GENERATION OF REGIONAL CLIMATE CHANGE SCENARIOS USING GENERAL CIRCULATION MODELS AND EMPIRICAL DOWNSCALING

Justin T. Schoof

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Geography,
Indiana University

September 2004

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Sara C. Pryor, Ph.D.

Doctoral Committee

_____
Scott M. Robeson, Ph.D.

_____
Rebecca J. Barthelmie, Ph.D.

_____
Oral Examination September 2004 David F. Parkhurst, Ph.D.

ACKNOWLEDGEMENTS

The research presented in this dissertation would not have been possible without the support of many people. I would like to acknowledge my Ph.D. supervisor, Dr. Sara Pryor, for her expertise, professionalism, and generous support throughout my years as a graduate student at Indiana University. I would also like to thank the other members of my research committee – Dr. Rebecca Barthelmie, Dr. Scott Robeson, and Dr. David Parkhurst – for their support and suggestions. Additional thanks to the faculty, graduate students, and staff in the Department of Geography for their support and camaraderie.

I would also like to acknowledge my family and friends for being supportive throughout the educational process. My mother, Lynn, has provided unconditional love and support, and displayed unwavering faith in my abilities, while my son, Kevin, has provided the inspiration for all of my past, present, and future achievements. I would also like to thank my girlfriend, Wendy, for her consistent support and encouragement.

ABSTRACT

Justin T. Schoof

Generation of Regional Climate Change Scenarios Using General Circulation Models and Empirical Downscaling

Coupled ocean-atmosphere general circulation models (GCMs) are the best tools available for examination of climate change due to increases in atmospheric greenhouse gas concentrations. Due to large computational requirements, these numerical models run at horizontal resolutions that are inadequate for climate impact studies and, hence, require parameterization of many small-scale processes important for characterization of regional climate. The aim of this research was to develop and evaluate a methodology for generating regional climate change scenarios for the Midwest region of the USA using GCM simulations and empirical downscaling methods. The research focuses on (1) identification of relationships between large-scale predictors and three surface parameters (local maximum and minimum daily surface air temperature and total daily precipitation) at 84 stations in the study region, (2) evaluation of variables simulated by two GCMs, and (3) development and evaluation of empirical downscaling tools to generate projections of the surface parameters for the 21st century.

The results of the analysis indicate that the large-scale atmospheric predictors explain a large proportion of the variability in the surface parameters, but that GCM simulations of the large-scale predictors do not exhibit an acceptable level of agreement with observations at the grid point level. Therefore, the downscaling models applied in

this study are based on (1) relationships between GCM simulated variables and the surface parameters and (2) spatially aggregated predictor information.

The downscaled climate change scenarios indicate strong warming at most stations consistent with projected increases in greenhouse gases. Averaged over all stations, the downscaled results indicate year-round warming, but the magnitude of the 21st century temperature change is inconsistent between results downscaled from the two GCMs used. These results show that, under the emissions scenarios used by the GCMs, important climate change impacts such as increases in heat wave frequency may be realized, although there is a high degree of uncertainty associated with these findings. The downscaled precipitation scenarios are less consistent than those for temperature (in terms of both the direction and magnitude of precipitation change and its spatial coherence), resulting in lower confidence for the precipitation scenarios relative to those for temperature.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION AND OBJECTIVES....1

1.1 BACKGROUND....1
1.2 RESEARCH OBJECTIVES....5
1.2.1 To evaluate the homogenized, gridded data sets (i.e., reanalysis products) commonly used in empirical downscaling....6
1.2.2 To use the historical climate record to identify physically meaningful large-scale variables accounting for a large proportion of the variance of surface climate parameters....7
1.2.3 To evaluate multiple GCM simulations and identify well-simulated variables for use in generating regional climate prognoses....8
1.2.4 To apply, compare, and evaluate multiple empirical downscaling methods....8

CHAPTER 2: STUDY REGION AND DATA....10

2.1 STUDY REGION....10
2.1.1 Current climate of the Midwestern United States....10
2.1.2 Recent climate change in the Midwestern United States....13
2.2 DATA....14
2.2.1 Reanalysis data....14
2.2.2 Radiosonde data....15
2.2.3 Surface station data....15
2.2.4 General circulation model (GCM) data....18

CHAPTER 3: RESEARCH METHODS....22

3.1 DEFINITIONS OF DERIVED VARIABLES....22
3.1.1 Teleconnection indices....22
3.1.1.1 The North Atlantic Oscillation (NAO)....23
3.1.1.2 The Pacific/North American (PNA) pattern....24
3.1.2 Geostrophic flow and vorticity component definitions....25
3.2 SYNOPTIC CLASSIFICATION METHODS....27
3.2.1 Correlation based-map typing....28
3.2.2 Eigenvector/clustering-based synoptic classification....29
3.2.2.1 Principal components analysis (PCA)....30
3.2.2.2 Cluster analysis....31
3.2.3 Comparison of synoptic classifications....33
3.3 ANALYSIS AND COMPARISON OF PROBABILITY DISTRIBUTIONS....34
3.4 BOOTSTRAP RESAMPLING....34
3.4.1 Evaluation of GCM-simulated grid point data....35
3.4.2 Evaluation of GCM-simulated weather types....36
3.5 'TAYLOR' DIAGRAMS....37
3.6 THE TWO-SAMPLE DIFFERENCE OF PROPORTIONS TEST....39
3.7 EMPIRICAL DOWNSCALING METHODS....40
3.7.1 The analog method....40
3.7.2 Stochastic weather generator....41
3.7.2.1 Precipitation occurrence....41
3.7.2.2 Precipitation amount....42
3.6.2.3 Simulation of other variables....43
3.7.3 MOS-based downscaling with multiple linear regression....44

CHAPTER 4: QUANTITATIVE EVALUATION OF 'OBSERVED' (REANALYSIS) DATA USED IN THE EVALUATION OF GCMS....45

1 INTRODUCTION....46
2 DATA / STUDY AREA....48
3 METHODS....52
3.1 Development of the synoptic classifications....52

3.2 Projection of RSONDE classification *55*
4 RESULTS 55
4.1 Analysis of the independent classifications *55*
4.1.1 Modes of variability as manifest in the results of principal components analysis 55
4.1.2 Synoptic types resulting from the cluster analysis 58
4.2 Analysis of the targeted classification *64*
5 DISCUSSION 70
6 CONCLUSIONS 73

CHAPTER 5: IDENTIFICATION OF DOWNSCALING PREDICTOR VARIABLES 78

5.1 INTRODUCTION 78
5.2 PREVIOUS STUDIES 78
5.3 STRATEGY FOR SELECTION OF DOWNSCALING PREDICTORS 83
5.4 SCALES OF VARIABILITY 84
5.5 RESULTS 87
5.5.1 Upper-air variables *87*
5.5.1.1 500-hPa geopotential height (Z_{500}) 87
5.5.1.2 850-hPa geopotential height (Z_{850}) 90
5.5.1.3 850-500 hPa layer thickness (THICK) 93
5.5.1.4 850-hPa specific humidity (Q_{850}) 96
5.5.1.5 850-hPa relative humidity (RH_{850}) 99
5.5.2 Surface variables *102*
5.5.2.1 Sea-level pressure (SLP) 102
5.5.2.2 Zonal component of geostrophic flow (GEOW) 105
5.5.2.3 Meridional component of geostrophic flow (GEOS) 108
5.5.2.4 Resultant geostrophic flow (GEOWS) 111
5.5.2.5 Westerly geostrophic vorticity (GEOZW) 114
5.5.2.6 Southerly geostrophic vorticity (GEOZS) 117
5.5.2.7 Total geostrophic vorticity (GEOZT) 120
5.6 SUMMARY 123

CHAPTER 6: EVALUATION OF GCM SIMULATIONS 130

6.1 INTRODUCTION 130
6.2 EVALUATION OF GCM SIMULATION MEAN FIELDS (1990-2001) 131
6.3 EVALUATION OF GCM GRID POINT DATA 135
6.3.1 GCM evaluation at annual timescales *137*
6.3.2 GCM evaluation at seasonal timescales *140*
6.3.3 GCM evaluation at monthly timescales *146*
6.4 SUMMARY 149
6.5 DISCUSSION 153

CHAPTER 7: EVALUATION OF GCM SIMULATIONS OF SYNOPTIC PHENOMENA 155

1 INTRODUCTION 156
1.1 Background *156*
1.2 Objectives *158*
1.3 Synoptic climate of the Midwestern USA and links to hemispheric teleconnection indices *159*
2 DATA 162
2.1 NCEP/NCAR Reanalysis *162*
2.2 GCM Simulations *163*
3 METHODOLOGY 164
3.1 TELECONNECTION INDICES 164
3.1.1 The North Atlantic Oscillation (NAO) index 164
3.1.2 The Pacific North American (PNA) index 166
3.1.3 Evaluation of large-scale teleconnection indices 166
3.2 Synoptic-scale map-pattern analysis *167*
3.2.1 Correlation-based map typing 168
3.2.2 Evaluation of synoptic-scale map-pattern classification 169
3.3 Relationships between teleconnections and Midwest US weather types *170*

4 RESULTS 172
4.1 Teleconnection indices *172*
4.1.1 North Atlantic Oscillation (NAO) 172
4.1.2 Pacific/North American (PNA) index 176
4.1.3 Summary 178
4.2 Synoptic-scale map-pattern analysis *179*
4.2.1 Observed map-patterns: 1990-2001 179
4.2.2 Comparison of map type frequencies in NNR and transient GCM simulations 183
4.2.3 Linking synoptic scale circulation and the teleconnection indices 187
4.2.4 Summary 189
5 CONCLUDING REMARKS 190

CHAPTER 8: EMPIRICAL DOWNSCALING OF GCM OUTPUT 198

1 INTRODUCTION 198
2 STUDY REGION 202
3 DATA 205
3.1 Reanalysis data *205*
3.2 Surface station data *205*
3.3 General circulation model (GCM) data *208*
4 METHODOLOGY 210
4.1 Identification of downscaling predictor variables *210*
4.2 Evaluation of grid point GCM simulations *212*
4.3 Downscaling methods *217*
4.3.1 MOS-based downscaling of the seasonal cycle 218
4.3.2 The analog method for downscaling daily variability 220
4.3.3 Stochastic weather generator for downscaling daily variability 221
4.3.3.1 Weather classification 222
4.3.3.2 Precipitation occurrence 222
4.3.3.3 Precipitation amount 223
4.3.3.4 Simulation of other variables 223
5 RESULTS 224
5.1 MOS-based downscaling of the T_{max} and T_{min} seasonal cycles *224*
5.1.1 Projected changes in daily T_{max} and T_{min}: 2040-2049 225
5.1.2 Projected changes in the daily T_{max} and T_{min}: 2080-2089 228
5.1.3 Summary 229
5.2 Daily sequences of T_{max}, T_{min}, and precipitation *230*
5.2.1 The analog method 231
5.2.1.1 Temperature scenarios 231
5.2.1.2 Precipitation scenarios 234
5.2.2.1 Temperature scenarios 236
5.2.2.2 Precipitation scenarios 237
6 SUMMARY AND CONCLUSIONS 239

CHAPTER 9: SUMMARY AND CONCLUSIONS 247

9.1 SUMMARY 247
9.1.1 To evaluate the homogenized, gridded data sets (i.e., reanalysis products) commonly using in empirical downscaling) *247*
9.1.2 To use the historical climate record to identify physically meaningful large-scale variables accounting for a large proportion of the variance of surface climate parameters. *248*
9.1.3 To evaluate multiple GCM simulations and identify well-simulated variables for use in generating regional climate prognoses. *250*
9.1.4 To apply, compare, and evaluate multiple empirical downscaling methods. *252*
9.2 DISCUSSION 254
9.3 FUTURE WORK 254

REFERENCES CITED 256

APPENDIX A 269

1 INTRODUCTION 270
2 DATA 271
3 DESCRIPTION OF WEATHER GENERATOR 272
3.1 Precipitation Occurrence Component *273*
3.2 Temperature and Radiation Component *273*
4 OBSERVED RELATIONSHIPS 275
4.1 Seasonal/Spatial Variability of Lag-0 Correlation Coefficients (M_0) *276*
4.2 Seasonal/Spatial Variability of Lag-1 Correlation Coefficients (M_1) *278*
4.3 Seasonal/Spatial Variability of A and B *280*
5 WEATHER GENERATOR IMPLEMENTATION 282
6 EVALUATION OF GENERATED DATA 282
6.1 Means and Standard Deviations *282*
6.2 Correlations Between Generated Variables *283*
6.3 Diurnal Temperature Range (DTR) *286*
7 DISCUSSION AND CONCLUSIONS 288

CHAPTER 1: INTRODUCTION AND OBJECTIVES

1.1 Background

Coupled atmospheric-oceanic general circulation models (AOGCMs, or GCMs) simulate the temporal evolution of the climate system in three spatial dimensions. Due to large computational requirements, these numerical models run at typical horizontal resolutions of 200 to 500 km, resulting in the parameterization of many processes of importance for regional climate (e.g., mesoscale circulation, cloud processes, surface-atmosphere interaction, etc.) and exclusion of others. Thus, while GCM simulations are considered robust at large (e.g., continental, seasonal) scales (McAveney et al. 2001), a relatively high degree of uncertainty still characterizes prognoses of climate at the regional scale, especially for hydrologic variables (e.g. Grotch and MacCracken 1991; Solman and Nunez 1999). The techniques available for deriving regional-scale climate scenarios from GCMs can be divided into three classes (Giorgi et al. 2001):

1) high and variable resolution atmospheric GCMs
2) regional climate models
3) statistical-dynamical and empirical-statistical methods

The first technique uses output from a coupled GCM as boundary conditions for high and variable resolution (typically) atmosphere-only GCM (AGCM) over a specific region for periods up to a few decades (e.g., Cubasch et al. 1995; Cubasch et al. 1996; Deque et al. 1998), and hence is very computationally demanding. Additionally, AGCM formulations are optimized for a specific resolution and some processes may therefore be poorly represented with increased resolution (Giorgi et al. 2001).

Regional climate modeling, which is often referred to as dynamical downscaling, uses initial conditions and lateral meteorological conditions from GCMs, along with surface boundary conditions to drive high resolution (~10 to 50 km grid spacing) regional climate models (RCMs). Dynamical downscaling was pioneered by Dickinson et al. (1989) and has since been applied in many regions including Europe (Jones et al. 1995; Hanssen-Bauer et al. 2003), North America (Giorgi et al. 1994; Leung and Ghan 1999), Australia (McGregor and Walsh 1994), Africa (Semazzi et al. 1994; Arnell et al. 2003), and Asia (Hirakuchi and Giorgi 1995). Like the variable and high resolution GCM approach, regional climate modeling is computationally expensive. Also, the need for regular boundary condition updates (approximately every 6 hours) requires GCM output at high temporal resolution (Giorgi et al. 2001). In contrast to the high and variable resolution approach, the regional climate modeling approach employs GCM output at its original spatial resolution. A variation of the dynamic downscaling approach is hybrid statistical-dynamical downscaling, which uses the output of RCMs or mesoscale atmospheric models to drive statistical transfer functions between the model output and local or regional climate parameters. For example, Fuentes and Heimann (2000) derived 22 synoptic weather classes and ran an RCM once for each class in an application to precipitation downscaling in the Alpine region.

Empirical-statistical downscaling (hereafter referred to as empirical downscaling) uses statistical relationships between large-scale atmospheric fields or grid point values and surface climate variables from historical data, and hence, unlike the other downscaling techniques, can be used to derive either station-scale or spatially aggregated information from multi-decadal GCM simulations. Because empirical downscaling relies

only on observed relationships and not on the physical systems and processes that govern climate (i.e., they are not based on fundamental equations, but rather on statistical associations), caution must be used in applying them under changed climate conditions. In spite of this caveat, empirical downscaling has been widely applied (see reviews by Hewitson and Crane 1996; Wilby and Wigley 1997) and although there have not been any studies comparing empirical and dynamical methods for the study region considered here, studies comparing techniques in other regions have indicated that the empirical downscaling approach provides similar skill for both thermal and hydrological parameters (Kidson and Thompson 1998, Murphy 1999). This dissertation therefore focuses on the development of regional climate change scenarios using empirical downscaling methods.

The empirical downscaling method has its roots in the numerical-statistical daily forecast methods of the National Weather Service (e.g., perfect prognosis (PP), Klein et al. 1959; model output statistic (MOS), Glahn and Lowry 1972), which produce weather forecasts by establishing statistical relationships between free atmosphere variables and surface variables. While the PP approach uses observed free atmosphere predictors, the MOS approach employs model output variables and can therefore account for the internal bias and inaccuracy of the large-scale fields (Klein 1982, Sailor and Li 1999). Most empirical downscaling studies have employed PP approaches, although many have failed to adequately demonstrate that the GCM accurately simulates the predictor variables. Regardless of which approach is adopted, there are several requirements for successful application of empirical downscaling:

1) The free atmosphere predictors must be relevant to the local climate variable of interest. Typically, surface air temperature and/or precipitation are modeled from atmospheric circulation variables (e.g., mean sea-level pressure or 500-hPa geopotential height), either at a nearby grid points (e.g., Sailor and Li 1999) or in weather classification schemes (e.g., Schoof and Pryor 2001). Some studies have criticized the blind use of the nearest grid point and advocated consideration of remote grid points to maximize the proportion of predictand variance explained by the predictor(s) (e.g., Brinkmann 2002). However, most studies avoid quantitative predictor selection and choose predictor variables on a theoretical basis. It should further be noted that the variables archived from GCM experiments place an additional limitation on the choice of downscaling predictors.
2) The free atmosphere predictor variables must be adequately simulated by the GCM. While this requirement is less important in studies which employ MOS-based empirical downscaling techniques, the ability of the GCM to simulate recent climate and climate variability is of paramount importance to any study that uses GCM output to examine climate change.
3) The statistical link between the free atmosphere predictor and local climate variable must be valid under changed climate conditions. While the inability to test this requirement is the major drawback of the empirical method, availability of long data series for model development may allow some testing of the robustness of relationships under a range of climate conditions.

