2001

Downscaling Temperature and Precipitation: A Comparison of Regression-Based Methods and Artificial Neural Networks

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

S C. Pryor
Indiana University Bloomington

International Journal of Climatology, 21: 773–790 (2001).

INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **21**: 773–790 (2001)
DOI: 10.1002/joc.655

DOWNSCALING TEMPERATURE AND PRECIPITATION: A COMPARISON OF REGRESSION-BASED METHODS AND ARTIFICIAL NEURAL NETWORKS

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

Received 9 May 2000
Revised 22 January 2001
Accepted 26 January 2001

ABSTRACT

A comparison of two statistical downscaling methods for daily maximum and minimum surface air temperature, total daily precipitation and total monthly precipitation at Indianapolis, IN, USA, is presented. The analysis is conducted for two seasons, the growing season and the non-growing season, defined based on variability of surface air temperature. The predictors used in the downscaling are indices of the synoptic scale circulation derived from rotated principal components analysis (PCA) and cluster analysis of variables extracted from an 18-year record from seven rawinsonde stations in the Midwest region of the United States. PCA yielded seven significant components for the growing season and five significant components for the non-growing season. These PCs explained 86% and 83% of the original rawinsonde data for the growing and non-growing seasons, respectively. Cluster analysis of the PC scores using the average linkage method resulted in eight growing season synoptic types and twelve non-growing synoptic types. The downscaling of temperature and precipitation is conducted using PC scores and cluster frequencies in regression models and artificial neural networks (ANNs).

Regression models and ANNs yielded similar results, but the data for each regression model violated at least one of the assumptions of regression analysis. As expected, the accuracy of the downscaling models for temperature was superior to that for precipitation. The accuracy of all temperature models was improved by adding an autoregressive term, which also changed the relative importance of the dominant anomaly patterns as manifest in the PC scores. Application of the transfer functions to model daily maximum and minimum temperature data from an independent time series resulted in correlation coefficients of 0.34–0.89. In accord with previous studies, the precipitation models exhibited lesser predictive capabilities. The correlation coefficient for predicted versus observed daily precipitation totals was less than 0.5 for both seasons, while that for monthly total precipitation was below 0.65. The downscaling techniques are discussed in terms of model performance, comparison of techniques and possible model improvements.

KEY WORDS: artificial neural networks; circulation-based downscaling; Midwestern United States; regression

1. INTRODUCTION

Climate most obviously impacts human activity in the Midwestern United States through effects on agricultural practices and productivity. Although agricultural productivity varies as a function of land use and management issues (Rounsevell *et al.*, 1999), forecasting future agricultural yields requires explicit attention to climatic variability (Smit *et al.*, 1996) and climatic conditions on spatial scales much smaller than those offered by the most highly resolved general circulation models (GCMs). Extended periods of high air temperature with below normal precipitation can cause moisture stress to agricultural crops, such as corn and soybeans, and negatively impact yields (Carlson, 1990). Therefore, it is important to quantify

* Correspondence to: Atmospheric Science Program, Department of Geography, Indiana University Student Building, 701 East Kirkwood Ave., Bloomington, IN 47405-7100, USA.

how agricultural yields change as a function of climate parameters, and to examine how local climate might change as a result of global climate change.

Local meteorological conditions are directly related to synoptic-scale meteorology (also referred to as macro-β scale; 2000–20000 km; Orlanski, 1975), and although synoptic-systems are not the only factor controlling local meteorology, the local climate is embedded in the synoptic-scale. The large-scale circulation largely determines local rainfall amounts, cloud cover and other air mass characteristics. However, the links between the synoptic-scale circulation and surface conditions may be complex and highly non-linear. In order to predict meteorological variables at the regional scale, the relationships between synoptic-scale circulation and the local climate must be well understood and quantified to formulate the transfer functions necessary to reconcile these scales.

The objectives of this study are to quantify the relationships between the synoptic-scale circulation and local climate parameters in the Midwestern United States by classification of the synoptic-scale circulation followed by application of downscaling procedures. Synoptic classification is used to determine the dominant (most frequently occurring) modes of the synoptic-scale climate. Synoptic indices from this classification are then used in downscaling procedures to determine the functional dependence of local meteorological variables on the synoptic-scale circulation. It should be noted that the utility of the transfer functions derived in this research, like all circulation-based downscaling, is limited by how closely the large-scale circulation relates to the surface variable(s) of interest, and by how well the statistical methods employed can identify such relationships.

2. DATA

To study relationships between the synoptic-scale circulation and surface climate, data are needed from stations throughout the region of interest and from a station that will serve as the focus of the downscaling analysis. In order to classify the circulation properly, a large number of observations are needed and the observations must be representative of both surface and upper atmospheric conditions. In addition, the stations used in the synoptic classification must be roughly evenly distributed in space in order to properly characterize the region.

The study area for this research extends from 35° to 45°N and from 80° to 95°W, and includes Indiana (IN), Illinois (IL), Ohio (OH), Kentucky (KY), Tennessee (TN), Michigan (MI), Missouri (MO), western Pennsylvania (PA), and southeastern Wisconsin (WI) (Figure 1). This study area was chosen because of the importance of the downscaled quantities for agricultural yields in the region (Mendelsohn *et al.*, 1994). In addition, the relatively flat region reduces the effects of terrain on temperature, precipitation and flow.

Data from seven stations (Table I) within the study area were extracted from the Radiosonde data of North America (1946–1996) database issued by the Forecast Systems Laboratory and National Climatic Data Center (FSL/NCDC, 1997) for inclusion in the study. It should be noted that the rawinsonde station network is quite sparse, resulting in a lack of information in the northwest portion of the domain (Figure 1). Hence, in graphical depictions of the synoptic classes, features in this sub-domain should be

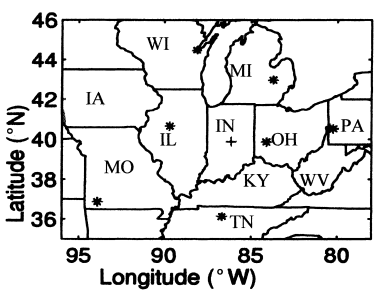

Figure 1. Map of study area showing rawinsonde stations (*) and Indianapolis, IN station (+)

Table I. Station information for rawinsonde stations

Station name	WMO station ID	Latitude (°N)	Longitude (°W)	Distance to Indianapolis (km)	Maximum length of data record (years)
Monett, MO	UMN 77349	36.88	93.90	720	26
Nashville, TN	BNA 72327	36.12	86.68	449	33
Flint, MI	ENT 72637	42.97	83.73	421	40
Peoria, IL	PIA 72532	40.67	89.68	325	37
Green Bay, WI	ORB 72645	44.48	88.13	529	51
Pittsburgh, PA	PIT 72520	40.53	80.23	520	32
Dayon, OH	DAY 72429	39.87	84.12	194	40

interpreted with caution. To ensure that the 18-year time period (1973–1990) used in this study is representative of longer periods, the probability distributions of surface air temperature and 500-hPa height at each rawinsonde station were compared to those from long-term time series for these sites. For each variable and each station, the longest possible record was extracted and compared to the 20-year record using a Kolmogorov–Smirnov (K-S) test (see Wilks, 1995). The results of the K-S test showed that the distributions of the long and short-term data were nearly identical. We therefore assume that the short-term dataset is representative of longer-term climate, and is thus sufficient to characterize accurately the synoptic-scale circulation over the Midwestern United States.

The variables chosen for use in the synoptic circulation classification are twice-daily (00:00 GMT, 12:00 GMT) observations of 500-hPa height, 700-hPa height, 850-hPa temperature, sea level pressure (SLP), plus one additional derived variable, column average relative humidity (RH). These variables were chosen to:

(i) Provide information about the vertical structure of the atmosphere.
(ii) Provide an estimate of water available for precipitation.
(iii) Use commonly available variables. It is anticipated that the techniques developed here could later be used to generate future local climate realizations from GCM simulations. Therefore, the input variables must be readily available from GCMs.

