletin of the American Meteorological Society* **82**: 247–267.
Marshall GJ. 2002. Trends in Antarctic geopotential height and temperature: a comparison between radiosonde and NCEP–NCAR reanalysis data. *Journal of Climate* **15**: 659–674.
McKendry IG, Steyn DG, McBean G. 1995. Validation of synoptic circulation patterns simulated by the Canadian Climate centre general circulation model for western North America. *Atmosphere–Ocean* **33**: 809–825.
OFCM. 1997. Federal Meteorological Handbook No. 3 — Rawinsonde and Pibal Observations, Office of the Federal Coordinator for Meteorology, FCM-H3-1997.
Overland JE, Preisendorfer RW. 1982. A significance test for principal components applied to a cyclone climatology. *Monthly Weather Review* **110**: 1–4.
Paatero P, Hopke PK. In press. Discarding or downweighting high-noise variables in factor analytic models. *Analytica Chimica Acta*.
Renfrew IA, Moore GWK, Guest PS, Bumke K. 2002. A comparison of surface layer and surface turbulent flux observations over the Labrador Sea with ECMWF analyses and NCEP reanalyses. *Journal of Physical Oceanography* **32**: 383–400.
Richman MB. 1986. Rotation of principal components. *Journal of Climatology* **6**: 293–335.
Schoof JT, Pryor SC. 2001. Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks. *International Journal of Climatology* **21**: 773–790.
Sokal RR, Michener CD. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* **38**: 1409–1438.
Stratton RA. 1999. A high resolution AMIP integration using the Hadley Centre model HadAM2b. *Climate Dynamics* **15**: 9–28.
Swail VR, Cox AT. 2000. On the use of NCEP–NCAR reanalysis surface marine wind fields for a long-term North Atlantic wave hindcast. *Journal of Atmospheric and Oceanic Technology* **17**: 532–545.
Trenberth KE, Stepanik DP. 2002. A pathological problem with NCEP reanalysis in the stratosphere. *Journal of Climate* **15**: 690–695.
Yarnal B. 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press: London.

 