4) The climate change signal must be fully represented by the predictor variable(s). While many downscaling studies have used circulation metrics as the predictor variables, changes in regional climate will likely have sources other than changes in circulation, and the need to include other parameters (e.g., large scale temperature as a surrogate for increased radiation, or atmospheric humidity to differentiate between moist and dry episodes with weather classes) is being recognized (Charles et al. 1999; Giorgi et al. 2001).

1.2 Research Objectives

The research presented in this dissertation contributes to two key themes of global climate change research:

1) Evaluation of GCMs to:
    a. Provide confidence bounds for climate prognoses
    b. Identify parameters and processes that are not well simulated and hence identify (i) model components that may require modification and (ii) identify robust variables or parameters ensembles that may be used for empirical downscaling
2) Development and evaluation of downscaling tools

The geographic focus of the research is the Midwest region of the USA, although the techniques used have much broader applications. The specific parameter foci for the downscaling are temperature and precipitation due to their direct relationship with anticipated climate change impacts in the study region (see Chapter 2). In the following subsections, I introduce specific research objectives.

1.2.1 To evaluate the homogenized, gridded data sets (i.e., reanalysis products) commonly used in empirical downscaling.

High-quality observed climate data are a critical tool for successful empirical downscaling. These data are needed not only for model specification and validation, but also for evaluation of GCMs and their ability to represent the current climate. In practice, climatologists often employ data derived from reanalysis products (e.g., NCEP/NCAR reanalysis: Kalnay et al. 1996; ECMWF reanalysis: Gibson et al. 1997). However, evaluation of these reanalyses relative to independent data sets has resulted in the documentation of several shortcomings, particularly in data sparse regions (e.g., Swail and Cox 2000; Marshall 2002; Trenberth and Stepaniak 2002). Previous evaluation of reanalysis data has focused on individual variables. Here I focus on ensembles representing fundamental features and scales and inter-relationships of the data set variables. Specifically, this research evaluates the ability of the NCEP/NCAR reanalysis data set to reproduce synoptic-scale climate and variability relative to radiosonde station data using multivariate synoptic classification methods. The data and methods needed to meet this objective are described in Sections 2.2 and 3.2, respectively. The results of the analysis are presented in full in Chapter 4.

1.2.2 To use the historical climate record to identify physically meaningful large-scale variables accounting for a large proportion of the variance of surface climate parameters.

An important aspect of any downscaling study is the identification and quantification of relationships between large-scale predictors and small-scale predictands. The availability of atmospheric reanalysis products now makes it possible to consider a large number of atmospheric predictor variables. A systematic evaluation of the relative skill of different atmospheric predictors is absent from the literature and is a top priority for the climate change research community (Winkler et al. 1997; Giorgi et al. 2001). Even after several downscaling review articles (Hewitson and Crane 1996; Kattenberg et al. 1996; Wilby and Wigley 1997), there is little consensus among downscaling studies regarding the proper choice of atmospheric predictor variables (Wilby and Wigley 2000). Further, predictor variables are often chosen without quantitative evaluation. The research presented here provides the first comprehensive and quantitative assessment of the utility of a suite of free atmosphere predictor variables (including sea level pressure, geopotential height, upper air temperature and specific humidity, and derived variables such as geostrophic flow and vorticity components (see Section 3.1)) for use in downscaling temperature and precipitation in the Midwestern USA. The data required to meet this objective are described in Section 2.2. The results are presented in full in Chapter 5.

1.2.3 To evaluate multiple GCM simulations and identify well-simulated variables for use in generating regional climate prognoses.

A fundamental assumption of any downscaling study is that the parent GCM responds realistically and properly to atmospheric changes associated with a given emissions scenario. However, due to differences in model formulation, resolution, boundary conditions, and parameterization of sub-grid scale processes, two GCMs can have very different responses to the same emission scenario. Hence, in this dissertation, simulations from two GCMs are evaluated at both the grid point level and in terms of synoptic phenomena. The grid point evaluation is conducted on multiple timescales, from daily to annual. The diagnosis of GCM-simulated synoptic phenomena is based on 1) the agreement between the observed and GCM simulated mean fields of several large scale variables, 2) the ability of the GCMs to reproduce teleconnections of interest to regional climate in the study area (e.g., the Pacific/North American (PNA) pattern and the North Atlantic Oscillation (NAO), 3) the ability of the GCMs to reproduce the circulation climatology of the study area, and 4) the ability of the GCMs to reproduce the links between the teleconnections and the circulation types. The description of the data and methods needed to meet this objective are described in detail in Sections 2.2 and 3.1-3.5, respectively. The results of this analysis are presented in Chapters 6 and 7.

1.2.4 To apply, compare, and evaluate multiple empirical downscaling methods.

Multiple methods exist for empirical downscaling of GCM output, each with relative strengths and weaknesses (Wilby and Wigley 1997; Xu 1999). Although many studies use a single method, application of multiple methods can help to identify the

relative merits of a particular approach. The most common approaches are based on transfer functions, which use direct relationships between the free atmosphere predictor and the surface variable of interest. These transfer functions vary in complexity from simple regression models to artificial neural networks (e.g., Kim et al. 1984; Schoof and Pryor 2001). Weather typing approaches, which relate the surface climate variable of interest to the large-scale state (usually defined in terms of circulation), are also commonly used. Another alternative is use of weather generators, a class of random number generator which produces realistic climate sequences, which can be conditioned on the large-scale state (e.g., Wilks 1992; Wilks 1999). In this study, several empirical downscaling methods are applied. The data and methods needed to meet this objective are described in Sections 2.2 and 3.6, respectively. The results are given in Chapter 8. An additional paper describing a specific weather generator, WGEN, and the impact of stochastic weather generator parameters on simulation results, is included in Appendix A.

CHAPTER 2: STUDY REGION AND DATA

2.1 Study region

As stated in the introduction, this research focuses on the development of regional climate scenarios for the Midwest region of the United States, which extends from approximately 35° to 50°N latitude and from 75° to 95°W longitude (Figure 2.1). This region is highly sensitive to climate change impacts for several reasons (Sousounis and Albercook 2000). First, the region is a major agricultural center, producing a large proportion of the nation's corn and soybeans. Secondly, with the exception of the polar ice caps, the Great Lakes are the world's largest source of fresh water and provide drinking water and hydroelectric power to the region. The lakes additionally serve as a major transportation system, linking the region with the Atlantic Ocean via the St. Lawrence Seaway. While continued increases in growing season length (e.g., Robeson 2002) could benefit agriculture in the study area, the region is also susceptible to a number of negative climate change impacts. Examples of these key issues include reductions in lake and river levels, increases in heat related stress and mortality, shifts in agricultural productivity, and negative ecological impacts, such as runoff of excess nutrients used for fertilization.

2.1.1 Current climate of the Midwestern United States

During the winter, the synoptic scale climate of the Midwestern United States is characterized by high day-to-day variability associated with the position of the polar jet stream, which is often located over the region, and the associated polar front. Mid-latitude cyclones, which form in the presence of these features, are the main sources of

precipitation in the region during the winter months. During summer, the polar front retreats into Canada and frontal passages are less common. Hence, during these months, a greater proportion of precipitation is convective in nature and moisture availability is largely determined by the position and intensity of the subtropical (Bermuda) high (often located within or to the east or southeast of the study area). Variations in the position and intensity of the subtropical high and other semi-permanent pressure systems affecting the region (e.g., the Hudson Bay low), as well as the tracking and intensity of synoptic scale phenomena in the study region, have been linked to larger pressure oscillations (or teleconnection patterns), specifically the North Atlantic Oscillation (NAO) and Pacific North American (PNA) pattern (see Chapter 3).

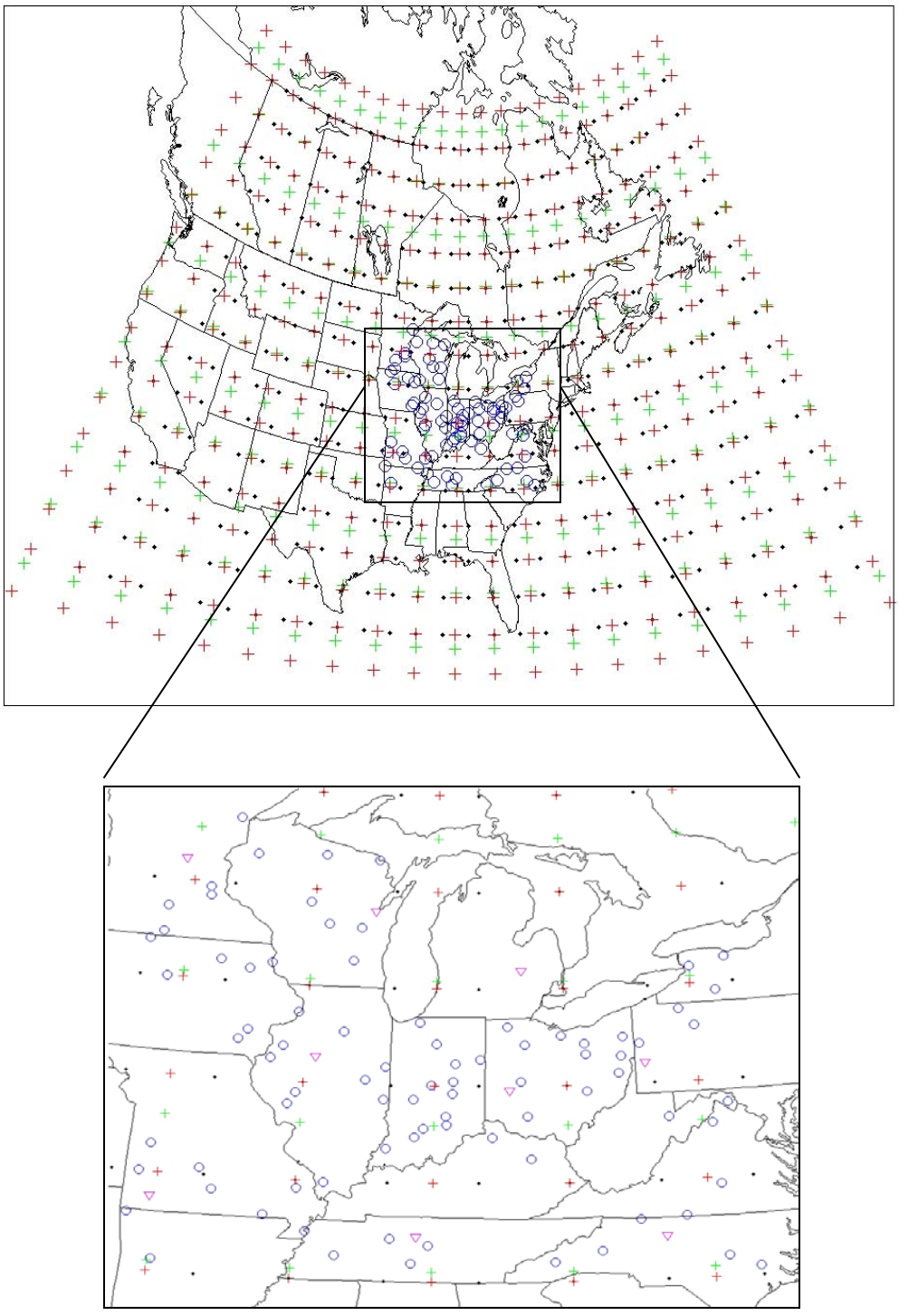

Figure 2.1 Map of the Midwest region of the United States, including the locations of the NCEP/NCAR and ECMWF reanalysis predictor grid points (•), HadCM3 grid points (+), CGCM2 grid points (+), radiosonde stations (▽), and surface (USHCN/D) stations (o). See Section 2.2 for a complete description of each data set.

2.1.2 Recent climate change in the Midwestern United States

The climate of the Midwestern United States has changed considerably over the past 100 years. The northern part of the region (i.e., the upper Great Lakes) has warmed by approximately 2°C, while the southern and eastern parts of the region have warmed slightly or even cooled (Folland et al. 2001). However, during the last quarter century, all parts of the region have warmed, with surface air temperature trends of 0.4 to 0.8°C per decade, and the largest warming during the winter (Folland et al. 2001). Robeson (2004) found that different parts of the daily maximum and minimum air temperature (T_{max} and T_{min}) probability distributions have exhibited different trends over the last 50 years. Although there was substantial spatial variability in the results, most stations in the study area (see Figure 2.1) have experienced the largest temperature trends in the highest percentiles of both T_{max} and T_{min} during late winter through early spring (>1.5°C and >1°C per 50 years, respectively). Stations in the extreme NW part of the study region have experienced much large trends in T_{min}, with the lowest minimum temperatures exhibiting trends greater than 3°C per 50 years during spring. During other times of the year, most stations have experienced small decreases in T_{max} and relatively little change in T_{min}. Precipitation has also increased over most of the region during the last century, with increases as large as 30% in some regions. In the southeastern part of the study area, precipitation has been relatively stable or decreased slightly (Folland et al. 2001).

2.2 Data

Several types of data are used to meet the research objectives stated in Chapter 1. They are described in detail in the following sections.

2.2.1 Reanalysis data

In this research, data are used from both the NCEP/NCAR reanalysis project (Kalnay et al. 1996) and the ECMWF reanalysis project (Gibson et al. 1997). These data are gridded at a horizontal resolution of 2.5° × 2.5°, with either daily or twice daily output on multiple atmospheric levels. The atmospheric model used for the NCEP/NCAR reanalysis is spectral with transformation to a Gaussian grid for calculation of nonlinear quantities and physics. The horizontal resolution is spectral triangular 62 (T62), which is roughly equivalent to a horizontal resolution of 200 km, though, for most variables, they are archived at the coarser resolution described above. The vertical domain is divided into 28 sigma (terrain-following) levels with enhanced resolution near the bottom and the top. The atmospheric model used for the ECMWF reanalysis has T159 spherical-harmonic representation for basic dynamic fields, with a reduced Gaussian grid of approximately uniform 125 km spacing for surface and other grid-point fields. The vertical domain is divided into 60 levels. The reanalysis data used in this study span either 1948-2001 (NCEP/NCAR reanalysis) or mid-1957-2001 (ECMWF reanalysis).

In the analysis of the synoptic climate of the study area in radiosonde and reanalysis data, I use NCEP/NCAR 500-hPa geopotential heights, 850-hPa air temperatures and 200-hPa wind speeds to broadly represent circulation patterns, thermal

advection, and jet stream strength, respectively. In the downscaling analysis I use ECMWF 850-hPa specific humidity and NCEP/NCAR 850-hPa relative humidity, 500 and 850-hPa geopotential height, 850-500 hPa thickness, and sea-level pressure along with geostrophic flow and vorticity components derived from the reanalysis sea-level pressure fields.

2.2.2 Radiosonde data

One of the primary foci of this research is evaluation of the synoptic climate of the study region as manifest in the reanalysis data relative to radiosonde observations. Nine radiosonde stations within the study area (Figure 2.1) have data available from the Radiosonde Data of North America 1946-1996 database (FSL/NCDC 1997). The variables are chosen to match those from the reanalysis (described above) and include twice-daily (0 GMT, 12 GMT) observations of 500-hPa heights, 850-hPa air temperatures, and 200-hPa wind speeds. While the length and quality of the record varies from station to station (Gaffen 1994; Gaffen et al. 2000), all stations have data available for the summer during the period 1971-1993. This time period therefore serves as the focus of the comparison. The synoptic classification methodology is described in detail in Section 3.2 and the comparison with reanalysis output is described in Chapter 4.

2.2.3 Surface station data

Construction of the downscaling models requires high quality surface station data from throughout the study region. The Daily United States Historical Climatology Network (Easterling et al. 1999) contains daily data from 1064 stations within the

contiguous United States and includes five elements: maximum and minimum daily air temperature, daily precipitation amount, daily snowfall amount, and snow depth. The stations chosen for inclusion in the USHCN/D are a subset of the 1221 stations in the monthly USHCN (Karl et al. 1990) and hence, most of the stations are long-term cooperative stations. However, some 1st order National Weather Service (NWS) stations have been added for increased spatial resolution and coverage.

Historical temperature and precipitation data are susceptible to a number of biases and inhomogeneities resulting from changes in the station environment or observing practices (e.g., urbanization, station moves, instrumentation and time of observation changes). The complex process of identifying and removing these effects has been the subject of intense study (e.g., Jones 1994; Jones et al. 1997; Peterson et al. 1998). In this study, stations are included if the data record from the station is at least 95% complete over the period 1958-2001 and at least 90% complete within each individual year. To ensure consistency between the station data and large-scale free atmosphere predictors, only stations with a consistent observation time are included in this study. With these constraints, 84 stations are available within the study region (Figure 2.1, Table 2.1). For each station, maximum and minimum daily air temperature and total daily precipitation are available for at least the period 1958-2001. These data are used in conjunction with the reanalysis products and GCM simulations to produce regional climate scenarios for the study region.