The surface data used for the downscaling analysis were extracted from the Solar and Meteorological Surface Observation Network (SAMSON) dataset issued by the National Climatic Data Center (NCDC/NREL, 1993). From the hourly surface observations, daily maximum and minimum air temperature and daily precipitation totals at the Indianapolis International Airport (hereafter referred to as Indianapolis) (39°44′N, 86°17′W) were derived (calculated from 12:00 GMT to 12:00 GMT) for the same temporal domain as the rawinsonde data (1973–1990). Daily maximum and minimum air temperature were chosen as predictands for this analysis because effects of climate on agriculture are often the result of extremes (e.g., frost and drought). In addition, climate change may be manifest as a change in the diurnal variability/range of temperature rather than the mean (Karl *et al.*, 1996).

3. METHODS

3.1. Synoptic classification

Synoptic classification is based on the concept that a large proportion of the variability at the synoptic scale can be explained by grouping observations into several classes. The synoptic classification presented here was undertaken as follows:

(i) The annual cycle was removed from the input variables (rawinsonde observations).

(ii) The observations were stratified into growing and non-growing seasons. This nomenclature is used because of the application of downscaling studies to agriculture. Rather than use actual growing season dates for local crops (which vary widely), physically based growing and non-growing seasons were determined from the surface air temperature record.
(iii) The rawinsonde data described in the previous section were subjected to principal component analysis (PCA) in order to generate representative scenarios and indices of the synoptic-scale circulation for each observational time period.
(iv) Time series of the circulation indices were clustered to determine synoptic classes where each cluster represents a frequently occurring 'weather pattern'.

Each of these analysis steps is described in more detail in the following sections.

3.1.1. Removal of the annual cycle. Because the study area is located in the mid-latitudes, the input variables exhibit annual cycles. To avoid the influence of the annual cycle on the synoptic classification, this cycle was removed by transforming the data into the frequency domain using a finite Fourier transform, applying a frequency domain filter and then using a reverse Fourier transform to bring the data back into the time domain.

3.1.2. Determination of seasons. While some classification studies have ignored the effects of seasonality on synoptic classification (e.g., Ashbaugh *et al*., 1984), separate classifications for different seasons are common in the meteorological literature (e.g., Pryor *et al*., 1995; Stratheropoulos *et al*., 1998). In this study, separation of the data into growing and non-growing seasons was necessary because:

(i) The synoptic-scale circulation in the growing-season differs from the synoptic-scale circulation in the non-growing season.
(ii) Agricultural yield is most dependent on conditions during the growing season.

The seasons were determined based on analysis of the surface air temperature at Indianapolis. The standard deviation of air temperature on each day of the year and the average standard deviation (over all days of the year) were determined from hourly data (with the seasonal cycle present). As shown in Figure 2, the average standard deviation of temperature by year day is highest during the non-growing season and intersects the annual average standard deviation in two places, dividing the year into two segments that will be referred to as the growing and non-growing seasons. Using this definition, the growing season begins on year day 100 and ends on year day 300.

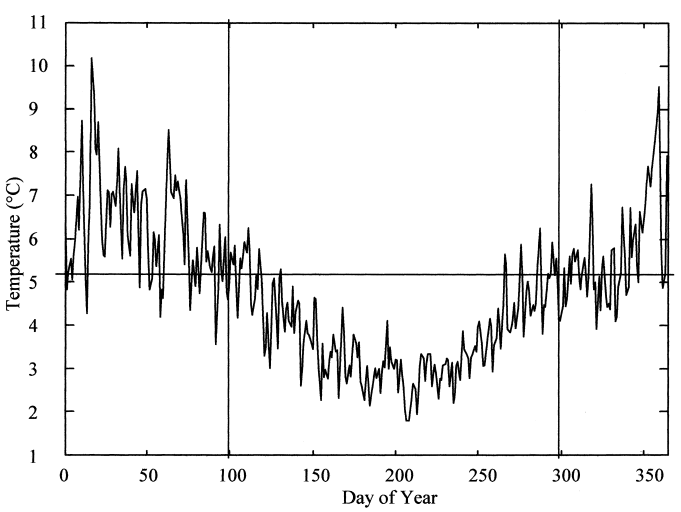

Figure 2. Standard deviation of daily surface air temperature for Indianapolis, IN. The horizontal line is the average standard deviation over the entire year. The seasons are determined by whether the mean daily standard deviation is above or below the annual mean

3.1.3. Principal components analysis. PCA has been widely applied in atmospheric science (e.g., McGregor, 1996; Stratheropoulos *et al*., 1998) as a means of data reduction and for identification of modes of variability. Subsequent to removal of the annual cycle and division of the data into growing and non-growing seasons, the 35 input variables from the rawinsonde database were subjected to correlation-based S-mode PCA (Richman, 1986). Since the first few components explain most of the variance of the original dataset, truncation rules are used to identify which of the subsequent components, order by variance explained, may be discarded. In this analysis, a combination of the logarithmic eigenvalue plot (Craddock and Flood, 1969), the Monte Carlo technique of Overland and Preisendorfer (1982) and a subjective criteria, that the components be readily interpretable, were used to identify whether the components should be retained. The unrotated solutions were then subjected to Varimax orthogonal and Harris–Kaiser (H-K) II oblique rotations, and the solutions compared using simple structure plots (Richman, 1986) and coefficients of congruence (White *et al*., 1991).

3.1.4. Cluster analysis. Cluster analysis is a data reduction/classification method that uses measures of distance to relate (and group) observations within a dataset (Gong and Richman, 1995). Here we use hierarchical methods of clustering, in which each observation is initially treated as a cluster, and then proximal clusters are merged based on intra- and inter-class similarity (e.g., Kalkstein *et al*., 1987). In this analysis, Euclidean distance is used to determine the data similarity and five algorithms, which differ in the way distances between observations are measured, were applied: (i) single linkage (which uses the minimum distance between two clusters), (ii) complete linkage (which uses the maximum distance between observations in the two clusters), average linkage (which uses either (iii) the average distance between observations in the two clusters or (iv) the average distance between points in the newly formed cluster), and (v) Ward's method (Ward, 1963), (which minimizes the within-cluster sum of squares).

3.2. Downscaling

In order to relate the large-scale weather patterns to the local scale, downscaling is necessary. The relationships between these scales can be determined by a number of methods including regression (Kilsby *et al*., 1998), canonical correlation analysis (Heyen *et al*., 1996; Xopalski *et al*., 2000), artificial neural networks (Hewitson and Crane, 1992; Gardner and Dorling, 1998; Cannon and Lord, 2000), singular value decomposition (Huth, 1999), stochastic weather generators (Hughes and Guttorp, 1994), and limited area models (Giorgi *et al*., 1990).

Because most downscaling studies are focused on application to GCM output, the predictors for such analyses must be accurately reproduced by GCMs. Thus, it is typical for downscaling studies to begin with atmospheric circulation variables (Huth, 1999). Because variations in local climate are not attributable only to circulation variables, this study employs circulation, temperature and humidity variables.

In this study, regression models and artificial neural networks (ANNs) are used to downscale daily maximum (T_{max}) and daily minimum (T_{min}) temperature and daily ($PREC_d$) and monthly ($PREC_m$) precipitation. These downscaling techniques are more cost-efficient and easier to implement than nested mesoscale models and have been shown to have comparable accuracy to those obtained with dynamical downscaling methods (e.g., Kidson and Thompson, 1998; Solman and Nuñez, 1999).

3.2.1. Multiple regression. Multiple linear regression (MLR) is the least computationally demanding downscaling technique and has been widely used in atmospheric science (e.g., Karl *et al*., 1990; Wigley *et al*., 1990; Winkler *et al*., 1997). In this application, the temperature variables (T_{max} and T_{min}) are downscaled using least squares MLR where the temperature variables are the predictands and synoptic indices (PC scores) are the predictors.

The probability distribution of precipitation is highly skewed and does not have the constant variance properties required for MLR. Therefore, precipitation was modelled using the Poisson regression model, solved using maximum likelihood methods, again with the synoptic indices (PC scores) as predictors.

3.2.2. Artificial neural networks (ANNs). ANNs are multi-layer perceptrons used to map relationships between input variables and dependent output variables. The goal of the neural network is to minimize the root mean squared error (RMSE) between the predicted and observed value of the dependent variable. An ANN is composed of an input layer, any number of hidden layers and an output layer. Typically, the number of hidden nodes is approximately 1.5 times the number of predictor variables (Eberhart and Dobbins, 1990).