Table 2.1 List of the 84 USHCN/D stations used in this study.

Station Name	Lat (°N)	Lon (°W)	Station Name	Lat (°N)	Lon (°W)
Gravette, AR	36.44	94.45	Lamar, MO	37.51	94.27
Subiaco, AR	35.30	93.66	Lebanon, MO	37.67	92.66
Anna, IL	37.47	89.24	Marble Hill, MO	37.30	89.97
Hoopston, IL	40.47	87.67	Mountain Grove, MO	37.16	92.27
Jacksonville, IL	39.74	90.20	Angelica, NY	42.30	78.02
La Harpe, IL	40.59	90.97	Buffalo, NY	42.92	78.74
Monmouth, IL	40.92	90.64	Rochester, NY	43.14	77.67
Ottawa, IL	41.34	88.92	Kinston, NC	35.22	77.54
Paris, IL	39.62	87.70	Morganton, NC	35.76	81.69
Urbana, IL	40.11	88.24	Mount Airy, NC	36.52	80.62
White Hall, IL	39.44	90.39	Smithfield, NC	35.52	78.35
Anderson, IN	40.11	85.72	Waynesville, NC	35.49	82.97
Berne, IN	40.67	84.95	Cadiz, OH	40.27	81.00
Columbus, IN	39.21	85.92	Chippewa Lake, OH	41.05	81.94
Greencastle, IN	39.64	86.85	Findlay, OH	41.05	83.67
Greenfield, IN	39.79	85.75	Hillsboro, OH	39.21	83.62
LaPorte, IN	41.61	86.72	Millport, OH	40.72	80.91
Marion, IN	40.57	85.67	Norwalk, OH	41.27	82.62
Oolitic, IN	38.89	86.55	Philo, OH	39.84	81.92
Princeton, IN	38.36	87.59	Urbana, OH	40.11	83.79
Rochester, IN	41.07	86.22	Warren, OH	41.21	80.82
Seymour, IN	38.99	85.91	Wauseon, OH	41.52	84.16
Shoals, IN	38.67	86.80	Wooster, OH	40.79	81.92
Whitestown, IN	40.01	86.35	New Castle, PA	41.02	80.37
Charles City, IA	43.05	92.67	Ridgway, PA	41.42	78.75
Clinton, IA	41.80	90.27	Warren, PA	41.86	79.16
Estherville, IA	43.42	94.84	Dickson, TN	36.07	87.39
Fairfield, IA	41.04	91.95	Jackson, TN	35.62	88.84
Fayette, IA	42.86	91.80	Lewisburg, TN	35.46	86.80
Fort Dodge, IA	42.51	94.2	Murfreesboro, TN	35.92	86.37
Washington, IA	41.29	91.69	Danville, VA	36.59	79.39
Farmers, KY	38.12	83.55	Farmville, VA	37.34	78.39
Williamstown, KY	38.66	84.62	Woodstock, VA	38.91	78.47
Iron Mountain, MI	45.79	88.09	Martinsburg, WV	39.41	77.99
Cloquet, MN	46.71	92.52	Parsons, WV	39.11	79.67
Fairmont, MN	43.64	94.47	Hancock, WI	44.12	89.54
Farmington, MN	44.67	93.19	Marshfield, WI	44.66	90.14
Minneapolis, MN	44.89	93.22	Minocqua Dam, WI	45.89	89.74
New Ulm, MN	44.30	94.45	Oshkosh, WI	44.04	88.55
Appleton City, MO	38.21	94.04	Prairie Du Chien, WI	43.04	91.16
Caruthersville, MO	36.21	89.67	Spooner, WI	45.82	91.89
Doniphan, MO	36.59	90.82	Watertown, WI	43.19	88.74

2.2.4 General circulation model (GCM) data

Output from coupled atmospheric-oceanic GCMs is used here as input to the downscaling analyses. GCMs have provided climate scenarios for the vast majority of previous regional climate change studies. However, many studies have used a single or equilibrium response GCM to assess possible regional climate change impacts. Because significant differences exist between climate models, primarily due to sub-grid scale parameterizations, application of a single model does not provide results that adequately reflect the uncertainty inherent in regional climate scenario generation (see discussion in Giorgi et al. 2001). Therefore, in this study, the output from two transient (i.e., based on gradual changes in forcing and response) GCMs is used: 1) the Hadley Centre 3rd generation coupled oceanic-atmospheric general circulation model (HadCM3; Gordon et al. 2000; Pope et al. 2000) and 2) the Canadian Centre for Climate Modelling and Analysis (CCCma) 2nd Generation coupled general circulation model (CGCM2; Flato et al. 2000; Flato and Boer 2001).

HadCM3 and CGCM2 were chosen to provide a range of characteristics associated with the current generation of coupled models and therefore differ in several respects. First, HadCM3 is a Cartesian model (i.e., the atmospheric and ocean are divided into grid boxes for which the model equations are solved), whereas CGCM2 is a spectral model (i.e, in any latitudinal band, the model quantities are represented a summation of a number of waves). Second, CGCM2 employs flux adjustments (e.g., nudging of ocean-atmosphere heat and water fluxes in an effort to achieve a stable and realistic representation of current climate conditions) as first described in Hansen et al. (1984). HadCM3, with slightly higher oceanic resolution (1.25°×1.25° vs. 1.8°×1.8°),

does not employ flux adjustments. While results from the Coupled Model Intercomparison Project (CMIP) suggest that the basic patterns of climate system response to external forcing are relatively robust in GCMs with and without flux adjustments (Meehl et al. 2000), models which do not need flux adjustments can be viewed as physically self-consistent representations of the climate system (McAveney et al. 2001).

The atmospheric component of HadCM3 has 19 levels with a horizontal resolution of approximately 2.5° latitude by 3.75° longitude, which produces a global grid of 96 × 73 points. This is equivalent to a surface resolution of about 417 km × 278 km at the equator, reducing to 295 km × 278 km at a latitude of 45° (comparable to a spectral resolution of T42). The atmospheric component of CGCM2 has ten vertical levels and employs a triangular spectral truncation having 32 longitudinal waves (T32/L10 or approximately 3.75° × 3.75°). The output from HadCM3 used in this analysis was archived at a spatial resolution of 2.5° latitude × 3.75° longitude, while that from CGCM2 was archived at approximately 3.75° × 3.75°. The model grid points are shown in Figure 2.1. Prior to use in this study, both GCM grids were interpolated to the 2.5° × 2.5° reanalysis grid (Figure 2.1) using an inverse distance based interpolation algorithm. Additionally, the output from HadCM3 was archived with a 360-day year. For comparison with observed data and output from CGCM2, this 360-day year was projected onto a 365-day year. The transient simulations from HadCM3 and CGCM2 span the periods 1990-2099 and 1990-2100, respectively.

Climate evolution and variability are critically dependent on changes in socio-economic factors, particularly those affecting emissions behavior. Therefore, projected

changes in these quantities are an integral part of climate scenario generation, but may not be consistent between models and model applications. The most comprehensive emissions scenarios currently available are those from the IPCC Special Report on Emissions Scenarios (SRES), which detail emissions of carbon dioxide (CO_2), as well as other atmospheric constituents not accounted for by previous emissions scenarios (IPCC 2000). The GCM simulations used here were conducted using the SRES A2 emissions scenario (IPCC 2000). The SRES scenarios encompass a range of parameters that dictate emissions including demographic, technological and economic changes. The SRES A2 scenario group results in global CO_2 emissions from industry and energy in 2100 that are almost four times the 1900 value and emissions from land use change by 2100 are close to zero, leading to a global CO_2 emission in 2100 of almost 28 GtC yr^{-1}. This emissions scenario equates to a moderate to high greenhouse gas cumulative emission for 1990 to 2100 as a result of projected population growth and fairly slow introduction of alternative technologies, and is used in this study to provide an upper bound on likely climate change and hence a high signal to noise ratio when comparing current and future climates. Given identical A2 forcing scenarios, the factors discussed above (e.g., parameterizations, model differences, etc.) dictate that the projections from these models differ markedly in terms of surface temperature and precipitation response (Cubasch et al. 2001). According to Chapter 9 of the IPCC Third Assessment Report (TAR; IPCC 2001), CGCM2 A2 exhibits a relatively large increase in global mean temperature over the 21st century (> 4°C), while HadCM3 A2 exhibits an increase of around 3.5°C. In terms of precipitation, control runs from both models were shown (along with the entire suite of models analyzed by the TAR) to over-estimate northern hemisphere winter mid-latitude

precipitation, while the transient simulations indicate approximately a 3% increase in global precipitation by the end of the 21st century.

CHAPTER 3: RESEARCH METHODS

Several methods are needed to meet the research objectives stated in Chapter 1. These methods are used 1) to quantify characteristics of the historical and GCM simulated climate of the study region, 2) to compare characteristics of the historical and GCM simulated climate, or 3) with GCM output to produce regional climate scenarios. The methodologies are described in the following sections. The specific details of their application are described in the chapters that follow.

3.1 Definitions of derived variables

3.1.1 Teleconnection indices

At the synoptic scale, climate variability in the study region is largely governed by the behavior of the polar jet stream and the relative intensity of several semi-permanent pressure systems. Variations in the intensity of these pressure systems and the tracking and intensity of synoptic scale phenomena in the study region have been linked to larger pressure oscillations (or teleconnection patterns), specifically the North Atlantic Oscillation (NAO) and Pacific North American (PNA) pattern. Both the NAO and PNA are prominent hemispheric-scale modes of climate variability that are strongly linked to surface temperature and precipitation at least partly via synoptic scale circulation patterns. For example, the NAO accounts for 31% of the variance in mid-latitude winter mean surface air temperatures (Hurrell 1996) and the PNA is closely linked with both thermal and hydrologic regimes within the Ohio Valley (Coleman and Rogers 2003; Sheridan 2003), accounting for as much as 40% of the variance in surface air temperature during winter season (Leathers et al. 1991).

The NAO and PNA teleconnection indices describe patterns of mean sea level pressure (MSLP) and tropospheric height variability, respectively, and hence represent the steering patterns of synoptic scale phenomena (Leathers et al. 1991; Hurrell 1995). They are used here both as an evaluation tool for GCMs and to explain differences in the synoptic scale climate in the study region across time or data sets. Both the NAO and PNA exhibit variability on a range of temporal scales. Stephenson and Pavan (2003) show that the NAO signal is dominated by short-term (i.e., year to year) variations and state that more than 70% of the NAO variance is explained by fluctuations with periods of less than a decade. Likewise, the PNA teleconnection varies on all timescales from days to decades (Blackmon et al. 1984; Yarnal and Leathers 1988).

3.1.1.1 The North Atlantic Oscillation (NAO)

The NAO describes a redistribution of atmospheric mass between the Arctic and the subtropical Atlantic, typically characterized by winter sea-level pressure anomalies associated with the Azores high and the Icelandic low (Walker 1924; Walker and Bliss 1932). During the positive phase of the NAO, both of these pressure centers are stronger than average and shifted to the northeast (Glowienka-Hense 1990), resulting in stronger westerlies in the mid-latitude Atlantic Ocean. Although it is centered east of North America, the positive phase of the NAO is generally associated with higher temperatures and increased moisture content in the Midwestern USA in response to the strengthening of the flow around the subtropical Atlantic high generating southerly flow over the eastern USA (Dickson and Namias 1976; Yarnal and Leathers 1988; Yin 1994).

There is no unique way to define the spatial structure of the NAO (Hurrell et al. 2003). It has been defined in terms of pressure differences between point locations, area weighted pressure extremes or in terms of principal components of pressure fields (see Osborn et al. 1999; Hurrell et al. 2003). In this study, the NAO index is computed as the difference in standardized (i.e., transformed to have a mean of zero and variance of one by subtracting the mean and dividing by the standard deviation) SLP between grid points near the NAO "centers of action" or:

$$NAO = SLP_{Azores} - SLP_{Iceland} \qquad \textbf{3.1}$$

where SLP_{Azores} is the standardized SLP at Ponta Delgada, Azores (37.7°N, 25.7°W) and $SLP_{Iceland}$ is the standardized SLP at Stykkisholmur, Iceland (65.1°N, 22.7°W).

3.1.1.2 The Pacific/North American (PNA) pattern

The mean flow over the Pacific-North American sector is characterized by a trough in the east-central North Pacific, a ridge over the Rocky Mountains, and a trough over eastern North America (Leathers et al. 1991; Kang et al. 2002). The PNA teleconnection index reflects deviations from this mean flow, suggesting more meridional (positive phase) or zonal (negative phase) flow over North America. During the positive phase, negative geopotential height anomalies are located south of Alaska and in the southeastern United States, while positive geopotential height anomalies are located near Hawaii and over the Rocky Mountains resulting in meridional flow over North America. During the negative PNA phase, the anomalies at the pressure centers are reversed and the flow over North America is more zonal.

The PNA index used in this study is computed using the equation of Wallace and Gutzler (1981):

$$PNA = \frac{1}{4}\left[Z(20^\circ N,160^\circ W) - Z(45^\circ N,165^\circ W) + Z(55^\circ N,115^\circ W) - Z(30^\circ N,85^\circ W)\right] \quad \textbf{3.2}$$

where Z are standardized (by season within the study period) 500 hPa geopotential height values.

3.1.2 Geostrophic flow and vorticity component definitions

Several studies have identified geostrophic flow and vorticity components as useful downscaling predictors (Wilby 1998; Wilby 1998b). These derived variables can be estimated for any regular grid from sea level pressure fields using the method of Dessouky and Jenkinson (1975). Specifically, given the hypothetical grid shown in Figure 3.1, the flow and vorticity components are computed as follows.

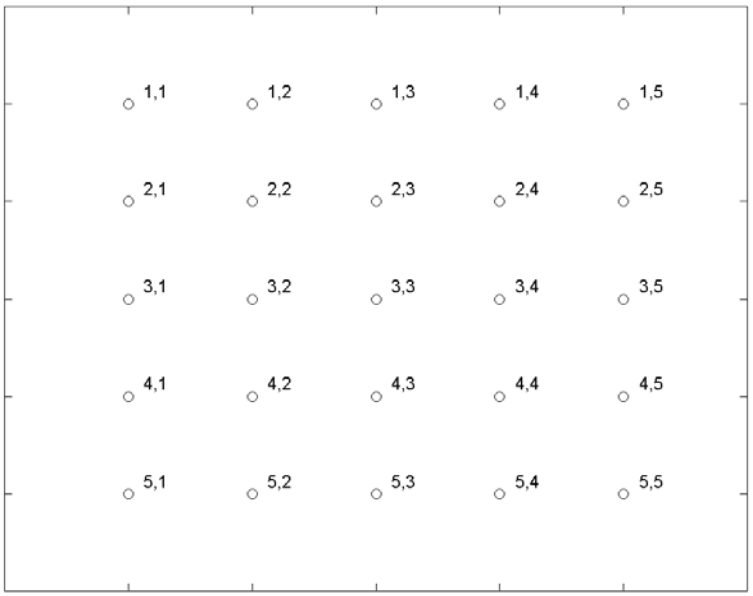

Figure 3.1 Hypothetical 2.5° × 2.5° grid for demonstration of geostrophic flow and vorticity calculations.

The westerly component of surface geostrophic wind, GEOW, at the central point ($P_{3,3}$) is given by:

$$GEOW = \frac{1}{4}(P_{5,1} + 2P_{5,3} + P_{5,5}) - \frac{1}{4}(P_{1,1} + 2P_{1,3} + P_{1,5}) \quad \textbf{3.3}$$

where P is the mean sea-level pressure (SLP) and the subscript denotes the grid point.

The southerly component of the surface geostrophic wind, GEOS, at the central grid point (3,3) is given by:

$$GEOS = c_1 * (\frac{1}{4}(P_{1,5} + 2P_{3,5} + P_{5,5}) - \frac{1}{4}(P_{1,1} + 2P_{3,1} + P_{5,1})) \quad \textbf{3.4}$$

where c_1 is a constant equal to $\frac{1}{\cos(latitude(3,3))}$ which accounts for the relative difference in grid spacing in the north-south and east-west directions.

The resultant geostrophic flow, GEOWS, at the central grid point (3,3) is then computed as:

$$GEOWS = \sqrt{W^2 + S^2} \quad \textbf{3.5}$$

The westerly component of shear vorticity, GEOZW, at the central grid point (3,3) is given by:

$$\begin{aligned} GEOZW = 2 * (c_2 * (\frac{1}{4}(P_{5,1} + 2P_{5,3} + P_{5,5}) - \frac{1}{4}(P_{3,1} + 2P_{3,3} + P_{3,5})) \\ - c_3 * (\frac{1}{4}(P_{3,1} + 2P_{3,3} + P_{3,5}) - \frac{1}{4}(P_{1,1} + 2P_{1,3} + P_{1,5}))) \end{aligned} \quad \textbf{3.6}$$

where c_2 is a constant equal to $2*\frac{\sin(latitude(3,3))}{\sin(latitude(5,3))}$ and c_3 is a constant equal to $2*\frac{\sin(latitude(3,3))}{\sin(latitude(1,3))}$. Both constants account for the relative difference in grid spacing in the north-south and east-west directions.