Although the weights of an ANN are similar to non-linear regression coefficients (Hewitson and Crane, 1992), the simple non-linear functions that send information between nodes in a neural network allow the network to approximate extremely non-linear functions (Gardner and Dorling, 1998; Hornik *et al.*, 1989). Also, while MLR has a closed form solution, neural networks use an iterative process and while regression assumes a functional form, ANNs allow the data to define the functional form (Warner and Misra, 1996).

The network works in two modes: mapping mode and learning mode. In learning mode, training of the network begins with arbitrary values for the hidden and output node weights. An iterative process in which the weights are adjusted at each step is used to train the network. The output value is compared with the known output at the end of each iteration, or epoch, and the weights are adjusted accordingly. In mapping mode, an observation is presented to the input layer. These values are not changed and are passed to the hidden layer nodes, where the inputs for each hidden node are summed. A mapping function, usually a sigmoid function, is then applied to the weighted sum and the value of the mapping function is sent to the output nodes. The output nodes perform the same calculation as the hidden nodes and produce the value of the dependent variable(s).

The ANN learns patterns using the backpropagation learning algorithm, also termed 'steepest descent'. This algorithm repeatedly runs through the training data, comparing the predicted values and the observed values. The backpropagation learning algorithm has two parameters: the learning rate (η) and the momentum factor (α). The learning rate determines how much the weights are allowed to change each time they are updated. The momentum factor determines how much the current weight change is affected by the previous weight change. The weights of the neural network are adjusted as follows:

$$w_{i,j}(\text{new}) = w_{i,j}(\text{old}) + \eta\delta_i o_j + \alpha[\Delta w_{i,j}(\text{old})] \tag{1}$$

where $w_{i,j}$ is the weight associated with the jth node in the ith layer, η is the momentum factor, α is the learning rate, o_j is the output from the jth output node, and δ_i is the error signal determined by:

$$\delta_i = (t_i - o_i)o_i(1 - o_i) \tag{2}$$

where t_i is the observed value for the ith output node.

The neural network sums the weight adjustments over an epoch and then adjusts the weights. After an initial period of rapid adjustment of the weights, the ANN reaches a stable solution indicating that the model has 'learned' the data structure and may be applied for prognostic analysis of 'new' data.

4. RESULTS

4.1. Synoptic classification

4.1.1. PCA. On the basis of the truncation methods described in the previous section on PCA, seven components were retained for the growing season and five components were retained for the non-growing season. These PCs explain 86% and 83% of the variance of the original rawinsonde derived data sets from the two seasons, respectively. The unrotated solutions were then subjected to Varimax orthogonal and H-K II oblique rotations, which were compared using simple structure plots (an example is given in Figure 3). As shown in Figure 3, both the orthogonal and oblique solutions exhibit stronger simple structure than the unrotated solution. In addition to inferior simple structure, the unrotated solution also exhibited Buell patterns (Buell, 1975).

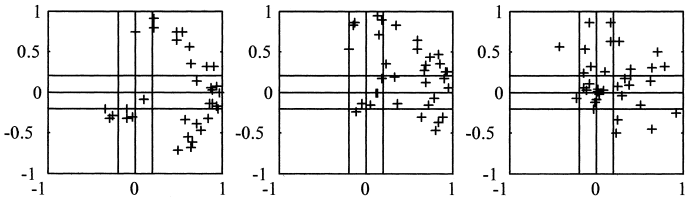

Figure 3. Simple structure plots for the unrotated (left), orthogonally rotated (middle) and obliquely rotated (right) solutions for the first two growing season components

Because the PC scores were to be used in a cluster analysis, it was desirable that the PCs be orthogonal. To ensure that the orthogonal (Varimax) solution attained adequate simple structure, coefficients of congruence (White *et al.*, 1991) between the orthogonal and oblique components were calculated. The results indicate that each of the Varimax orthogonally rotated components is similar to one of the H-K II obliquely rotated components (with the exception of the first and seventh growing season components) (Table II). Based on this analysis and the simple structure plots, it was assumed that the Varimax solution yields acceptable simple structure.

In order to interpret the PC solution, maps of the Varimax rotated component loadings were produced. The plots for the growing season are shown in Figure 4 and are indicative of the relationship between each of the PCs and the original input variables. Based on Richman and Gong (1993), loadings in excess of ± 0.4 are deemed significant and are shaded to aid in the interpretation of the loading maps.

As Figure 4 shows, the first component (referred to as PC1) is related to temperature, with large loadings on the geopotential heights and on 850-hPa temperature. The second component (PC2) is primarily pressure related, with high loadings on sea level pressure. Component 3 is a moisture component, with large loadings on relative humidity. The remaining four components (PCs 4–7) are also related to temperature, pressure and moisture, but in different parts of the domain. While the non-growing season components (figures not shown) exhibit similar patterns, the magnitude of the loadings and mean patterns of the variables differ between the seasons, thus confirming the need for separate analyses by season.

Table II. Coefficients of congruence between Varimax orthogonally rotated components and H-K obliquely rotated components for growing season (top) and non-growing season (bottom)[a]

Growing season	Varimax #1	Varimax #2	Varimax #3	Varimax #4	Varimax #5	Varimax #6	Varimax #7
H-K #1	0.76	0.10	−0.13	0.38	0.06	−0.10	−0.82
H-K #2	0.29	0.89	−0.23	0.25	−0.08	0.54	−0.05
H-K #3	−0.19	−0.08	0.98	−0.25	0.18	−0.09	−0.01
H-K #4	0.34	−0.13	−0.22	0.97	−0.08	0.08	−0.34
H-K #5	−0.14	0.13	0.16	−0.13	0.98	−0.10	−0.11
H-K #6	−0.31	−0.17	−0.01	−0.37	−0.05	0.94	0.10
H-K #7	0.54	0.02	−0.14	0.14	−0.10	0.20	0.56

Non-growing season	Varimax #1	Varimax #2	Varimax #3	Varimax #4	Varimax #5
H-K #1	0.98	−0.01	−0.31	−0.15	−0.04
H-K #2	0.05	0.97	−0.29	−0.20	−0.26
H-K #3	−0.13	−0.15	0.99	0.26	0.21
H-K #4	−0.11	−0.15	0.29	0.98	0.23
H-K #5	−0.03	−0.17	0.18	0.17	0.99

[a] Underlined coefficients of congruence represent a high degree of similarity between the Varimax components and the H-K components.