The southerly component of shear vorticity, GEOZS, at the central grid point (3,3) is given by:

$$GEOZS = c_4 * ((\frac{1}{4}(P_{1,5} + 2P_{3,5} + P_{5,5}) - \frac{1}{4}(P_{1,3} + 2P_{3,3} + P_{5,3})) - (\frac{1}{4}(P_{1,3} + 2P_{3,3} + P_{5,3}) - \frac{1}{4}(P_{1,1} + 2P_{3,1} + P_{5,1}))) \quad \textbf{3.7}$$

where c_4 is a constant equal to $\frac{2}{\cos(latitude(3,3))}$, again to account for the relative differences in north-south and east-west grid spacing.

The total shear vorticity, GEOZT, is taken as the sum of GEOZW and GEOZS. The units for the geostrophic flow variables are hPa per 10° of latitude at the grid point latitude. The vorticity variables are expressed in units of hPa per 10° of latitude at the grid point latitude, per 10° of latitude.

3.2 Synoptic classification methods

Classification of large-scale weather information is useful for both 1) examination of surface climate under different forcing conditions and 2) evaluation of the synoptic scale climate within GCMs. Several methodologies are available for the development of synoptic-scale weather classifications (Yarnal 1993). In this dissertation, two methods are

used: 1) correlation-based methods and 2) eigenvector/clustering-based methods. While correlation based methods are readily amenable to both development of targeted classifications for the GCM output relative to observations and comparison of the observed and GCM classifications using bootstrapping techniques, the eigenvector/clustering technique is easily applied to multivariate data sets.

3.2.1 Correlation based-map typing

Correlation-based methods of synoptic-scale weather classification were first used by Lund (1963), who classified map patterns using the Pearson product-moment correlation of gridded fields. Kirchhofer (1974) improved on the correlation-based map typing technique by requiring that sub-map scale patterns also meet the correlation threshold (El-Kadi and Smithson 1992). The Kirchhofer score of similarity, S, for two grid point maps, containing x and y, is given by:

$$S = \sum_{i=1}^{n} (z_{xi} - z_{yi})^2 \qquad \textbf{3.8}$$

In (3.8), z_{xi} and z_{yi} are the normalized (spatially) grid-point values of x and y, respectively, and n is the number of grid points. Normalization of the data ensures that the value of S indicates similarity in map patterns and is not contaminated by the magnitudes of the values being classified (Blair 1998). The relationship between the Kirchhofer score, S, and the correlation coefficient, r (Willmott 1987; Blair 1998) is given by:

$$S(n-1)^{-1} = 2[1 - r(x,y)] \ , \ 0 \leq S(n\text{-}1)^{-1} \leq 4 \qquad \textbf{3.9}$$

By setting a threshold for the required correlation, (3.9) can be used to determine the appropriate threshold for S. It is then necessary to calculate S for all combinations of pairs of observations being classified. Sub-grid scale similarity is ensured by also requiring each of the rows and columns of the grids to meet the similarity criteria. The observation with the maximum number of threshold exceedances is referred to as the first key day. This key day and all 'similar' observations are then removed from the analysis. All days similar to those days are also removed. This process is then repeated until there are no more days left. Once all of the key days have been identified, each observation that has been removed is assigned to the key day for which it has the highest S value.

Different thresholds for grid and sub-grid similarity and minimum group size have previously been applied for different variables (Yarnal 1985; McKendry et al. 1995). In the study presented in Chapter 7, the correlation threshold for the entire gird is set at 0.75, the individual row and column thresholds are set at half of the entire grid threshold (0.375), and the map-pattern groups are required have more than five members. These values are chosen to ensure a manageable number of weather types and a high percentage of classified days.

3.2.2 Eigenvector/clustering-based synoptic classification

As the name implies, the eigenvector/clustering-based synoptic classification method consists of two steps: eigenvector analysis and cluster analysis. Eigenvector analysis identifies modes of variability within the input data by identifying orthogonal linear combinations of input variables and projecting the input data onto those modes. The relationships between these modes and the input observations are then clustered to

yield synoptic weather types. In this study, as reported in detail in Chapter 4, rotated principal components analysis (PCA) (Richman 1986) is followed by a two-step clustering algorithm (Davis and Kalkstein 1990).

3.2.2.1 Principal components analysis (PCA)

The principal component model expresses the input data (Z) in terms of two matrices: the principal component scores matrix (F, representing the relationship between the principal components and the individual observations) and the principal component loadings matrix (A, representing the relationship between the principal components and the individual input variables):

$$Z = FA^T \qquad \textbf{3.10}$$

The first principal component is a linear combination of the original variables defined such that the maximum amount of variance is explained. The second principal component explains the greatest amount of residual variance with the addition constraint that it must be orthogonal to the first principal component. Additional principal components are defined similarly. A full mathematic description of PCA, including derivation of the principal component scores and loadings matrices is available in Richman (1986).

In this analysis PCA is applied primarily as a data reduction technique and to eliminate correlation between input variables for subsequent cluster analysis. However, the PC loadings and PC scores are also valuable tools in the interpretation of the classification results. While the full PCA solution explains all of the variance present in the original data, truncated PCA solutions can usually explain a large proportion of that

variance. The number of PCs to retain 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). To aid in the interpretation of the solution, the resulting PCs are rotated using (orthogonal) Varimax rotation (Richman 1986; Bonell and Sumner 1992; Brinkmann 1999). The truncated principal component scores matrix is then subjected to cluster analysis to yield meaningful climatological classes.

3.2.2.2 Cluster analysis

Cluster analysis refers to the classification of a number of objects into a smaller number of homogeneous (or semi-homogeneous) groups. Clustering techniques can be classified as either hierarchical or nonhierarchical (see Gong and Richman 1995). Hierarchical methods either merge objects to form clusters (agglomerative clustering) or split objects to form clusters (divisive clustering), while non-hierarchical methods begin with a specified number of clusters and find the optimal classification of data into the specified number of clusters. Since the number of synoptic weather types (the eventual outcome of the cluster analysis presented here) is not known a priori, hierarchical methods are first applied and then followed by a non-hierarchical method to allow reassignment of poorly placed observations.

Hierarchical techniques begin with the number of clusters equal to the number of observations and combines similar observations until there is one large cluster. Different hierarchical algorithms use different techniques to calculate distance between clusters containing more than one observation. Single linkage, for example, uses the minimum

distance between the clusters. Complete linkage uses the maximum distance between the clusters. Average linkage uses either the average distance between the two clusters or the average distance between points in the newly formed cluster while the method of Ward (1963) minimizes the within-clusters sum of squares.

There have been a number of studies comparing clustering algorithms. Kalkstein et al. (1987) compared three clustering algorithms (average linkage, complete linkage, and Ward's method) and found that average linkage yielded the most realistic synoptic groupings. A similar result was obtained by Schoof and Pryor (2001). Although Ward's method has been used widely (e.g., Stone 1989; Fovell and Fovell 1993), the method tends to produce clusters of similar size, and hence, may under represent the frequency of the dominant synoptic type and over-represent important extreme events. Several studies have used a combination of clustering methodologies. For example, Davis and Kalkstein (1990) and Zelenka (1997) used average linkage followed by k-means cluster analysis, a non-hierarchical technique that forms new clusters given the number of clusters to form and the results of a separate cluster analysis. This combination allows the determination of the seeds needed for nonhierarchical clustering using hierarchical clustering.

Based on previous success in climatological analyses (Kalkstein et al. 1987; Schoof and Pryor 2001) the average linkage clustering method (Sokal and Michener 1958) is applied in this study (see Chapter 4). 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 re-classification 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.

3.2.3 Comparison of synoptic classifications

There are several methods available for the comparison of climatological classifications (Huth 2000). First, classifications can be applied to two datasets independently, allowing the dominant weather-types from both datasets to be identified. The second method of comparing classifications requires that the classes are predetermined, which is often difficult in climatological studies. While 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 dataset, followed by projection of the results onto the second dataset. This method classifies the second dataset according to the structure found in the first, providing further insight into the physical basis of both classifications. To meet the first research objective (see Chapter 1), a combination of the approaches outlined above

is used. This approach allows examination of independent classifications of the input fields and an examination of the improvement (in agreement) achieved by projecting the one data set onto the classes defined using the other (see Chapter 4).

3.3 Analysis and comparison of probability distributions

Differences in observed (i.e., derived from variables in the reanalysis data sets) and GCM-simulated (i.e., derived from GCM output) teleconnections will be manifest as changes in the probability distributions of the teleconnection indices. Therefore, the teleconnection indices from GCM and reanalysis output are evaluated and compared using two statistical methods as described in Chapter 5: the Kolmogorov-Smirnov (K-S) and χ^2 goodness-of-fit tests (D'Agostino and Stephens 1986). The K-S statistic is the largest absolute difference between empirical cumulative distribution functions (ECDFs). The χ^2 test is based on binned counts of the data being tested. It identifies differences over the discretized bins and suggests rejection of the null hypothesis when the sum of squared differences between binned counts is sufficiently large. In Chapter 6, both tests are conducted under the null hypothesis that the daily teleconnection indices from different GCMs and time periods are drawn from the same distribution.

3.4 Bootstrap resampling

The 12-year period of overlap (1990-2001) between the reanalysis data and GCMs used here does not represent a full climatology and is therefore insufficient for verification and validation of the GCMs based on observed data. Bootstrap resampling, which partially addresses this shortcoming, allows determination of the sampling

distribution for statistics from the relatively short period of record associated with historic climate data by repeatedly resampling the observed data (with replacement) and computing the appropriate statistics at each iteration, resulting in a sampling distribution which can be used to make inferences about agreement, or lack thereof, between observations and GCM simulations. Bootstrapping is used to address two different aspects of GCM simulations in this study: 1) grid point evaluation of the means and variances of model variables, and 2) evaluation of the frequency of synoptic scale weather types.

3.4.1 Evaluation of GCM-simulated grid point data

The method proposed by Chervin (1981), and later used and refined by Portman et al. (1992), is employed in this study to compare the means and variances of several variables as observed simulated by the GCMs described in Section 2.2. The method consists of two test variates, r_1 and r_2, related to the means and variances of the variable being tested, respectively. For any sample size and a priori confidence level, the critical values of the two variates can be found and compared to those computed to assess GCM performance at the grid point scale.

The first variate, a scaled difference of means of the GCM and observed time series is given by:

$$r_1 = \frac{(\mu_a - \mu_b)}{\sqrt{\sigma_a^2 + \sigma_b^2}} \qquad \textbf{3.11}$$

where μ_a and μ_b are the means of the GCM-simulated and observed time series, respectively, and σ_a^2 and σ_b^2 are the variances of the GCM-simulated and observed time

series, respectively. The composite estimate of variance is used in the denominator because the variances cannot be assumed to be equal (Chervin 1981). This quantity is equal to zero if the means are equal and the null hypothesis cannot be rejected.

The second variate is a simple ratio of standard deviations given by:

$$r_2 = \frac{\sigma_a}{\sigma_b} \qquad \textbf{3.12}$$

where σ_a and σ_b are the standard deviations of the GCM-simulated and observed time series, respectively. Used in conjunction, r_1 and r_2 provide information about the relationships between the distributions of observed and GCM-simulated variables.

3.4.2 Evaluation of GCM-simulated weather types

The bootstrap resampling methodology described above is also used to evaluate the ability of the GCMs to represent the circulation regimes of the study area. After application of the Kirchhofer method (Section 3.2.1) to the observed and GCM-simulated 500 hPa geopotential heights fields, the observed record is viewed as one possible realization of the current climate and other potential realizations are extracted using the bootstrap resampling method described above to generate a 'climatology' comprising 1000 random samples, which is presented in the form of confidence intervals for the mean map type frequencies.

3.5 'Taylor' diagrams

The ability of a model to properly simulate any field depends on the models ability to simulate both the pattern and its magnitude. For pattern similarity, the most commonly used measure is the correlation coefficient (r), given by

$$r = \frac{\frac{1}{N}\sum_{i=1}^{N}(m_n - \overline{m})(o_n - \overline{o})}{\sigma_m \sigma_o} \qquad \textbf{3.13}$$

where N is the number of grid points, m and o are the modeled and observed quantities, respectively, and σ_m and σ_o are the standard deviations of m and o, and the overbar indicates the mean. The statistic used to identify differences in the two fields is the root-mean-square (RMS) difference, given by

$$E = \left[\frac{1}{N}\sum_{i=1}^{N}(m_n - o_n)^2\right]^{1/2} \qquad \textbf{3.14}$$

which can be divided into systematic (bias) and non-systematic (pattern RMS error) terms, $\overline{E}$ and E', respectively.

The statistics described above are not completely independent (Taylor 2001), and are related according to:

$$E'^2 = \sigma_m^2 + \sigma_o^2 - 2\sigma_m \sigma_o R \qquad \textbf{3.15}$$

Taylor (2001) recognized that 3.15 closely resembles the law of cosines

$$c^2 = a^2 + b \ - 2ab\cos(\phi) \qquad \textbf{3.16}$$

where a, b, and c are the sides of a triangle and ϕ is the angle facing side c. This relationship allows the graphical representation of three quantities (r, E', and the ratio of the modeled to observed standard deviation) on a single two-dimensional plot (Figure 3.2).

An example of the Taylor diagram is shown in Figure 3.2. The cosine of the angle of any point from the horizontal axis is the correlation, r. The radial coordinate is the ratio of the model to observed standard deviation, providing a comparison of the model and observed amplitude of variations. When plotted in this way, the distance between any point and the origin (o) is proportional to the RMS difference. Because Taylor diagrams allow comparison of results from multiple models or model variations, they have been adopted by both the IPCC (McAveney et al. 2001) and the Coupled Model Intercomparison Project (CMIP, Covey et al. 2003) for model evaluation studies and are used here (Chapter 6) to compare the GCM-derived mean fields with those from the reanalysis.

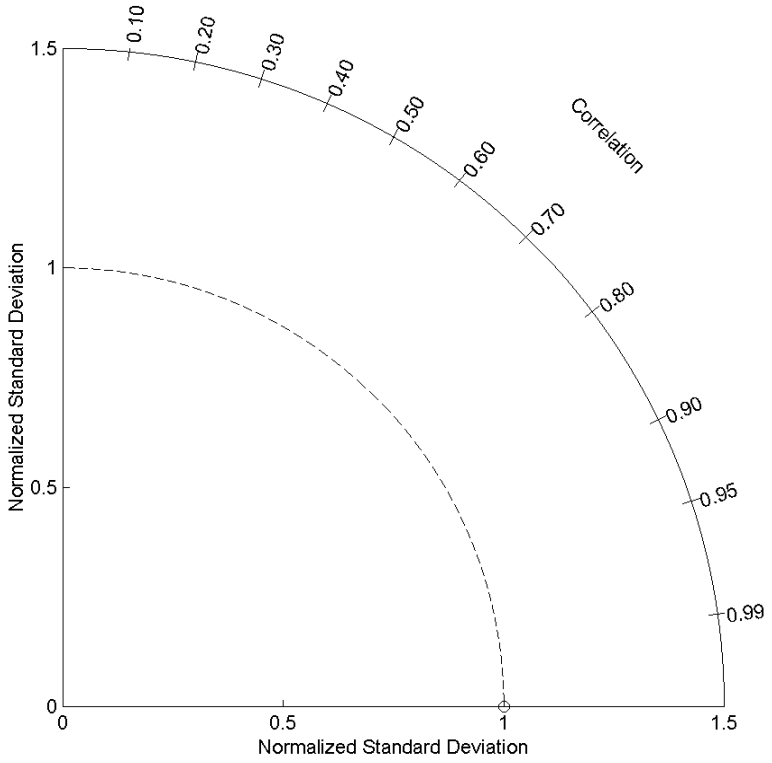

Figure 3.2. An example of the Taylor diagram.

3.6 The two-sample difference of proportions test

While the GCM simulation of the teleconnections and weather types can be evaluated using the methods described in Section 3.3 and Section 3.4, respectively, an additional statistical test is used to determine if the GCMs correctly simulate the links between the teleconnections and weather types. This is accomplished with the two-sample difference of proportions test (Ott 1993; Sheridan 2003).

First, each day is assigned as either positive (more than one standard deviation greater than the seasonal mean of the teleconnection), negative (more than one standard deviation less than the seasonal mean of the teleconnection), or neutral (within one standard deviation of the seasonal mean of the teleconnection). The dependence of the weather types on teleconnection phase is then tested using a z-score:

$$z = \frac{\pi_1 - \pi_2}{\hat{\sigma}} \qquad \textbf{3.17}$$

where π_1 and π_2 represent the proportions of positive and negative phase days for the teleconnection, and $\hat{\sigma}$ is:

$$\hat{\sigma} = \sqrt{\hat{\pi}(1-\hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \qquad \textbf{3.18}$$

where n_1 and n_2 are the total number of days in the positive and negative teleconnection phase, respectively, and $\hat{\pi}$ is

$$\hat{\pi} = \frac{n_1\pi_1 + n_2\pi_2}{n_1 + n_2} \qquad \textbf{3.19}$$

Proportions resulting in a large (absolute) z-score (i.e., not expected more than 5% of the time if the population proportions are the same) suggest that a particular map type is more likely to occur in a particular phase of the teleconnection. The test is not performed

if $n\hat{\pi}$ or $n(1-\hat{\pi})$ is less than 5 because the normal approximation to the binomial does not hold for small samples.

3.7 Empirical downscaling methods

As stated in Section 1.2.4., multiple methodologies exist for empirical downscaling. In this dissertation, three methods are developed, applied, and evaluated. These methods are described in the following sections. The results of the downscaling analysis are presented in Chapter 8.

3.7.1 The analog method

The analog method is the simplest of the three empirical methods applied in this study. As the name implies, in the analog method historical observations are examined to find the most similar 'state' (usually defined in terms of the sum of the squared differences over all grid points) to that produced by the GCM. The values of the surface variables are then modeled as those that occurred with the historical analog. While only a few studies have applied the analog method to climate downscaling (Zorita et al. 1995; Cubasch et al. 1996; Zorita and von Storch 1999), it has long been used in weather forecasting and short-term climate prediction (Lorenz 1969; Kruizinga and Murphy 1983; van den Dool 1994).

The analog method requires a long historical record if a suitable analog is to be found for each GCM simulation (van den Dool 1994). Regional focus typically alleviates this problem and useful analogs have been found for most downscaling applications (Zorita and von Storch 1999). Given the simplicity of the analog approach, refinements

have been suggested, including defining the large scale state in terms of multiple variables and including information from the previous day (Cubasch et al. 1996). The specific characteristics of the analog method, as applied in this dissertation, are described in detail in Chapter 8.

3.7.2 Stochastic weather generator

Stochastic weather generators (SWGs) are models of daily weather sequences which can be viewed as statistical representations of the local climate (Wilks 1999). SWGs have been widely used for agricultural (i.e., crop model) applications when observed data is inadequate due to missing or incomplete data (Wilks and Wilby 1999) and for development of regional climate change scenarios (Wilks 1992; Semenov and Barrow 1997). Typically, SWGs are comprised of three components: a precipitation occurrence component, a precipitation amount component, and an additional algorithm for other variables (often conditioned on precipitation occurrence).

3.7.2.1 Precipitation occurrence

Stochastic weather generators simulate precipitation occurrence using either Markov chain models or spell length distributions. In the first approach, in the case of a first order Markov process, the occurrence of precipitation depends on two parameters: p_{01}, the probability of a wet day following a dry day, and p_{11}, the probability of a wet day following a wet day. Depending on the wet/dry status and weather type of the previous day, a uniform [0,1] random number is compared to the appropriate transition probability. If the random number is less than the transition probability, a wet day is simulated.

Otherwise, a dry day is simulated. Wilks (1999) found this type of model correctly reproduced the precipitation occurrence characteristics for stations in the central and eastern USA. For other regions, higher order models have been advocated.

The spell length approach to modeling precipitation occurrence uses distributions fitted to observed lengths of wet and dry spells. Alternating wet and dry spells are then simulated by randomly sampling from the appropriate spell length distributions. Different sequences of wet and dry spells can be simulated depending on the choice of statistical distribution. For example, if a geometric distribution is used, the resulting series will be identical to that produced by a first-order Markov chain model (Wilks and Wilby 1999).

3.7.2.2 Precipitation amount

Once the precipitation occurrence algorithm produces a wet day, a precipitation amount must be drawn from a chosen statistical distribution. The most common choice for precipitation amount simulation has been the gamma distribution (Wilks and Wilby 1999), given by:

$$f_r = \frac{(r/\beta)^{\alpha-1}\exp(-r/\beta)}{\beta\Gamma(\alpha)}; r, \alpha, \beta > 0 \qquad \textbf{3.20}$$

where the distribution parameters are α (the shape parameter) and β (the scale parameter), r is the daily precipitation amount, and Γ is the gamma function, given by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt \qquad \textbf{3.21}$$

The mean wet-day precipitation amount is given by $\mu = \alpha\beta$, and the variance is given by $\sigma^2 = \alpha\beta^2$.

Several studies have indicated that the mixed-exponential distribution provides a better overall fit (and a particularly better fit for large precipitation amounts) than the gamma distribution (Foufoula-Georgiou and Lettenmaier 1987; Wilks 1999). The mixed-exponential distribution is a probability mixture of two single parameter exponential distributions, with probability density function:

$$f_r = \frac{\alpha}{\mu_1}\exp\left[\frac{-r}{\mu_1}\right] + \frac{1-\alpha}{\mu_2}\exp\left[\frac{-r}{\mu_2}\right] \qquad \textbf{3.22}$$

where r is again the daily precipitation amount, with the mean and variance of wet day precipitation amount given by $\alpha\mu_1 + (1\text{-}\alpha)\mu_2$ and $\alpha\mu_1^2 + (1\text{-}\alpha)\mu_2^2 + \alpha(1\text{-}\alpha)(\mu_1\text{-}\mu_2)^2$, respectively.

3.6.2.3 Simulation of other variables

SWGs produce variables other than precipitation with a first order multiple autoregression, first described by Matalas (1967) and given by:

$$X_i = AX_{i-1} + B\varepsilon_i \qquad \textbf{3.23}$$

where X_i is a matrix containing the current day's standardized values of the variables and X_{i-1} is a matrix containing the previous day's standardized values of the variables, ε_i is a vector of independent values from a standard Normal distribution, and A and B are matrices given by

$$A = M_1 M_0^{-1} \qquad \textbf{3.24}$$

$$BB^T = M_0 - M_1 M_0^{-1} M_1^T \qquad \textbf{3.25}$$

where M_o is the matrix of lag-0 cross correlations and M_1 is the matrix of lag-1 cross correlations. While A can be directly computed, B is computed by defining a new matrix

$Z=BB^T$ (see Greene 2000). Then $Z=CLC^T$, where C is the matrix of eigenvectors of BB^T and L has the eigenvalues of BB^T on the diagonal and zeros elsewhere. B can then be computed as $B=CL^{1/2}C^T$.

To produce stationary time series, annual cycles of the daily means and standard deviations of the input variables are constructed using harmonic analysis (usually with separate harmonics for wet and dry days). The time series are then reduced to standardized and stationary residual elements by subtracting the daily means and dividing by the standard deviations, as defined by the harmonics. After generation of the residual series with Eq. 3.23, the daily harmonics described above are used to produce dimensional values of the input variables, based on wet/dry status.

3.7.3 MOS-based downscaling with multiple linear regression

As stated in the introduction, the MOS (model output statistics) approach to empirical downscale is appropriate even if the agreement between large scale observations and GCM simulations is not perfect (see Chapter 6). MOS-based downscaling is applied in this study by using multiple linear regression (MLR) to relate free atmosphere GCM output to surface observations. The MLR model is given by:

$$y = \beta x + \varepsilon \qquad \textbf{3.26}$$

where y is the predictand, x is the matrix of predictors, β are the model coefficients, and ε is the error term.

CHAPTER 4: QUANTITATIVE EVALUATION OF 'OBSERVED' (REANALYSIS) DATA USED IN THE EVALUATION OF GCMs

Schoof JT, Pryor SC. 2003. Evaluation of the NCEP/NCAR reanalysis in terms of synoptic scale phenomena: A case study from the Midwestern USA. *International Journal of Climatology* **23**: 1725-1741.

Abstract

We evaluate the ability of the NCEP/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. While 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 which 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 may also reflect differences in the intensity of synoptic scale phenomena as manifest in the data sets.

Keywords: 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; D'Andrea *et al.*, 1998; Boyle, 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 mid-latitudes 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 dependent on the robustness of observationally based synoptic classifications of current conditions to be used in the evaluation.

To properly classify atmospheric conditions at the synoptic-scale, long-term data records are needed from stations, or grid points, throughout the region of interest. The NCEP/NCAR reanalysis 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 NCEP/NCAR reanalysis 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 NCEP/NCAR reanalysis. This research is warranted because although the NCEP/NCAR data set has been extensively evaluated 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., Swail and Cox, 2000; Hines *et al.,* 2000; Marshal, 2000; Trenberth and Stepanik, 2002; Renfrew *et al*., 2002). Also, previous analysis of the data set has typically focused on individual and co-located variables rather than ensembles representing fundamental features/scales and inter-relationships of the data set variables.

2 Data / Study Area

The primary objective of this study is to evaluate the ability of the NCEP/NCAR reanalysis (hereafter referred to as 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 thirty-five grid points (2.5° × 2.5° spacing) are available from the following two databases, respectively:

1) The Radiosonde Data of North America (1946-1996) database issued by the Forecast Systems Laboratory and National Climatic Data Center (FSL/NCDC, 1997). This dataset is hereafter referred to as RSONDE.

2) The NCEP/NCAR 40-year Reanalysis data series (Kalnay *et al*., 1996), hereafter referred to as NCEPR.

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

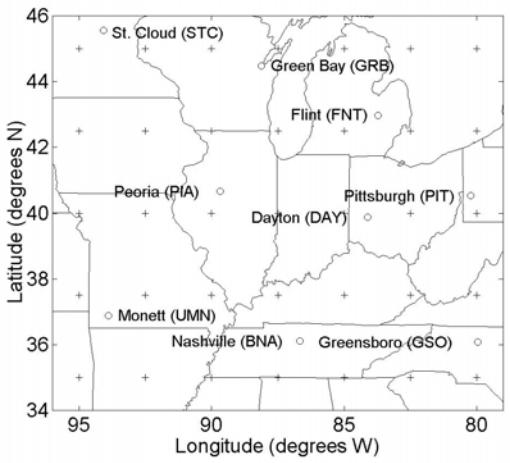

Figure 1: Map of study area showing the 35 NCEP/NCAR reanalysis grid points ('+') and the 9-station rawinsonde network ('○', station name included in figure).

Note that while 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 to broadly represent circulation patterns, thermal advection, and jet stream strength, respectively. The classification presented here is for the summers of 1971-1993, where summer is defined following Schoof and Pryor (2001) as year-day 100 (April 10) to year-day 300 (October 27). 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. While 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, while there are small differences in the standard deviations of the input variables (Figure 2).

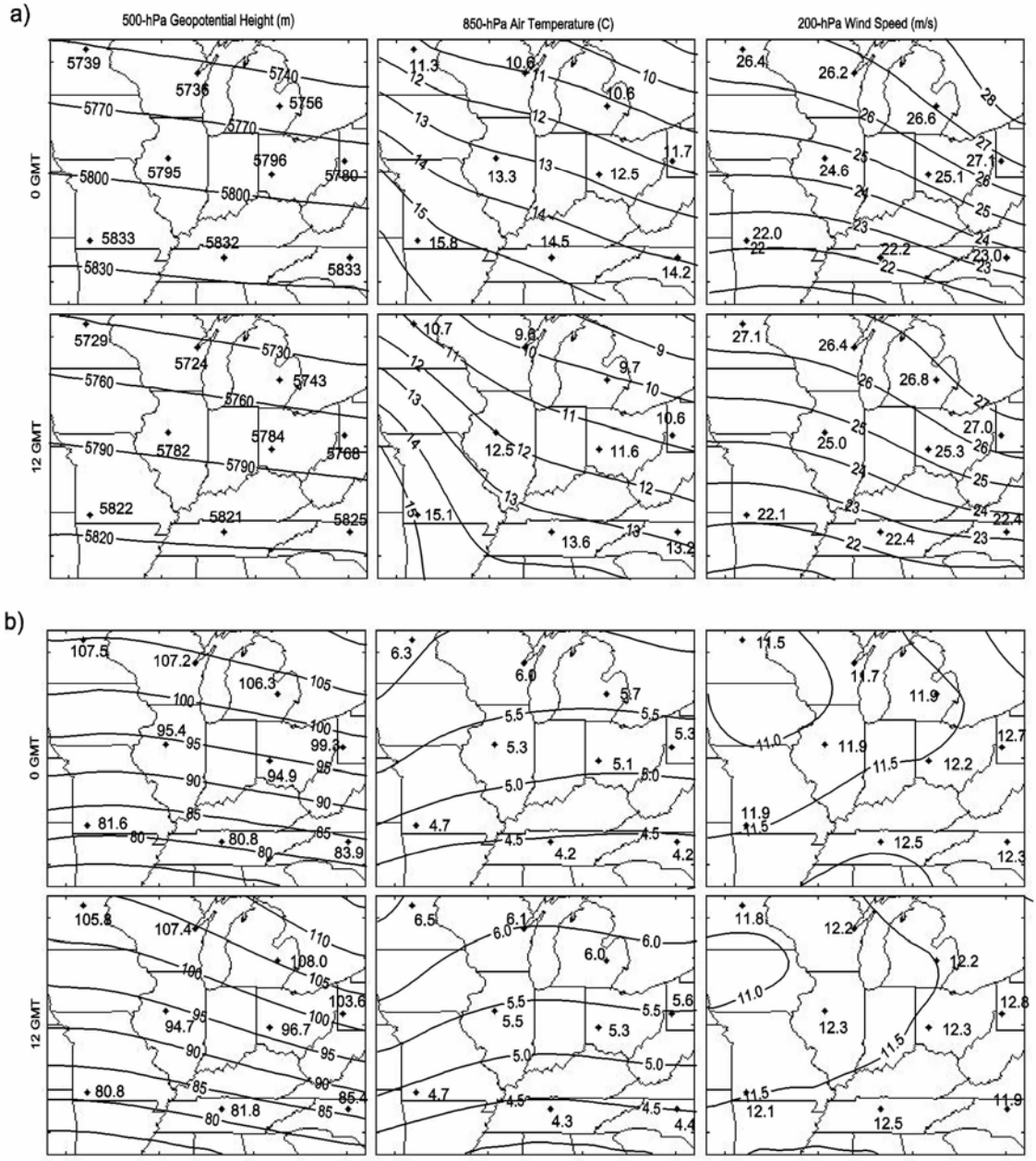

Figure 2: (a) Means and (b) standard deviations of RSONDE and NCEPR input data at 0 GMT and 12 GMT. In each plot, the contours represent NCEPR values. RSONDE values are shown in text adjacent to each station.

3 Methods

There are several methods available for the comparison of weather-type classifications (Huth, 2000). First, classifications can be applied to two datasets independently, allowing the dominant weather-types from both datasets to be identified. The second method of comparing classifications requires that the classes are predetermined, which is often difficult in climatological studies. While 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 dataset, followed by projection of the results onto the second dataset. This method classifies the second dataset 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 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 evenly weight the input variables, 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 re-classification 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 analyzed 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. While observational errors cannot be directly quantified (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 (0.5ºC 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 analysis 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 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 to 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 principal components 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, 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, while to compare the two PCA solutions, Pearson product-moment correlation coefficients (r) between RSONDE PCs and NCEPR PCs are computed.

As shown in Table 1, each of the PCs from the NCEPR solution shows a high degree of correspondence (|r| of PC scores which 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 1 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 behavior opposite 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 ($r = -0.66$). While 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 NE part of the study area and NCEPR PC8 loads strongly on wind speeds in the SE 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, while NCEPR 7 and 9 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.

Table 1: Pearson product-moment correlations between the NCEPR and RSONDE PC scores. Entries marked with '*' are significant at the 95% confidence level. Entries marked with '**' are significant at the 99% level.

	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**

4.1.2 Synoptic types resulting from the cluster analysis

Principal component scores represent the relationship between each PC and the individual observations. Therefore, clustering the PC scores will yield groups that contain similar 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 re-classification 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 one 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 chosen solution corresponds to large increases in the variance of the first few PCs.

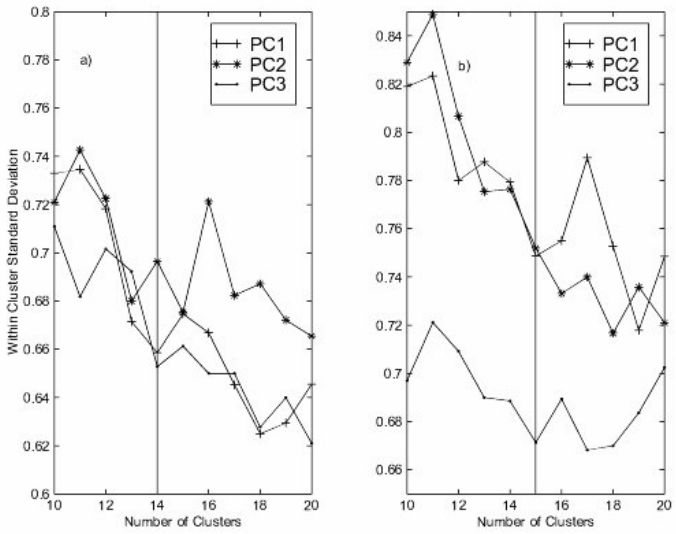

Figure 3: Clustering 'cut-off' diagrams for (a) RSONDE and (b) NCEPR. The vertical lines represent the chosen clustering level, consistent with the increases in the standard deviations of the leading PC scores.

Table 2: Description of RSONDE weather-type classification. The table shows the class size in % (column 2), class size in number of member observations (column 3), relation to principal components (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 (high or low anomalies and regions are shown with a + or -, respectively, in columns 10-12), and seasonal variability (percentage of observations in the bin specified by the day number at the top of columns 13-16).