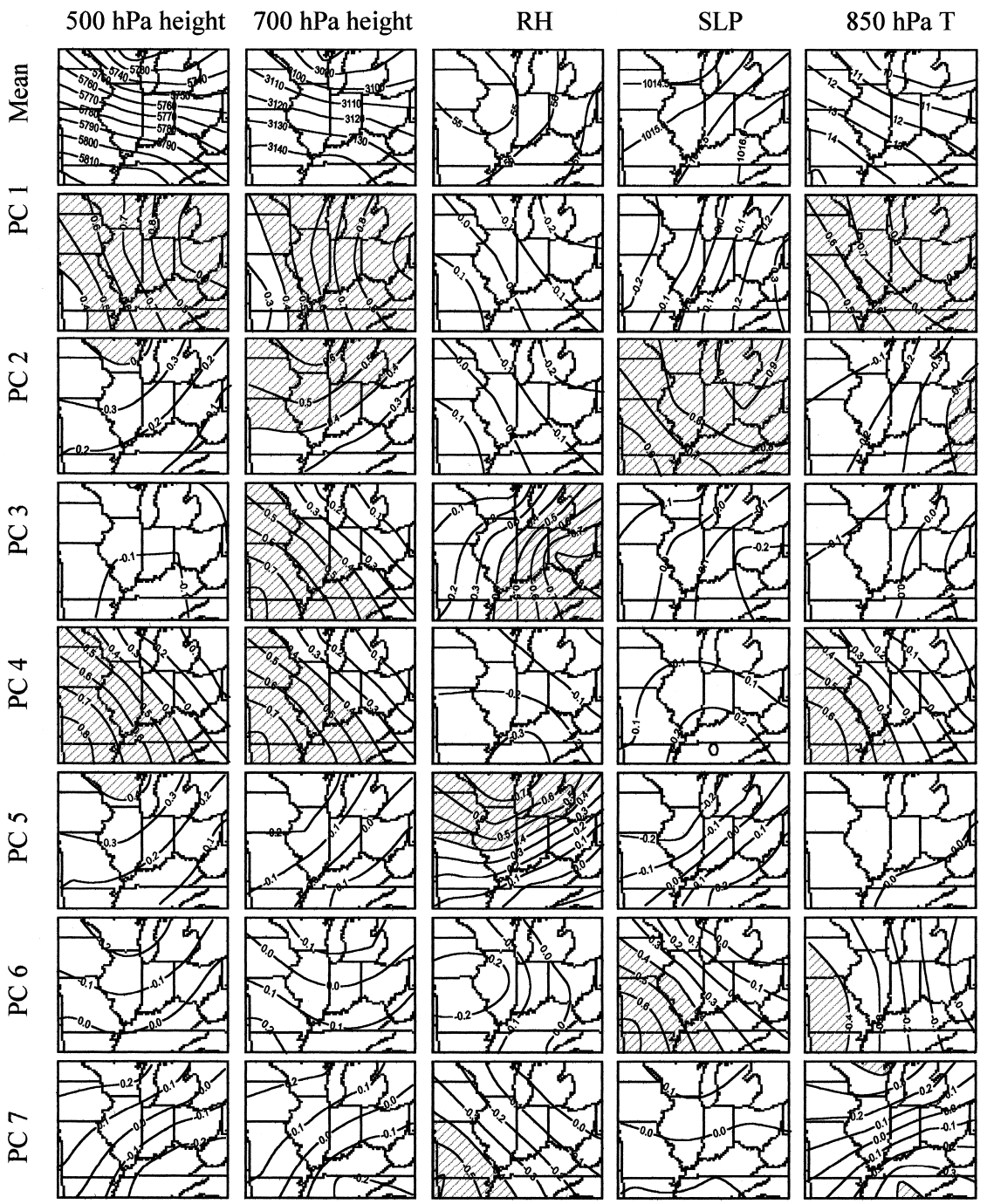

Figure 4. Growing season PC loadings. Values larger than ± 0.4 are shaded to aid in interpretation

4.1.2. Cluster analysis of PC scores. PC scores represent the relationship between each PC and the individual observations. Therefore, clustering the PC scores will yield groups that represent frequently occurring weather patterns.

Of the five hierarchical clustering algorithms applied to the PC scores, in accord with Kalkstein *et al.* (1987), average linkage I yielded the most realistic distribution of clusters (e.g., not biased towards a single dominant class), and thus was chosen as the clustering algorithm for this research.

To determine the correct number of clusters, we used the inflection point on a plot of the distance between clusters versus the number of clusters and the constraint that the clusters represent physically reasonable meteorological scenarios. This method results in eight growing season clusters and twelve non-growing season clusters. Each of the clusters represents a class of synoptic-scale circulation common in the Midwestern United States. The arithmetic means of the input variables in each cluster were then determined and a key day chosen based on proximity to the mean conditions. A sample key day from each growing season cluster is shown in Figure 5. The clusters are described as follows:

Cluster 1: is characterized by a mid-latitude cyclone to the north and west of the study area. The air mass affecting most of the domain is consistent with the moist air within the warm sector of the cyclone. At Indianapolis, cluster 1 is warm and cloudy. There is often precipitation affecting the western part of the domain.

Cluster 2: is associated with a large high-pressure centre to the southeast of the domain and associated dry conditions. Cluster 2 exhibits much lower relative humidity (although more variable in the south) and a much weaker pressure gradient than the previous cluster. Average sea level pressure values over the area are up to 5 hPa higher for cluster 2 than for cluster 1. Cluster 2 is among the driest of all clusters.

Cluster 3: is characterized by a weak cold front oriented W–E across the southern part of the study area and a strong high to the north. Temperatures at the 850-hPa level are highly variable over the domain, with the highest values in the SW and the lowest in the NE. Relative humidity (RH) values exhibit a similar pattern. At Indianapolis, cluster 3 is characterized by frontal activity, usually affecting only the eastern portion of the domain.

Cluster 4: is associated with a weak high over the study area. There is a general NE–SW temperature gradient. Both the SLP and geopotential height fields exhibit weak gradients. At Indianapolis, cluster 4 is associated with warm, dry conditions.

Cluster 5: is characterized by a cold front and associated precipitation in the central portion of the domain and/or along the east coast of the US. This front is generally accompanied by high pressure centred over the central US. Due to the passage of the cold front, 850-hPa temperatures are relatively low for observations in this cluster. In addition, the geopotential height gradients are relatively steep. At Indianapolis, this cluster is associated with above average air temperatures and precipitation.

Cluster 6: is a warm, moist cluster characterized by a stationary front affecting the study area. While the mean 850-hPa temperature varies considerably over the domain (>7°C), gradients of both geopotential height and SLP are weak. There is some precipitation associated with the stationary front. At Indianapolis, this cluster exhibits above average precipitation and temperature.

Cluster 7: is associated with a strong front in the eastern part of the domain. This area also exhibits the highest mean RH for the growing season clusters ($>75\%$). This cluster also has low 850 hPa temperatures, with the entire domain having mean 850-hPa temperature <11°C. Due to the position of the synoptic features, this cluster is generally associated with precipitation in the eastern half of the domain. This cluster exhibits lower than average surface air temperature and high precipitation at Indianapolis.

Cluster 8: is very similar to cluster 7 in terms of frontal locations. However, cluster 8 exhibits higher average SLP, lower average RH and slightly colder 850-hPa temperatures. In addition, for cluster 8 observations, there is often a low-pressure centre to the north of the domain. This low, and the associated frontal activity, are responsible for precipitation in the northwestern part of the domain. At Indianapolis, the observations in cluster 8 are associated with cold, dry conditions.

The mean and standard deviation of several surface variables at Indianapolis are shown by cluster in Table III for the growing season.