	Size		PC characteristics						Input data characteristics			Seasonal variability (Year day)			
Class	Size (%)	Size (n)	PC1	PC2	PC3	PC4	PC5	PC6	T850	HT500	WS200	100-150	151-200	201-250	251-300
1	5.6	113	-	+			-		-	-	+(E)	30.1	10.6	15.0	44.3
2	14.6	297	+	-		-			+	+	-	5.1	33.0	47.8	14.1
3	7.1	144		-	-	+	-	+	+	+	-(E), +(W)	19.4	18.6	13.2	48.6
4	9.8	200		+			+	-	+(W)		+(E), -(W)	10.5	42.0	40.0	7.5
5	6.7	137	+	+	+	+		+				19.7	11.7	21.2	47.4
6	7.2	146		+		-					+(E), -(W)	14.4	17.1	49.3	19.2
7	12.5	254				+		-			+(NW), -(SW)	4.5	31.7	54.4	9.4
8	6.2	126			+		-		-	-		42.0	2.4	1.6	54.0
9	5.1	103	-		+	-		-	-	-	-(W)	88.3	7.8	0	3.9
10	6.4	131			+		+		-(E), +(W)	-		35.9	30.5	10.7	22.9
11	6.7	136	-	+	-		+	+	-(E), +(W)	-	+(E)	37.5	31.6	14.7	16.2
12	12.1	247		-	-				+	+	-(E)	27.5	35.6	23.9	13.0

Table 3: Description of NCEPR weather-type classification. The table shows the class size in % (column 2), class size in number of member observations (column 3), relation to principal components (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 (high or low anomalies and regions are shown with a + or -, respectively, in columns 10-12), and seasonal variability (percentage of observations in the bin specified by the day number at the top of columns 13-16).

	Class Size		PC Characteristics									Input data characteristics			Seasonal variability (Year day)			
Class	Size (%)	Size (n)	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	T850	HT500	WS200	100-150	151-200	201-250	251-300
1	5.5	112			+						-	-	-		35.7	2.7	0.9	60.7
2	7.2	146	+	+		+					+	+(W)	+	+	3.4	24.0	42.5	30.1
3	7.8	158							-	+	+	+(SE)		+(SE)	5.7	45.6	42.4	6.3
4	3.8	77	-		+		-	+		-	+	-	-	-(E)	68.8	7.8	0	23.4
5	8.6	174		+	-	-	+	-			+	+	+(E)	-(S), +(N)	5.8	28.2	54.0	12.1
6	8.3	169				-	+		-		-	+(SE)		-	16.0	29.6	33.1	21.3
7	3.9	80	-	-	-	+	+	-	+		+	-(W), -(SE)	-	+	21.3	26.3	21.3	31.3
8	5.8	11.8	-	-		-			+	+		-(W), -(SE)	-	-	38.1	11.9	17.8	32.2
9	6.0	121	-			+		+		+	-	-(SW), +(NW)	-	+	49.6	20.7	2.5	27.2
10	9.0	183		+	-	-	-				-	+	+(E)	-	15.3	30.1	28.4	26.2
11	3.9	80	-		+			-	-			-	-		72.5	2.5	1.3	23.8
12	7.0	142	-	-			-			-		-	-(W)	-(NE)	56.3	23.9	8.5	11.3
13	4.5	92		+	-	+			+	-	-	+	+	+	15.2	15.2	10.9	58.7
14	7.6	154	+	-				+			+	+	+(W)	-(NE)	5.8	38.3	42.9	13.0
15	11.1	228	+	-			-					+	+(W)	-(NE)	4.8	33.3	50.0	11.8

The 14-class RSONDE and 15-class NCEPR solutions are described in detail in Tables 2 and 3, 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).

While Tables 2 and 3 highlight some of the characteristics of the individual classifications, the similarities and differences between the classifications are of primary interest in this study. Table 4 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 4) (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 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 2 and 3 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 2 and 3). 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 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.

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, 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, are shown in Table 5.

The application of this third classification slightly improves the agreement between the NCEPR classification and that derived from the RSONDE data (Table 6). 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 RSDONE 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, for example; 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). While 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, while there is some similarity between the two solutions, a solution approaching one-to-one correspondence is not obtained. This result suggests that while 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.

Table 4: 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, while small numbers indicate few shared observations.

RSONDE ⇒ NCEPR ⇓	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

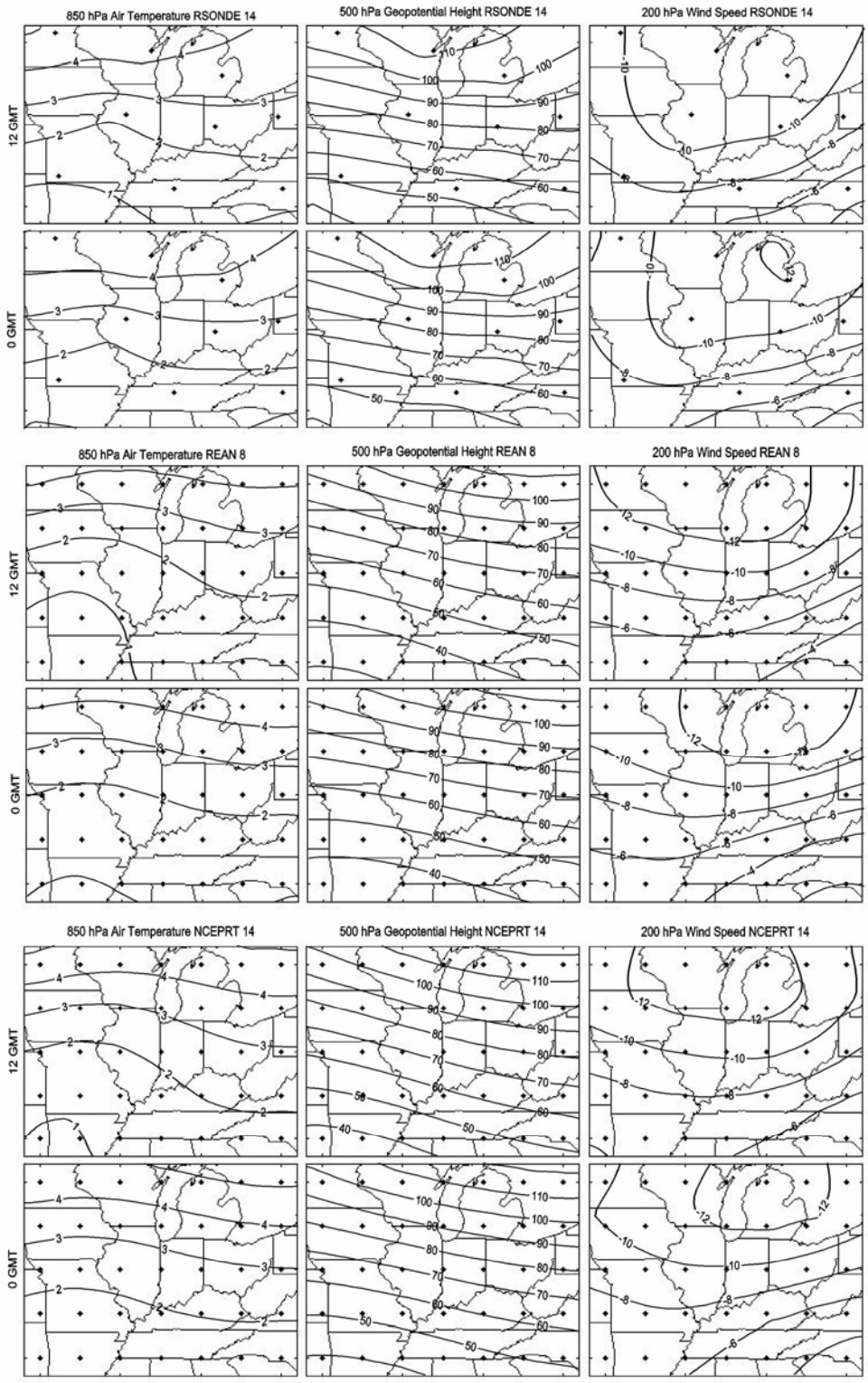

Figure 4: Input data anomalies (deviations from the mean of all observations) for RSONDE 14, NCEPR 8, and NCEPRT 14. The figures provide an example of good agreement between the classifications. Each plot shows the mean anomaly for each input variable for observations assigned to this class. The contours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

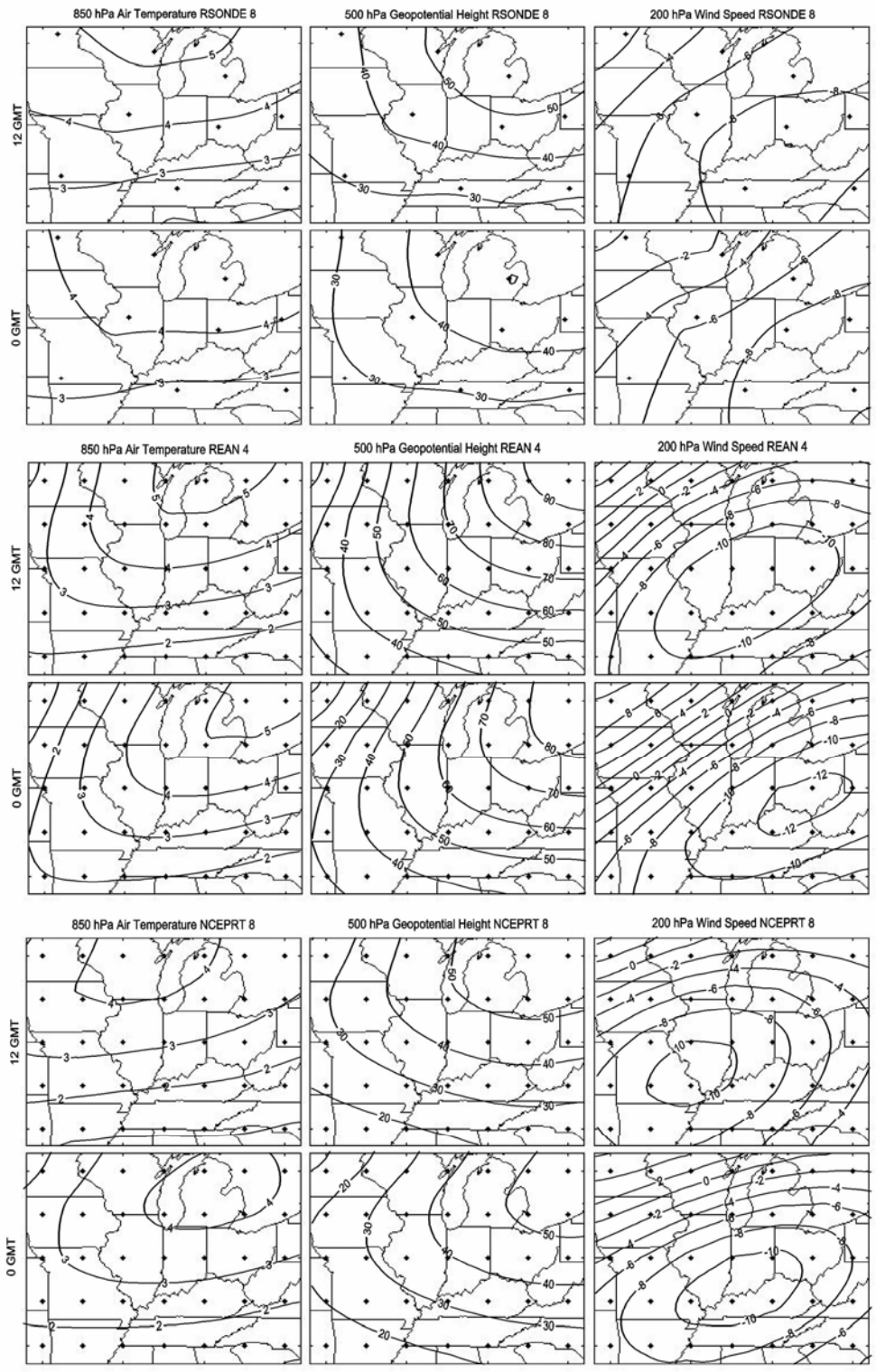

Figure 5: Input data anomalies (deviations from the mean of all observations) for RSONDE 8, NCEPR 4, and NCEPRT 8. The figures provide an example of poor agreement between the classifications. Each plot shows the mean anomaly for each input variable for observations assigned to this class. The contours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

Table 5: Description of NCEPRT weather-type classification. The table shows the class size in % (column 2), class size in number of member observations (column 3), relation to principal components (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 (high or low anomalies and regions are shown with a + or -, respectively, in columns 10-12), and seasonal variability (percentage of observations in the bin specified by the day number at the top of columns 13-16).

	Class size		PC characteristics									Input data characteristics			Seasonal variability (Year day)			
Class	Size (%)	Size (n)	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	T850	HT500	WS200	100-150	151-200	201-250	251-300
1	7.0	142	-				+	-	+			-	-	+	15.5	24.7	41.6	18.3
2	12.2	250	+	-			-					+E	+	-	4.8	32.8	48.8	12.0
3	6.2	125		+	-	+			+	-	-	+	+E	+NW, -SE	16.8	14.4	10.4	58.4
4	11.0	223				-	+		-	+		+SW		-W, +E	15.7	33.2	32.3	18.8
5	11.4	232	+	+	+	+				+					3.02	37.9	48.7	10.3
6	7.0	143		-					+	+		-SW		-NW, +SE	36.4	12.6	15.4	35.7
7	9.7	197		+		-	+							+N	5.1	30.0	50.3	14.7
8	6.1	123			+			-				-	-E		31.7	4.9	2.4	61.0
9	5.1	103	-	-	+	-	-					-	-	+SW, -N	62.1	9.7	6.8	21.4
10	5.0	102			+			+			+	-	-E		68.6	5.9	0	25.5
11	6.8	138	-		-	+		+		+					47.1	19.6	2.2	31.2
12	12.5	256			-							+	+E		28.1	31.6	25.0	15.2

Table 6: 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, while small numbers indicate few shared observations.

RSONDE ⇒ NCEPRT ⇓	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

5 Discussion

In Section 4, we showed that although classifications based on radiosonde data and the NCEP/NCAR reanalysis 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:

- Within-type variability as a result of both grid-point density and classification of continuous data into discrete classes (e.g. Yarnal (1993) and Brinkmann (1999b)). The NCEPR data is on a 2.5° × 2.5° grid, but the radiosonde data used in this study was 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), 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 2.5º × 2.5º box, while 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 NCEP/NCAR reanalysis. While 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 NCEP/NCAR reanalysis, 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 NCEP/NCAR reanalysis.
- 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-1980, 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 2, 3, and 5 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, 2002) 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, 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 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 NCEP/NCAR reanalysis 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 principal component analysis (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. While 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 analyze 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 NCEP/NCAR reanalysis 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), U.S. 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. The authors gratefully acknowledge the constructive criticisms of two anonymous reviewers.

References Cited

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 Ven Te Chow (ed.), *Handbook of applied hydrology, a compendium of water-resources technology*: New York, McGraw-Hill Book Company, p 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 (accessed 8/28/03).

Hines KM, Bromwich DH, Marshall GJ. 2001. 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 procedure 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 50 year 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 (Office of the Federal Coordinator for Meteorology). 1997. Federal Meteorological Handbook No. 3 – Rawinsonde and Pibal Observations, 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. 2002. Discarding or downweighting high-noise variables in factor analytic models. *Analytica Chimica Acta* (submitted).

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.

CHAPTER 5: IDENTIFICATION OF DOWNSCALING PREDICTOR VARIABLES

5.1 Introduction

Empirical downscaling studies have used a wide range of free atmosphere predictor variables, although there has been little systematic work evaluating the relative skill associated with different predictors (Winkler et al. 1997; Giorgi et al. 2001) in concert with the validity of their simulation by GCMs. Because the relationships between free atmosphere and surface variables vary temporally and spatially, it is difficult to make generalizations about predictor-predictand relationships in an empirical downscaling framework. Nevertheless, development of systematic approaches to predictor selection and evaluation are necessary pre-requisites to providing a framework for quantitative inter-comparison of downscaling studies. Hence, in this dissertation, predictor variables are evaluated in terms of i) their relationships with the surface variables of interest (this chapter), and ii) their representation with the chosen GCMs simulations (Chapters 6 and 7).

5.2 Previous studies

For both temperature and precipitation downscaling, most studies have used circulation parameters (e.g., geopotential height, sea level pressure) as predictor variables (e.g., Bardossy and Plate 1992; Heyen et al. 1996; Corte-Real et al. 1998; Crane and Hewitson 1998; Kilsby et al. 1998). However, changes in surface climate elements are not attributable solely to circulation changes (Yarnal 1985; Huth 1999). For example, low frequency changes in 20th century central European surface air temperature are only

partially explained by changes in circulation (Werner and von Storch 1993) and changes in surface temperature under increased atmospheric carbon dioxide concentrations may be largely controlled by changes in the radiative properties of the atmosphere, which can be incorporated into downscaling studies by including GCM-simulated upper air temperature fields (Schubert 1998). Hewitson and Crane (1996) demonstrate that inclusion of atmospheric humidity drastically alters the outcome of precipitation downscaling, although it may not appear to be an important predictor in terms of variance explanation in the historical record. Downscaling methodologies which systematically choose predictors based on variance explanation may therefore exclude variables which carry the climate change signal (Giorgi et al. 2001).

In a central European air temperature downscaling application, Huth (2002) found that large-scale free atmospheric temperatures are better predictors of mean daily surface temperature than large-scale circulation fields, while the combinations of circulation and temperature produced the best results. For precipitation downscaling, several studies have identified the importance of vorticity and other air flow indices derived from 500-hPa geopotential height or sea-level pressure fields (e.g., Conway et al. 1996; Goodess and Palutikof 1998; Kilsby et al. 1998; Wilby 1998). While Cayan and Roads (1984) found that vorticity was a poor predictor in an application to area averaged precipitation along the West coast of the USA, it is physically linked to divergence, and hence vertical motion, and therefore, has a physical link to precipitation processes. Furthermore, vorticity has been show to exert a strong influence on both the probability and magnitude of single site precipitation in the British Isles (Wilby et al. 1998) and is therefore included

as a potential downscaling predictor in this study (see Chapter 3 for derivation of vorticity variables).

Based on previous research (see Table 5.1), in addition to the vorticity variables, 500 hPa geopotential height (Z500) and sea-level pressure (SLP) are used in this study for representation of the large-scale circulation. Other upper-air predictors are derived from the 850-hPa level, which is often above the boundary layer and therefore is the lowest level in the GCMs for which simulations should be useful and consistent (Sailor and Li 1999).

While many downscaling studies simply choose a variable and downscale using GCM output from the nearest grid point, spatially remote points have been demonstrated to increase the proportion of predictand variance explained (Wilby and Wigley 2000; Brinkmann 2002). For example, it has been shown that daily station-scale precipitation in the eastern USA is most highly correlated with geopotential heights some distance west of the station location, associated with a short-wave trough (Stidd 1954; Klein 1963; Brinkmann 2002), while station-scale air temperature in the Great Lakes region has been shown to be most highly correlated with geopotential height to the east and southeast of the station location consistent with the advection of warm air (Brinkmann 1999; Brinkmann 2000).

The widespread availability of reanalysis data has resulted in tremendous growth in the number of potential large-scale downscaling predictors in recent years. In this chapter, an array of large scale variables from reanalyses (Table 5.1) are evaluated in terms of their relationships with surface temperature and precipitation at the 84 stations listed in Table 2.1. Each of these variables has been previously used in the empirical

downscaling framework (Table 5.1). The results will be used with GCM evaluation results (Chapters 6 and 7) to make decisions regarding the inclusion of particular variables for GCM downscaling.

Table 5.1. Predictor variables and examples of their use in previous downscaling studies. Predictands are shown in parentheses (T=surface air temperature, P=surface precipitation).

Predictor	Abbrev.	Previous studies
Upper-atmosphere variables		
850 hPa geopotential height (m)	Z_{850}	Sailor and Li 1999 (T, P)
500 hPa geopotential height (m)	Z_{500}	Hewitson and Crane 1996 (P) Weichert and Burger 1998 (T, P, vapor pressure) Sailor and Li 1999 (T, P) Palutikof et al. 2002 (T, P)
850-500 hPa thickness (m)	THICK	Cavazos 1997 (P) Kidson and Thompson 1998 (T, P) Cavazos 1999 (P)
850 hPa specific humidity (kg/kg)	Q_{850}	Crane and Hewitson 1998 (P) (uses Q at 1000 and 500 hPa) Murphy 1999 (T, P)
850 hPa relative humidity (%)	RH_{850}	Sailor and Li 1999 (T, P)
Surface variables		
Mean sea-level pressure (hPa)	SLP	Zorita et al. 1995 (P) Hewitson and Crane 1996 (P) Heyen et al. 1996 (sea level) Schubert 1998 (T_{max}, T_{min}) Palutikof et al. 2002 (T, P)
Zonal component of geostrophic flow (hPa /10° latitude at grid point latitude)	GEOW	Buishand and Brandsma 1997 (T, P)
Meridional component of geostrophic flow (hPa /10° latitude at grid point latitude)	GEOS	Wilby and Wigley 1997 (P) Kidson and Thompson 1998 (T, P)
Strength of the resultant geostrophic flow (hPa /10° latitude at grid point latitude)	GEOWS	Wilby 1998 (T, P)
Westerly shear vorticity (hPa /10° latitude at grid point latitude, per 10° latitude)	GEOZW	
Southerly shear vorticity (hPa /10° latitude at grid point latitude, per 10° latitude)	GEOZS	
Total shear vorticity (hPa /10° latitude at grid point latitude, per 10° latitude)	GEOZT	

5.3 Strategy for selection of downscaling predictors

The selection of downscaling predictors begins with examination of the relationships between the surface variable of interest and observed free atmosphere predictors. As is typical in downscaling studies, the focus is on maximizing the amount of the variance of the surface variables explained by individual free atmosphere predictors or combinations of predictors. While this practice raises the concern of overfitting statistical models to the current climate and it is generally accepted that models which exhibit the best agreement with observations are the optimal models for downscaling transient GCM output, there is no guarantee that observed predictor-predictand relationships will be valid under changed climate conditions.

Based on research indicating the value of non-local grid points, surface temperature and precipitation are examined in terms of free atmosphere variables over a large domain (Figure 2.1) to examine the improvement achieved through use of remote grid points. It is also anticipated that the evolution (or differences between daily values) of the predictor variables may also be of particular value in identifying high frequency changes responsible for changes in surface parameters (e.g., frontal passages). For this reason, each of the predictor variables are also examined in terms of one- and two-day differenced values and one- and two-day lagged values. The approach described above results in a total of 60 potential predictor variables (i.e., the 12 variables in Table 5.1 plus their lagged and differenced variations). For each predictor, the relationships with surface variables are assessed on a seasonal basis to minimize the effect of the seasonal cycle on the correlations and to examine potential seasonal variations. It should be noted that this approach does not allow examination of evolution of the meteorological seasons

and that the analyses presented here are intended only for identification of statistical associations. The downscaling approaches described in Chapter 8 (with the exception of the stochastic weather generator) do not have seasonal formulations. Correlations between the predictors and station precipitation amounts are examined for the entire time series and for wet days only. Because the correlations are nearly identical in each case, only the results for the complete time series are presented.

Throughout this chapter, correlations are described in terms of their magnitude and statistical significance. In cases where significance of correlations is discussed, adjustments to the number of degrees of freedom have been made to account for temporal autocorrelation using the method of Dawdy and Matalas (1964). However, it should be noted that the sample sizes in this study (approximately 3000 observations for each season) result in very small correlations attaining statistical significance (typically $|r| \geq 0.06$ is significant at the 95% level). For visualization purposes, the station and grid point values have been contoured using bilinear interpolation to a fine grid as shown in the figures presented in the following sections.

5.4 Scales of variability

Prior to examining the correlations between individual grid point predictors and the station values of T_{max}, T_{min} and precipitation, it is important to consider the spatial scales at which surface temperature and precipitation vary. To this end, the spatial variability of the surface variables is examined by computing the spatial autocorrelation between values at all stations for each surface variable during each season. When plotted

as a function of distance (Figure 5.1), autocorrelation statistics provide insight into the spatial coherence of the surface variables and how that coherence may vary seasonally.

As shown in Figure 5.1, T_{max} and T_{min} are both more spatially coherent during the transition seasons than during winter and summer, as evidenced by larger autocorrelation values. For both T_{max} and T_{min}, the lowest degree of spatial coherence occurs during summer. Higher correlations during the transitions seasons are likely due to common non-stationarities among the stations (e.g., all stations exhibit warming during spring and cooling during autumn). Another inference that can be made from Figure 5.1 is that surface air temperature is more spatially coherent than surface precipitation. During each season, correlations between station precipitation values are lower than those for the temperature variables and, for stations separated by distance greater than 400 km, precipitation correlations are near-zero, particularly during the summer. Previous studies (e.g., Kidson and Thompson 1998; Schoof and Pryor 2001) have found that temperature downscaling results are superior to those for precipitation. The larger degree of spatial variability (relative to GCM grid size) may provide a partial explanation for those findings.

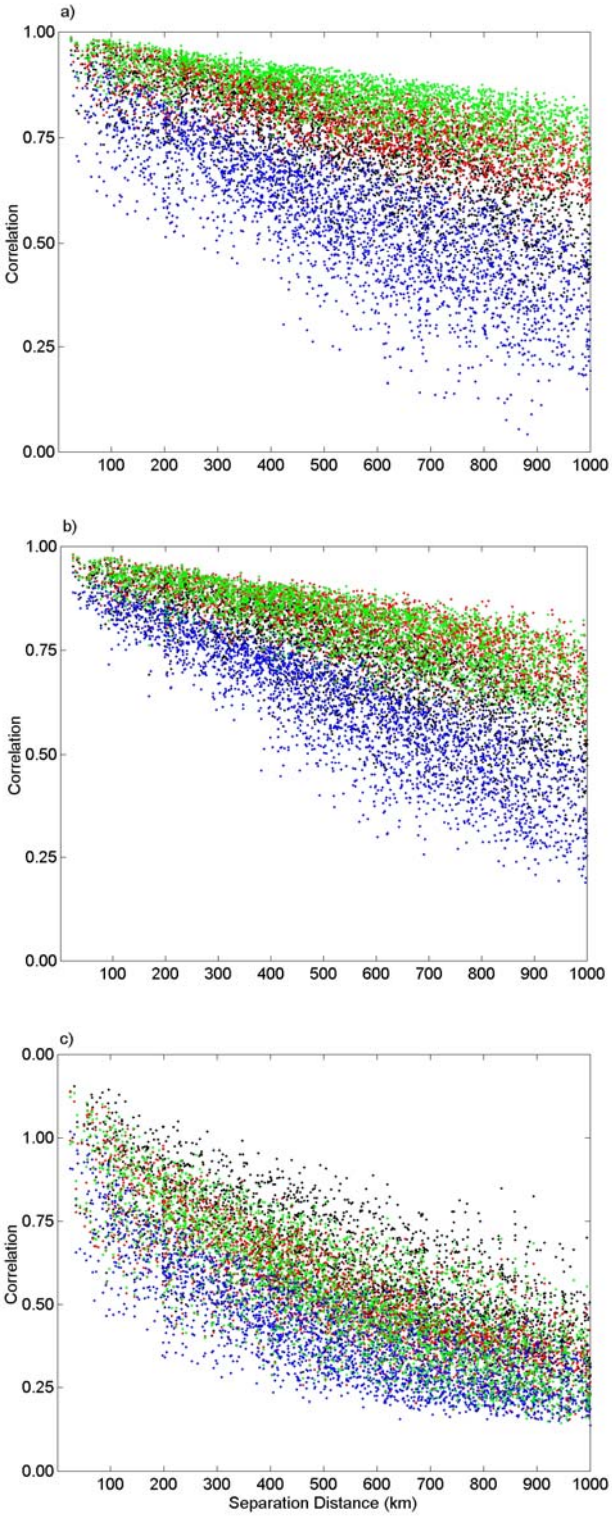

Figure 5.1 Correlation between surface variables as a function of distance: a) T_{max}, b) T_{min}, and c) precipitation. Points are shown for winter (•), spring (•), summer (•), and autumn (•).

5.5 Results

5.5.1 Upper-air variables

5.5.1.1 500-hPa geopotential height (Z_{500})

Grid point 500-hPa geopotential heights (Z_{500}) are highly correlated with both T_{max} and T_{min} in the study region during all seasons, although the correlations are stronger and the area of statistically significant correlations is larger during spring and autumn relative to winter and summer (see, for example, Figures 5.2 and 5.3) consistent with the results presented in Section 5.4. During spring and autumn, the correlations are positive and significant for all points in the domain. During the winter, negative correlations (as strong as -0.3) are found in the northwest and southeast corners of the study area.

Figure 5.2 shows the correlation between Z_{500} at the reanalysis grid points and the surface variables at a central station (Anderson, IN) and indicates that the correlations for the temperature variables are maximized slightly north and east of the station location. For precipitation, the correlations are substantially lower and generally positive east of the station location and negative west of the station location, with the weakest relationship during summer. This pattern of correlations implies that, in each season, precipitation at Anderson is most likely under conditions characterized by a long-wave trough to the west and a ridge to the east of the station location. This pattern is thus conducive to advection of moisture north from the Gulf of Mexico and a subsequent increase in precipitable water.

Figure 5.3 summarizes the maximum correlation (positive or negative) between the Z_{500} field and surface variables at each station and indicates that the correlations between Z_{500} and T_{max} and T_{min} are consistently high across all stations. For some

stations and seasons, the correlation between Z_{500} and precipitation is strongest for grid points west of the station (and therefore negative), while for others the correlation is strongest for grid points east of the station (and therefore positive). Correlations between Z_{500} and precipitation are low during the summer, relative to the other seasons. These results are in accord with a priori expectations that Z_{500} is not a particularly strong predictor of precipitation, which is expected to be governed more by moisture and vorticity advection into the station region.

At lag-1 and lag-2, correlations between Z_{500} and the surface temperature variables are very similar to those at lag-0, although the maximum values are shifted slightly to the west. For precipitation, lag-1 and lag-2 correlations are lower than those at lag-0. Correlations between differenced values of Z_{500} and the surface variables are lower than those at lag-0, lag-1, and lag-2, although values are still statistically significant at most stations.

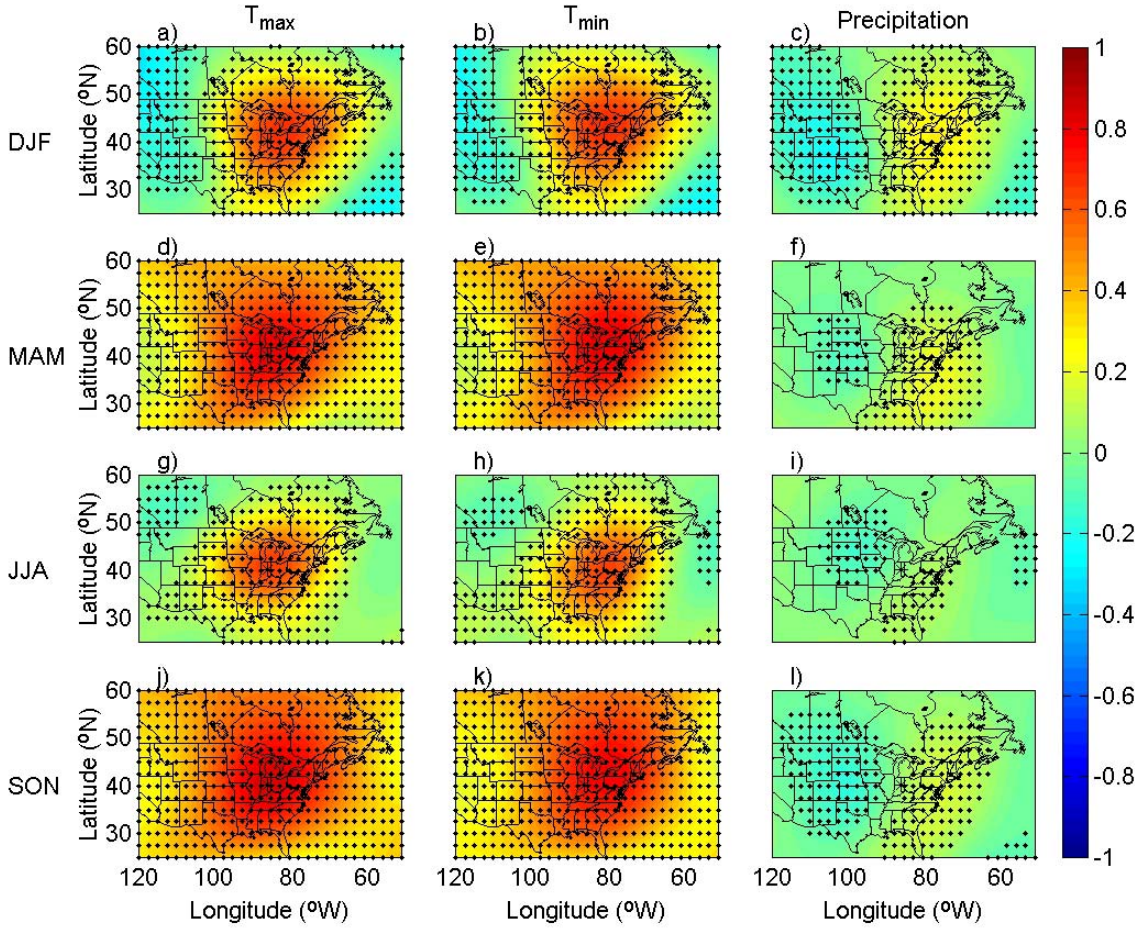

Figure 5.2 Correlation between the daily 500-hPa geopotential height field and surface variables at Anderson, IN (+): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