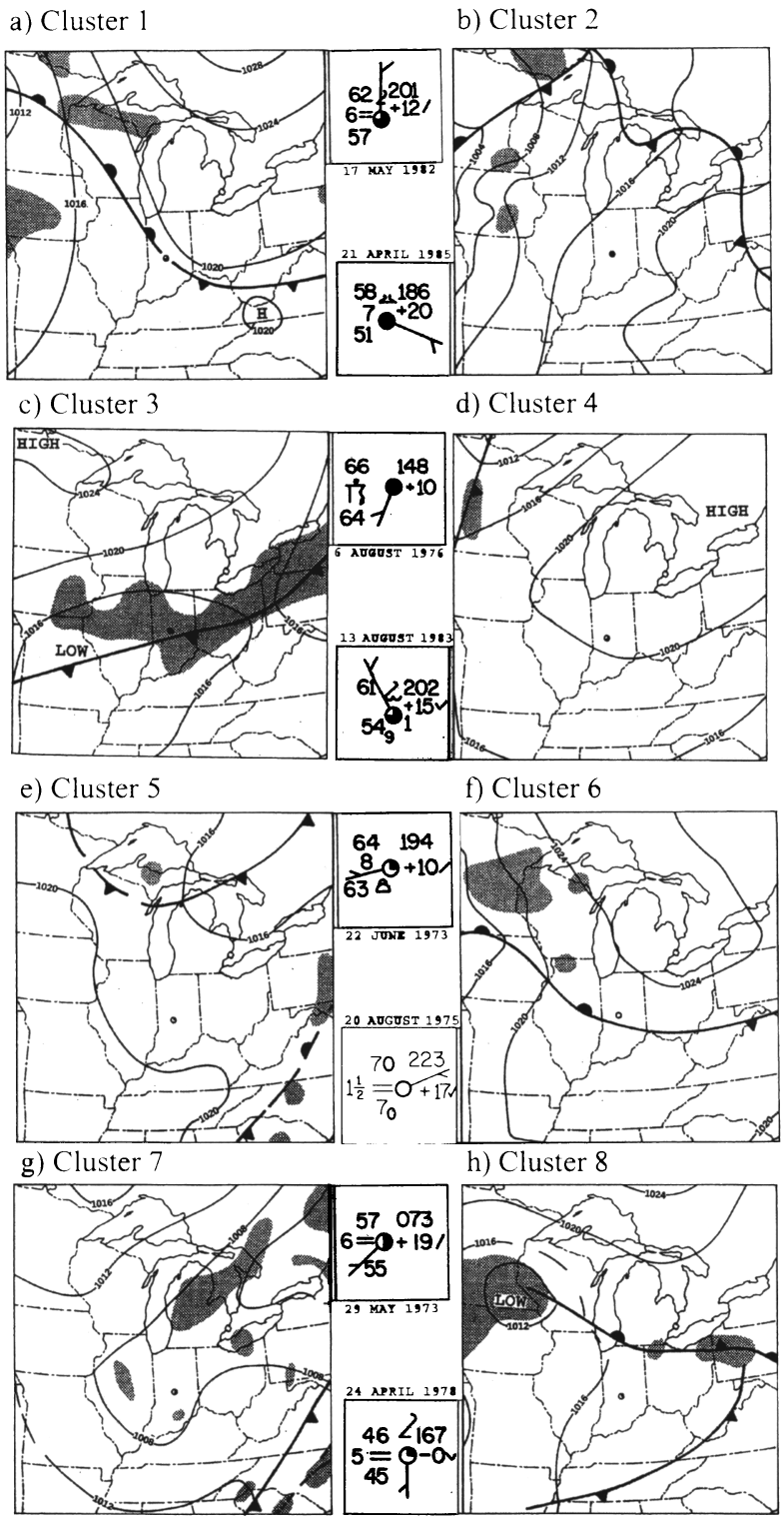

Figure 5. Plot showing cluster 'key days' for the growing season clusters: (a) cluster 1 key day: 17 May 1982; (b) cluster 2 key day: 21 April 1985; (c) cluster 3 key day: 6 August 1976; (d) cluster 4 key day: 13 August 1983; (e) cluster 5 key day: 22 June 1973; (f) cluster 6 key day: 20 August 1975; (g) cluster 7 key day: 29 May 1973; and (h) cluster 8 key day: 24 April 1978. For each key day a station model for conditions at Indianapolis is presented in the central column of the figure

Table III. Mean (μ) and standard deviation (σ) of surface meteorological variables at Indianapolis for the growing season clusters[a]

Cluster #		μ	σ		μ	σ		μ	σ
1	CC	5.88	4.03	T	20.35	6.22	T_d	15.03	6.09
	RH	74.13	18.01	P	987.14	4.57	WS	3.53	1.77
	Vis	13.71	7.98	PC	0.60	5.42			
2	CC	4.46	4.20	T	20.24	6.91	T_d	13.03	6.37
	RH	66.79	20.19	P	989.61	4.46	WS	3.20	1.57
	Vis	15.07	8.09	PC	0.29	2.21			
3	CC	4.01	3.96	T	14.33	7.62	T_d	7.60	6.04
	RH	67.37	19.13	P	992.26	4.31	WS	3.48	1.91
	Vis	21.86	8.20	PC	0.05	0.515			
4	CC	5.13	4.12	T	19.67	6.88	T_d	12.94	6.67
	RH	68.23	19.11	P	986.89	4.19	WS	3.28	1.57
	Vis	17.00	8.64	PC	0.24	2.02			
5	CC	6.68	3.85	T	19.31	6.37	T_d	14.76	6.35
	RH	77.35	17.52	P	985.92	4.1	WS	3.79	1.82
	Vis	15.30	8.93	PC	1.11	7.67			
6	CC	7.93	3.3	T	20.03	5.26	T_d	16.70	5.15
	RH	82.64	13.93	P	985.81	4.71	WS	3.57	1.61
	Vis	11.38	7.40	PC	0.90	5.71			
7	CC	7.95	3.31	T	15.00	6.12	T_d	11.94	6.22
	RH	83.66	14.56	P	985.25	5.42	WS	3.93	1.88
	Vis	11.52	8.27	PC	1.38	9.81			
8	CC	4.50	4.11	T	12.6	7.13	T_d	7.07	6.48
	RH	77.62	20.02	P	987.84	4.17	WS	3.31	1.73
	Vis	19.91	9.01	PC	0.17	1.18			

[a] Abbreviations are as follows: CC—cloud cover (tenths), T—air temperature (°C), T_d—dew point temperature (°C), RH—relative humidity (%), P—barometric pressure (hPa), WS—wind speed (m s^{-1}), Vis—visibility (km), and PC—precipitation (mm).

4.2. Downscaling

In development and testing of all the downscaling models, 75% of the data sets were randomly chosen to train the models and the remaining 25% was used for model evaluation.

4.2.1. Regression-based methods. Regression models for surface air temperature and daily precipitation at Indianapolis were constructed using the synoptic circulation classification indices as predictor variables (Table IV). Each temperature model was constructed with and without an autoregressive term, which represents the strong temporal autocorrelation inherent in daily temperature data. These models were then evaluated based on the accuracy of predictions for a test data set and in terms of the validity of the assumptions of MLR models.

In the following discussion, the following nomenclature is used to identify the models: the downscaled parameter (e.g., T_{max}), the season (e.g., GS to indicate growing season) and AU to indicate if the model were constructed using an autoregressive term (e.g., T_{max}GSAU indicates the model fits maximum temperature during the growing season and includes an autoregressive term).

Table IV shows the values of the regression coefficients, while Table V shows model evaluation statistics for each of the multiple regression models; correlation coefficients, mean absolute error (MAE), and the systematic (bias) and unsystematic (random) components of RMSE (Willmott, 1981). The performance of the models varied, but the models with autoregressive terms typically performed much better than those based solely on PC scores. However, it should be noted that if the models are to be used in a predictive sense (e.g., with GCM output), inclusion of an autoregressive term could result in propagation of errors.

As shown in Figure 4, the first PC from the growing season has high loadings on 850-hPa temperature. Thus, PC scores for this component have the largest regression coefficient for both T_{max}GS and

Table IV. Description of the regression models, input values and model forms[a]

Predictand	Equation	Predictors
T_{max}GS (°C)	$T_{max}GS = 1.86P1 - 0.36P2 + 0.63P_3 + 1.24P4 - 0.25P5 + 0.31P6 - 0.83P7 + 24.95$	00:00 GMT PC scores
T_{max}GSAU (°C)	$T_{max}GSAU = 1.07P1 - 0.41P2 + 0.14P3 + 0.52P4 - 0.30P5 + 0.07P6 - 0.34P7 + 0.76AU + 5.99$	00:00 GMT PC scores and an autoregressive term
T_{min}GS (°C)	$T_{min}GS = 2.79P1 - 0.54P2 + 1.16P3 + 1.18P4 + 0.39P5 - 0.66P6 - 0.47P7 + 13.70$	12:00 GMT PC scores
T_{min}GSAU (°C)	$T_{min}GSAU = 1.17P1 + 0.20P2 - 0.15P_3 + 0.23P_4 + 0.54P5 - 0.78P6 + 0.21P7 + 0.80AU + 2.74$	12:00 GMT PC scores and an autoregessive term
T_{max}NS (°C)	$T_{max}NS = 1.93P1 - 2.37P2 - 0.16P3 + 0.10P4 - 0.53P5 + 7.16$	00:00 GMT PC scores
T_{max}NSAU	$T_{max}NSAU = 1.22P1 - 1.25P2 - 0.22P3 + 0.24P4 - 0.10P5 + 0.69AU + 2.30$	00:00 GMT PC scores and an autoregressive term
T_{min}NS (°C)	$T_{min}NS = 3.80P1 - 3.10P2 - 0.15P3 + 0.36P4 - 0.32P5 - 2.87$	12:00 GMT PC scores
T_{min}NSAU	$T_{min}NSAU = 2.15P1 - 1.12P2 - 0.21P3 - 0.32P4 + 0.11P5 + 0.68AU - 0.94$	12:00 GMT PC scores and an autoregressive term
$PREC_d$GS (mm)	$PREC_dGS = 0.04P_{12}1 - 0.30P_{12}2 + 0.71P_{12}3 - 0.09P_{12}4 + 0.20P_{12}5 - 0.10P_{12}6 - 0.45P_{12}7 - 0.10P_{00}1 - 0.07P_{00}2 + 0.19P_{00}3 + 0.14P_{00}4 - 0.09P_{00}5 + 0.02P_{00}6 + 0.04P_{00}7 + 0.84$	00:00 GMT PC scores ($P_{00}\#$) and 12:00 GMT PC scores ($P_{12}\#$)
$PREC_d$NS (mm)	$PREC_dNS = 0.31P_{12}1 - 0.29P_{12}2 + 0.42P_{12}3 + 0.59P_{12}4 + 0.29P_{12}5 - 0.09P_{00}1 - 0.07P_{00}2 + 0.04P_{00}3 - 0.08P_{00}4 + 0.02P_{00}5 + 0.57$	00:00 GMT PC scores ($P_{00}\#$) and 12:00 GMT PC scores ($P_{12}\#$)

[a] The predictors in the regression equations ($P\#$) indicate the PC scores. GS indicates the growing season and NS indicates the non-growing season. The term AU refers to the autoregressive term.

T_{max}GSAU. However, while T_{max}GS has relatively large coefficients for PC3 and PC4, T_{max}GSAU is strongly dependent on the scores of PC1 and PC4. Because PC3 is a moisture-driven component, it likely contains some information about persistence associated with synoptic-scale systems. Therefore, the inclusion of the autoregressive term in T_{max}GSAU accounts for some of the same variation in growing season daily maximum surface air temperature and the coefficient of PC3 is much smaller.

The largest coefficient for T_{min}GS is also associated with PC1. In addition, PC3 has a relatively large coefficient. Because moisture is often related to the passage of synoptic-scale systems, PC3 is not nearly as important if the autoregressive term is used (T_{min}GSAU). T_{max}NS has large regression coefficients for

Table V. Model evaluation statistics for the regression models[a]

Model	μ		σ		r	MAE	RMSE	$RMSE_s$	$RMSE_u$
	OBS	PRED	OBS	PRED					
T_{max}GS	24.65	24.92	6.03	2.67	0.51	4.15	5.22	4.69	2.30
T_{max}GSAU	24.65	24.88	6.03	5.25	0.85	2.31	3.18	1.58	2.76
T_{min}GS	13.78	13.80	6.49	3.40	0.60	4.35	5.21	4.44	2.72
T_{min}GSAU	13.78	13.85	6.49	5.52	0.88	2.39	3.08	1.63	2.61
T_{max}NS	8.25	7.71	6.67	2.93	0.34	6.68	8.16	7.68	2.75
T_{max}NSAU	8.25	7.46	8.67	6.64	0.81	4.04	5.17	3.39	3.90
T_{min}NS	−3.09	−2.93	9.38	5.08	0.55	4.87	7.84	6.59	4.25
T_{min}NSAU	−3.09	−3.66	9.38	9.59	0.80	3.23	6.09	1.83	5.81
$PREC_d$GS	2.69	0.66	7.19	1.12	0.36	2.97	7.16	7.09	1.04
$PREC_d$NS	2.76	0.49	6.06	0.99	0.50	2.72	6.07	6.01	0.86

[a] The observed values are indicated by OBS, and the model predictions by PRE. The table includes the predicted and observed means (μ) and S.D.s (σ), Pearson product-moment correlation coefficient (r), MAE, RMSE, and the systematic and unsystematic components ($RMSE_s$ and $RMSE_u$). The models are as described in Table IV.

PC1, with a large negative coefficient for PC2. Unlike the growing season temperature models (T_{max}GS, T_{max}GSAU), the addition of the autoregressive term does not reduce the importance of the moisture driven component (PC3).

The coefficients of T_{min}NS exhibit relationships similar to those of T_{max}NS. In particular, the coefficient of PC1 is directly related to daily minimum temperature, while PC2 is inversely related to daily minimum temperature. The role of PC2 is somewhat diminished by the addition of the autoregressive term.

If regression models are to be used in a prognostic sense, it is important that the model assumptions are analysed. In this study, the variance, distribution and serial independence of the residuals were tested using both statistical and graphical methods. These analyses showed that each of the regression models violated at least one of the assumptions tested.

As in previous studies (e.g., Weichert and Bürger, 1998), the performance of the precipitation model based on the PC scores (Table IV) was poorer than that of the models for temperature. The model evaluation statistics (Table V) show that the Poisson regression models do not provide an acceptable level of predictive ability. There are several possible reasons for such poor performance:

(i) The Poisson regression model requires that the predictand be Poisson distributed. In this case, the distribution of daily precipitation is such that the mean and variance are not equal. This suggests that the distribution of daily precipitation deviates from a Poisson distribution.
(ii) It is difficult to predict a series that is dominated by zeros. The growing season precipitation data set is 70% zeros, while that of the non-growing season is 63% zeros.

4.2.2. ANNs. Several ANNs were constructed and evaluated using the results of the synoptic circulation classification (both PC scores and cluster frequencies) (analogous to those in Table IV). For each ANN, experimentation was necessary to determine the optimal values for the learning rate and momentum factor, as well as the correct number of hidden nodes. The final values of these parameters were determined by training and testing the ANNs numerous times and comparing predicted values with observed values. The chosen values of these parameters, as well as the sample sizes for the training and test data sets for each ANN are shown in Table VI. In all cases, larger values of the learning rate resulted in models that predicted the same value of the predictand, regardless of the network input.

Table VII shows the model evaluation statistics associated with each ANN. Because the coefficients (weights) of ANNs are not indicative of the relative importance of the predictors (as is the case with regression), ANNs are considerably more difficult to interpret. However, in terms of the model evaluation statistics the temperature ANNs perform similarly to their MLR counterparts.

According to Cannon and McKendry (1999), neural network outputs tend to be unstable. In order to test whether the models developed here exhibit instability, a resampling technique was employed. For each

Table VI. Learning rate, momentum factor, number of hidden nodes and sample sizes for the training and testing of the ANNs

Model	Learning	Momentum	# Hidden nodes	Train	Test
T_{max}GS (°C)	0.009	0.9	10	1896	632
T_{max}GSAU (°C)	0.01	0.5	12	1896	632
T_{min}GS (°C)	0.012	0.9	10	2455	671
T_{min}GSAU (°C)	0.012	0.8	10	2455	671
$PREC_d$GS (mm)	0.05	0.3	18	1813	604
$PREC_m$GS (mm)	0.3	0.05	12	68	22
T_{max}NS (°C)	0.007	0.9	8	1316	435
T_{max}NSAU (°C)	0.007	0.7	9	1316	435
T_{min}NS(°C)	0.012	0.9	7	1995	611
T_{min}NSAU (°C)	0.012	0.9	8	1995	611
$PREC_d$NS (mm)	0.05	0.3	15	1230	409
$PREC_m$NS (mm)	0.001	0.1	15	68	22

Table VII. Model evaluation statistics for the ANNs. Entries are analogous to those in Table V

Model	μ		σ		r	MAE	RMSE	$RMSE_s$	$RMSE_u$
	OBS	PRED	OBS	PRED					
T_{max}GS	24.77	24.84	6.07	3.12	0.58	3.86	4.98	4.27	2.55
T_{max}GSAU	24.77	24.66	6.07	5.13	0.85	2.40	3.18	1.70	2.68
T_{min}GS	13.69	13.73	6.51	3.59	0.63	4.24	5.08	4.25	2.79
T_{min}GSAU	13.69	13.87	6.51	5.60	0.89	2.27	2.95	1.53	2.53
T_{max}NS	7.63	7.60	8.16	3.10	0.36	6.12	7.61	7.03	2.89
T_{max}NSAU	7.63	7.18	8.16	6.49	0.80	3.81	4.87	2.98	3.85
T_{min}NS	−2.84	−2.73	7.59	5.00	0.64	4.52	5.81	4.37	3.83
T_{min}NSAU	−2.84	−2.82	7.59	6.92	0.87	2.76	3.76	1.57	3.42
$PREC_d$GS	2.84	3.43	8.17	0.05	0.09	4.71	8.19	8.19	0.05
$PREC_d$NS	2.79	2.63	6.01	0.06	0.11	3.80	6.01	6.00	0.06
$PREC_m$GS	80.27	90.98	66.6	40.38	0.65	42.18	52.00	41.89	30.82
$PREC_m$NS	57.18	68.02	28.78	13.28	0.54	22.88	26.60	24.15	11.16

predictand, multiple ANNs were constructed using different training and testing samples. Model evaluation statistics (including correlation coefficient, mean absolute error, and systematic and unsystematic components of the RMSE) from the original ANN were then compared to the distributions of those produced by the resampling method. Results show little variation in terms of the correlation coefficient between predicted and observed values, MAE, or the systematic and random components of the RMSE, indicating that the models derived here are relatively stable.

As a result of the poor performance of the daily precipitation models, monthly precipitation ANNs were also constructed using the cumulative cluster frequencies as predictors, although corresponding regression models were not produced because of the colinearity of the cluster frequencies. Table VII also shows the model evaluation statistics for the daily ($PREC_d$XX) and monthly ($PREC_m$XX) total precipitation neural networks. As shown, the performance of the daily precipitation ANNs failed to match the results achieved by Poisson regression. The poor performance can, again, be partially attributed to the large number of zeros in the time series of the predictand. The weights of the neural network become small to accommodate the large number of zeros, but are unable to recover when the value of the precipitation is large. As shown, the ANN developed with the aggregated monthly data performed better than any of the daily precipitation models.

4.3. Summary of statistical downscaling approaches

The performance of the ANNs for maximum daily surface air temperature is clearly superior to that of MLR for the non-autoregressive models (compare Tables V and VII). Not only is the correlation coefficient between predicted and observed values higher for the ANNs, the ANNs typically (e.g., T_{max}GS) resolved a larger portion of the systematic relationship between predicted and observed values than corresponding regression models. Alternatively stated, the temperature models based on ANNs (with the exception of T_{max}GSAU) show smaller bias than the regression models, and hence, have greater prognostic utility. In addition, the neural networks generally simulated the mean and standard deviation of the predictand better than the regression models. Inclusion of the autoregressive term generally resulted in neural networks and regression models that were comparable in terms of model evaluation statistics.

Also shown in Tables V and VII, the non-growing season daily maximum surface air temperature models did not perform as well as their growing season counterparts. The poorest performance of the downscaling models was associated with the non-autoregressive multiple linear regression model for non-growing season T_{max}. This model had a RMSE greater than 7°C. Performance of neural networks and regression models was very similar for non-growing season T_{max}. However, the regression models were more successful than neural networks in terms of simulating the mean of the T_{max}. Both methods failed to capture the variance.

The results of the downscaling analysis suggest that the PC scores also contain information about T_{min} during the two seasons. The two regression models for T_{min}GS performed similarly to the T_{max}GS models. The neural networks for T_{min}GS performed slightly better than the regression models. The correlation coefficients between predicted and observed precipitation in the test data sets for the neural networks were slightly higher than those for the regression models. In addition, the ANNs exhibited smaller RMSE and MAE, as well as a smaller systematic error component.

Both downscaling methods failed to produce an acceptable model for daily precipitation. In both seasons, the downscaling failed to predict precipitation events of large magnitude. In order to predict the large number of zero values, the Poisson regression models have small coefficients for all PC scores. Similarly, the neural networks failed to determine weights that would allow the prediction of such a large number of rain-free observations and the prediction of large precipitation events. Because these models always predicted small precipitation values, the regression errors are highly correlated with the observed precipitation value. This is also evident in the large systematic errors associated with these models. Cannon and Lord (2000) employ a histogram equalization process to aid in the prediction of extreme ozone events. It is conceivable that such a technique might provide some improvement for the precipitation models in this study.

Correlation coefficients for the non-growing season precipitation models were substantially higher than those for the growing season for both downscaling techniques. In addition, RMSE were lower for the $PREC_d$NS models. However, the errors in these models were highly correlated with the value of the predictand. Because the growing season had larger observed precipitation values than the non-growing season, the RMSE and MAE for the growing season are larger than those for the non-growing season.

The performance of the total monthly precipitation ANNs models was vastly superior to the total daily precipitation ANNs. Most months recorded some precipitation so the monthly precipitation models were not biased by a large number of observations with zero precipitation. The ANNs for monthly total precipitation for the growing and non-growing seasons had a RMSE of 52.0 and 26.6 mm, respectively, relative to a seasonally averaged monthly precipitation of 84 and 60 mm. The most accurate predictions are for months with moderate amounts of precipitation. This suggests that the ANNs have learned the patterns that are common in the data, but have failed to locate patterns associated with values of observed precipitation that deviate from the mean monthly precipitation. This may reflect the relatively small sample training data set ($n = 68$) and infrequent occurrence of high precipitation months. As described above, there are several possible reasons for the poor performance of the precipitation models, including violation of Poisson regression assumptions. However, the performance might also be the result of the synoptic classification that produced the predictors. It is possible that, in addition to the input variables, information is needed about atmospheric stability (e.g., CAPE, vertical vorticity), or a relative measure of humidity, such as 850-hPa dew point depression.

4.4. Comparison with previous downscaling studies

While this is the first study to compare regression-based methods and ANNs for maximum and minimum air temperature and precipitation prediction in the Midwest United States, there have been several studies using comparable methods in other domains. Therefore, while the results are not directly comparable, it is worthwhile to relate the performance of the models presented here with those presented in other studies that closely equate to this study.

Easterling (1999) used free atmospheric variables (850 and 500-hPa heights, 500-hPa temperature, 500-hPa relative humidity, and 500-hPa u and v wind components), along with PCA and regression analysis, to predict the maximum and minimum monthly surface air temperature and monthly precipitation at 32 stations in the central United States. The resulting transfer functions (for all stations and months) produced similar results to those of this study. The correlation coefficients for predicted and observed values ranged from 0.8 to 0.9 for monthly maximum air temperature, 0.7 to

0.8 for monthly minimum air temperature and 0.4 to 0.7 for monthly precipitation, while the comparable results for this study are 0.34–0.85 for daily maximum temperature, 0.63–0.89 for daily minimum temperature and 0.54–0.65 for monthly precipitation. As in this study, the regression models for monthly precipitation in Easterling's study performed better during the winter months.

Kidson and Thompson (1998) compared regression-based methods and a regional nested model for the downscaling of maximum and minimum daily and monthly surface air temperature and total precipitation in New Zealand. They found little difference in the explanatory skill of the two methods, suggesting the prediction of these variables is not limited by the methods, but rather by the relationships between the variables and the circulation. Although the results varied spatially due to the extreme relief in the study area, results were also similar to those attained in this study. The transfer functions yielded correlation coefficients ranging from 0.5 to 0.7 for daily maximum air temperature, 0.3 to 0.6 for daily minimum air temperature, and 0.1 to 0.5 for both daily and monthly precipitation totals.

Cavasos (1997) used neural networks to downscale daily precipitation for 20 stations in northeastern Mexico for eight winters (1985–1993) using sea level pressure, 500-hPa height and the 1000–500-hPa thickness. Unlike this study, Cavasos (1997) used time-lagged PC scores as predictors. In addition, the inclusion of 20 stations allowed Cavasos to analyse the spatial variation in model results as they relate to physiographic features of the study area. Although the derived correlation coefficients between predicted and observed values were as high as 0.7, the author also reported difficulty in capturing the magnitude of extreme events, such as El Niño.

Weichert and Bürger (1998) used 500 hPa-height and 850-hPa temperature fields to simulate surface temperature, vapour pressure, and precipitation for a station in Germany. The study compared linear (canonical correlation analysis) and non-linear techniques (ANNs). The model performance was comparable to that achieved in this study. However, while the overall performance of the non-linear technique was not superior to that of linear downscaling, Weichert and Bürger (1998) found clear differences between the two methods. In particular, they found that the ANNs accounted for some heavy rainfall events, while they were not identified by the linear technique. Such model behaviour was not observed in this study.

5. SUMMARY

Synoptic indices for the Midwestern United States were developed using PCA of 500 and 700-hPa height, 850-hPa temperature, sea level pressure, plus column average relative humidity from seven rawinsonde sites. Clustering of the resulting PC scores yielded eight and twelve synoptic classes for the growing and non-growing seasons, respectively. The circulation indices and cluster frequencies were then used to downscale maximum and minimum daily air temperature and daily and monthly precipitation totals at Indianapolis using regression techniques and ANNs. While the temperature models devised here perform relatively well, the precipitation models fail to capture the variability of precipitation as governed by the synoptic scale circulation. This suggests that other variables are necessary to capture precipitation events, especially on short time scales.

The ability to relate the synoptic scale circulation to surface climate in this study is limited by the density of the rawinsonde station network. Thus, future work will include the use of NCEP/NCAR gridded meteorological fields of temperature and moisture, which are derived from physically constrained interpolation of the rawinsonde network.

In the longer term, the synoptic classification methodology developed here will be applied to the results of GCM simulations of present climate to evaluate model performance. Contingent on this analysis, an analysis of $2 \times$ carbon dioxide (CO_2) simulations will be performed to assess the effect of global scale climate change on the synoptic scale conditions over the Midwestern United States, and transfer functions (downscaling models) such as those devised here will be applied to the $2 \times CO_2$ simulations to provide forecast climate scenarios for central Indiana.

ACKNOWLEDGEMENTS

50 percent of this research was funded by the National Institute for Global Environmental Change through the US Department of Energy (Cooperative Agreement No. DE-FC03-90ER61010). Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the DOE. Thanks to Mike Hollingsworth for the drafting of several figures. The clarity of this paper was substantially improved by the constructive criticisms of two anonymous reviewers.

REFERENCES

Ashbaugh LL, Myrup LO, Flocchini RG. 1984. A principal components analysis of sulfur concentrations in the Western United States. *Atmospheric Environment* **18**: 783–791.

Buell CE. 1975. The topography of empirical orthogonal functions. In [preprints, Fourth Conference on Probability and Statistics in Atmospheric Science]. Tallahassee, FL, American Meteorological Society; 188–193.

Cannon AJ, Lord ER. 2000. Forecasting summertime surface-level ozone concentrations in the Lower Fraser Valley of British Columbia: An ensemble neural network approach. *Journal of the Air and Waste Management Association* **50**: 322–339.

Cannon AJ, McKendry IG. 1999. Forecasting all India summer monsoon rainfall using regional circulation principal components: A comparison between neural network and multiple regression models. *International Journal of Climatology* **19**: 1561–1578.

Carlson RE. 1990. Heat stress, plant-available soil moisture, and corn yield in Iowa: A short- and long-term view. *Journal of Production Agriculture* **3**: 293–297.

Cavasos T. 1997. Downscaling large-scale circulation to local winter rainfall in North-eastern Mexico. *International Journal of Climatology* **17**: 1069–1082.

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.

Easterling DR. 1999. Development of regional climate scenarios using a downscaling approach. *Climatic Change* **41**: 615–634.

Eberhart R, Dobbins B. 1990. *Neural Network PC Tools: A Practical Guide*. Academic Press: San Diego, CA.

FSL/NCDC. 1997. Radiosonde data of North America, 1946–1996. Available from http://www.ncdc.noaa.gov/.

Gardner MW, Dorling SR. 1998. Artificial neural networks (the multi layer perceptron)—A review of applications in the atmospheric sciences. *Atmospheric Environment* **32**: 2627–2636.

Giorgi F, Marinucci MR, Visconti G. 1998. Use of a limited-area model nested in a general circulation model for regional climate simulation over Europe. *Journal of Geophysical Research* **95**: 18413–18431.

Gong X, Richman MB. 1995. On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. *Journal of Climate* **8**: 897–931.

Hewitson BC, Crane RG. 1992. Large-scale atmospheric controls on local precipitation in tropical Mexico. *Geophysical Research Letters* **19**: 1835–1838.

Heyen H, Zorita E, von Storch H. 1996. Statistical downscaling of monthly mean North Atlantic air-pressure to sea level anomalies in the Baltic Sea. *Tellus* **48A**: 312–323.

Hornik K, Stinchcombe H, White H. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* **2**: 359–366.

Hughes JP, Guttorp P. 1994. A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water Resources Research* **30**: 1535–1546.

Huth R. 1999. Statistical downscaling in central Europe: Evaluation of methods and potential predictors. *Climate Research* **13**: 91–101.

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.

Karl TR, Wang WC, Schlesinger ME, Knight RW, Portman D. 1990. A method of relating general circulation model simulated climate to observed local climate. Part I: Seasonal statistics. *Journal of Climate* **3**: 1053–1079.

Karl JK, Knight RW, Easterling DR, Quayle RO. 1996. Indicies of climate change for the United States. *Bulletin of the American Meteorological Society* **77**: 279–292.

Kidson JW, Thompson CS. 1998. A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate* **11**: 735–753.

Kilsby CG, Cowpertwait PSP, O'Connell PE, Jones PD. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology* **18**: 523–539.

McGregor GR. 1996. Identification of air quality affinity areas in Birmingham, UK. *Applied Geography* **16**: 109–122.

Mendelsohn R, Nordhaus WD, Shaw D. 1994. The impact of global warming on agriculture: A Ricardian analysis. *The American Economic Review* **84**: 753–771.

NCDC/NREL. 1993. Solar and meteorological surface observation network (SAMSON). Available from http://www.ncdc.noaa.gov/. Data sets were used from CD ROM.

Orlanski I. 1975. A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society* **56**: 527–530.

Overland JE, Preisendorfer RW. 1982. A significance test for principal components applied to a cyclone climatology. *Monthly Weather Review* **110**: 1–4.

Pryor SC, Davies TD, Hoffer TE, Richman MB. 1995. The influence of synoptic scale meteorology on transport of urban air to remote locations in the southwestern United States of America. *Atmospheric Environment* **29**: 1609–1618.

Richman MB. 1986. Rotation of principal components. *Journal of Climatology* **6**: 293–335.
Richman MB, Gong X-F. 1993. Relationships between the definition of the hyperplane width and the fidelity of the resulting PC loading patterns. In *Proceedings of the Seventeenth Annual Climate Diagnostics Workshop*. US Department of Commerce, Washington, D.C.; 372–379.
Rounsevell MD, Evans SP, Bullock P. 1999. Climate change and agricultural soils: Impacts and adaptation. *Climatic Change* **43**: 683–709.
Smit B, McNabb D, Smithers J. 1996. Agricultural adaptation to climatic variation. *Climatic Change* **33**: 7–29.
Solman SA, Nuñez MN. 1999. Local estimates of global climate change: A statistical downscaling approach. *International Journal of Climatology* **19**: 835–861.
Stratheropoulos M, Vassiliadis N, Pappa A. 1998. Principal component and canonical correlation analysis for examining air pollution and meteorological data. *Atmospheric Environment* **32**: 1087–1095.
Ward JH. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**: 236–244.
Warner B, Misra M. 1996. Understanding neural networks as statistical tools. *The American Statistician* **50**: 284–293.
Weichert A, Bürger G. 1998. Linear versus nonlinear techniques in downscaling. *Climate Research* **10**: 83–93.
White DM, Richman MB, Yarnal B. 1991. Climate regionalization and rotation of principal components. *International Journal of Climatology* **11**: 1–25.
Wigley TM, Jones PD, Briffa KR, Smith G. 1990. Obtaining sub-grid scale information from coarse resolution general circulation model output. *Journal of Geophysical Research* **95**: 1943–1953.
Wilks DS. 1995. *Statistical Methods in the Atmospheric Sciences*. Academic Press: San Diego, CA.
Willmott CJ. 1981. On the validation of models. *Physical Geography* **2**: 184–194.
Winkler JA, Palutikof JP, Andresen JA, Goodess CM. 1997. The simulation of daily temperature time series from GCM output. Part II: sensitivity analysis of an empirical transfer function methodology. *Journal of Climate* **10**: 2514–2532.
Xopalski E, Luterbacher J, Burkard R, Patrikas I, Maheras P. 2000. Connection between the large-scale 500 hPa geopotential height fields and precipitation over Greece during wintertime. *Climate Research* **14**: 129–146.