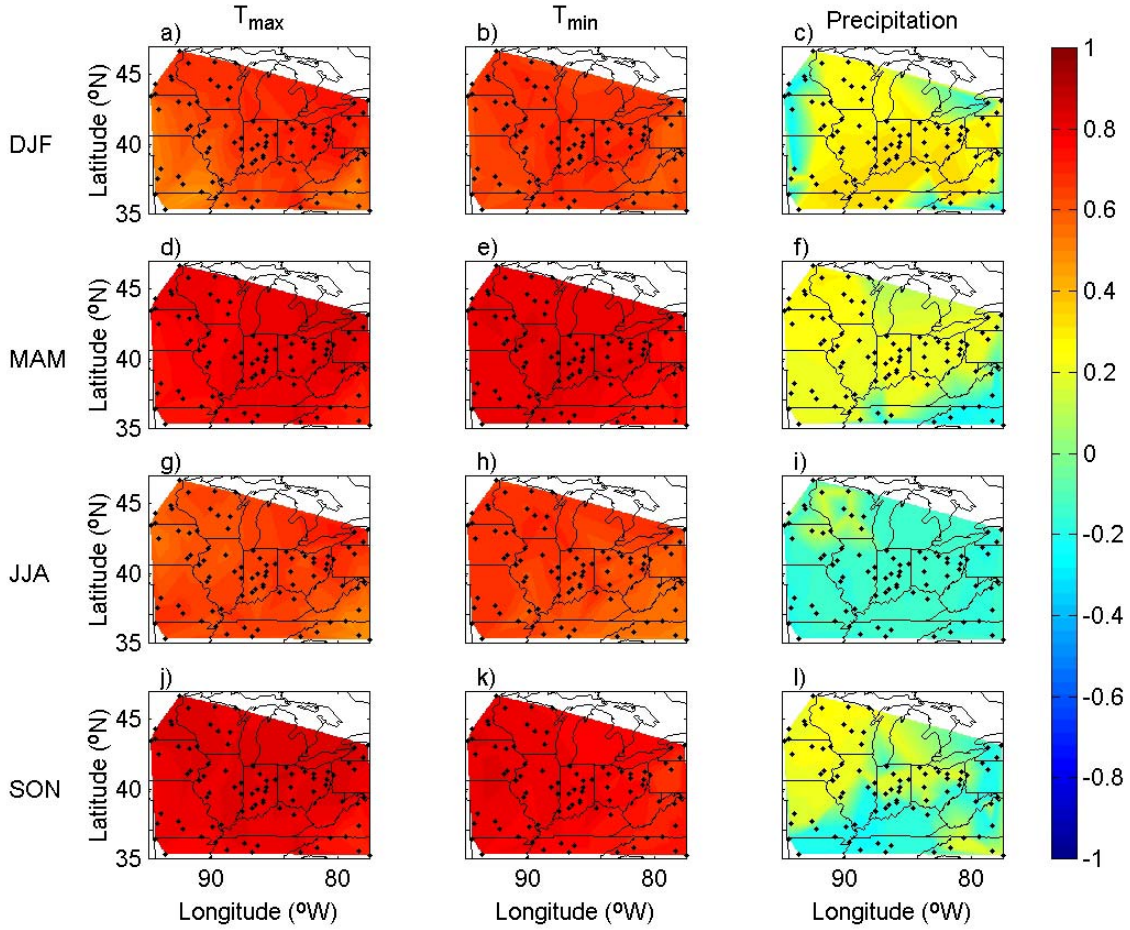

Figure 5.3 Maximum correlation (of positive or negative sign) between the daily 500-hPa geopotential height field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.1.2 850-hPa geopotential height (Z_{850})

Plots of the correlation between 850-hPa geopotential height (Z_{850}) and the surface variables exhibit patterns similar to those for Z_{500} described in Section 5.5.1.1, although the correlations with temperature are much weaker with 850 hPa geopotential height. Given the weaker correlations between Z_{850} and the temperature variables

(relative to Z_{500}), a greater prevalence of negative correlations is evident in Figure 5.4 relative to Figure 5.2. For precipitation, correlations with Z_{850} are slightly lower than those for Z_{500}. The spatial pattern of precipitation correlations shown in Figure 5.4 is consistent with increased precipitation in the presence of a shortwave trough west of the station location, as first suggested by Brinkmann 2002).

Figure 5.5 shows the maximum correlation (positive or negative) between the Z_{850} field and surface variables at each station and indicates that the correlations between Z_{850} and T_{max} and T_{min} are not as strong or consistent across stations as those for Z_{500}. For stations in the western part of the study area during winter, the negative correlations with temperature found west of the station location are stronger than the positive correlations located east of the station location (Figure 5.5). For precipitation, the correlation is strongest for grid points west of the station (and therefore negative), with the exception of a few stations during winter and summer. During spring and summer, correlations between Z_{850} and precipitation are weaker, yet still statistically significant, at all stations.

Correlations between lag-1 Z_{850} and the surface temperature variables are similar to the lag-1 values during spring, summer, and fall. During winter, however, the negative correlations with both T_{max} and T_{min} cover a much larger area and are larger in magnitude. At lag-2, the strongest winter correlations between Z_{850} and T_{max} and T_{min} are negative for almost all stations, while those for other seasons are similar to lag-0. Precipitation correlations are weaker at lag-1 and lag-2 than at lag-0. Correlations between differenced values of Z_{850} and T_{max} and T_{min} are lower than lag-0 values. For precipitation, the correlations are of similar or slightly greater magnitude.

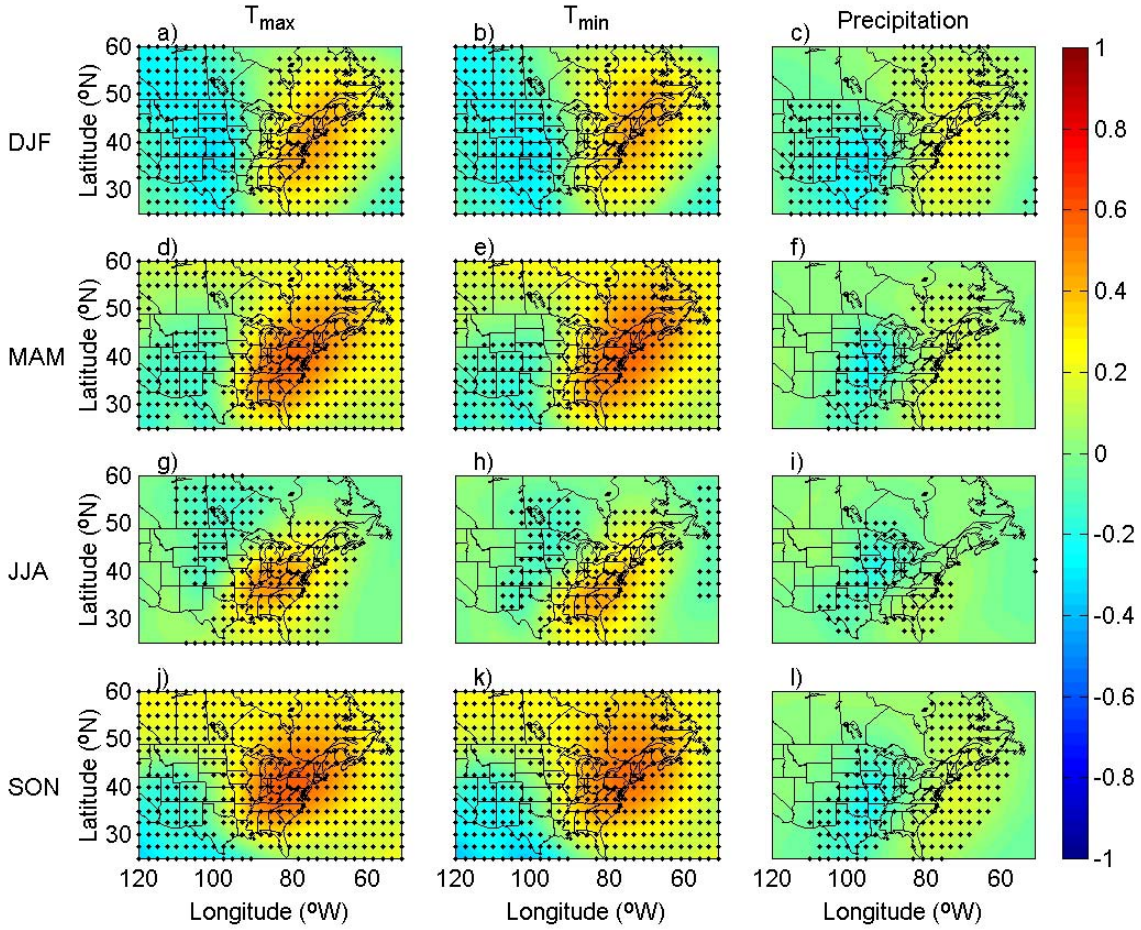

Figure 5.4 Correlation between the daily 850-hPa geopotential height field and surface variables at Anderson, IN (+): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

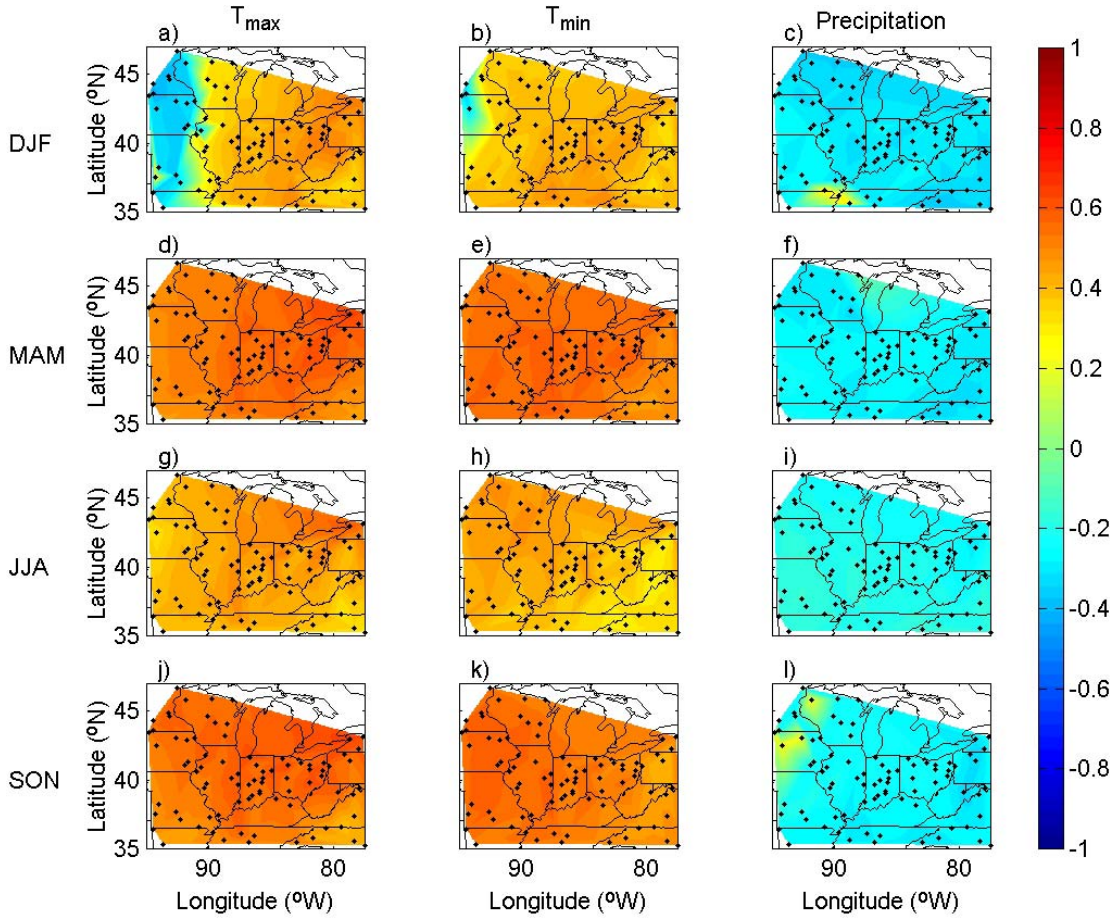

Figure 5.5 Maximum correlation (of positive or negative sign) between the daily 850-hPa geopotential height field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.1.3 850-500 hPa layer thickness (THICK)

The thickness of the atmosphere, from lower to mid levels is often used as a proxy for integrated layer temperature. Defined simply as the difference in the height of two pressure surfaces (in this case 500- and 850-hPa), the variable is expected to exhibit behavior similar to that of Z_{500}, although addition information is also provided by thickness. For example, 1000-500 hPa thickness can be used to estimate the transition

between liquid and frozen precipitation, and the likelihood of extreme freezing or extreme heat events at the surface (Chaston 1997). Here, I use the 850-500 hPa thickness because 1000-hPa geopotential height is not available from both GCMs. As shown in Figures 5.6 and 5.7, the correlations between 850-500 hPa layer thickness (THICK) and the surface variables are indeed very similar to those for Z_{500} as presented in Section 5.5.1.1. However, for all three surface variables, the correlations with THICK are slightly higher than those for Z_{500}. For T_{max} (and T_{min} for most stations), THICK is the variable with the highest correlation for each station in each season, with correlations higher than 0.8 at most stations during spring and autumn and higher than 0.60 at most stations during the winter and summer (Figure 5.7). As with Z_{500}, the correlations with T_{max} and T_{min} are maximized at locations slightly north and east of the station location with positive correlations that are larger and more extensive (spatially) during the transition seasons relative to summer and winter (Figures 5.6 and 5.7). Correlations between THICK and station precipitation are of similar magnitude to those for Z_{500}, with positive correlations during the winter and spring, mostly negative correlations during summer, and a mixture of positive and negative correlations during autumn (Figure 5.7c, f, i, and l).

Correlations between lag-1 THICK and the surface temperature variables are similar to the lag-1 values, with lag-1 correlations exceeding lag-0 correlations at about one third of the USHCN/D stations. At lag-2, the correlations are lower, but still greater than 0.50 at most stations during spring and autumn and greater than 0.40 at most stations during the summer and winter. Precipitation correlations are weaker at lag-1 than at lag-0 during the winter, but similar in magnitude and pattern in other seasons. At lag-2,

correlations with precipitation are lower than at lag-0 and lag-1 with most values less than 0.20. Although non-trivial, correlations between differenced THICK values and the surface variables are lower than those at lag-0, lag-1, and lag-2 and the area covered by statistically significant correlations is much smaller.

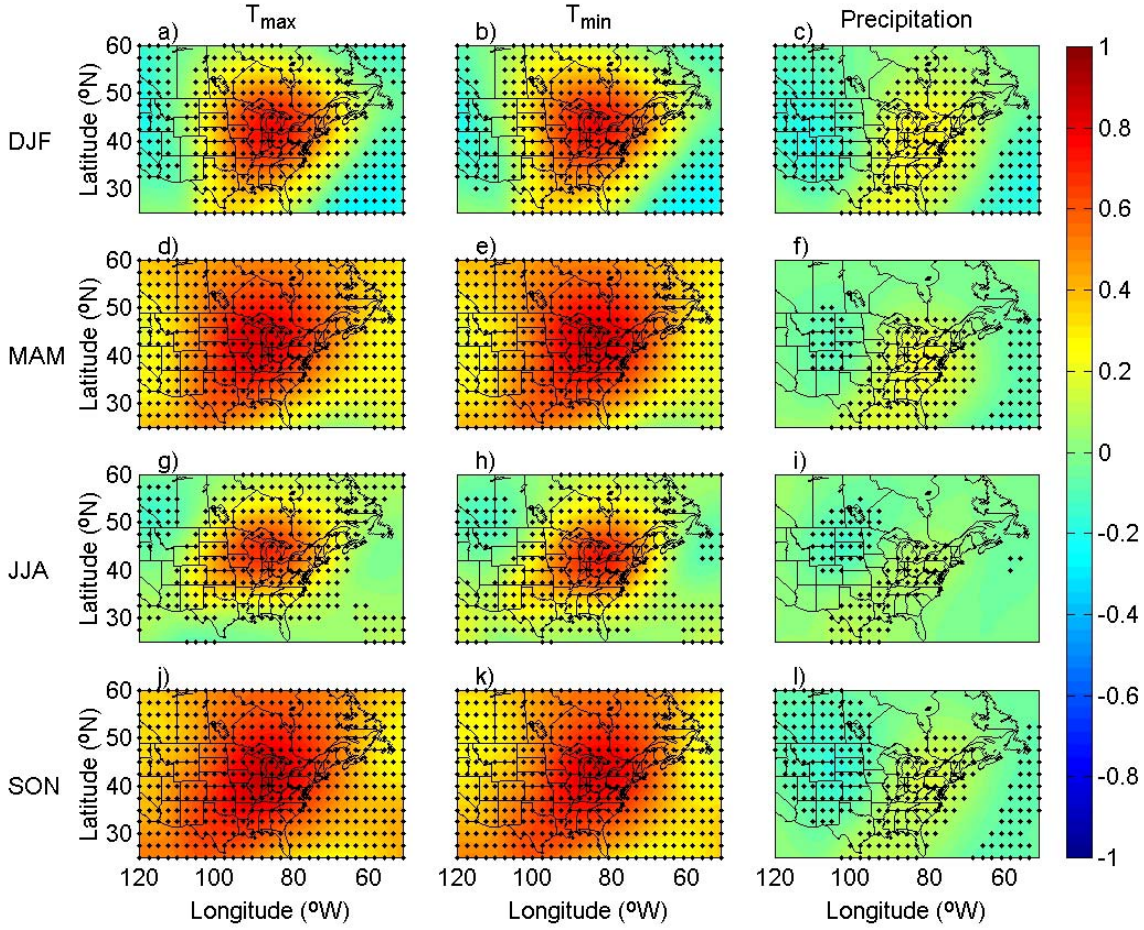

Figure 5.6 Correlation between the daily 850-500 hPa layer thickness field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

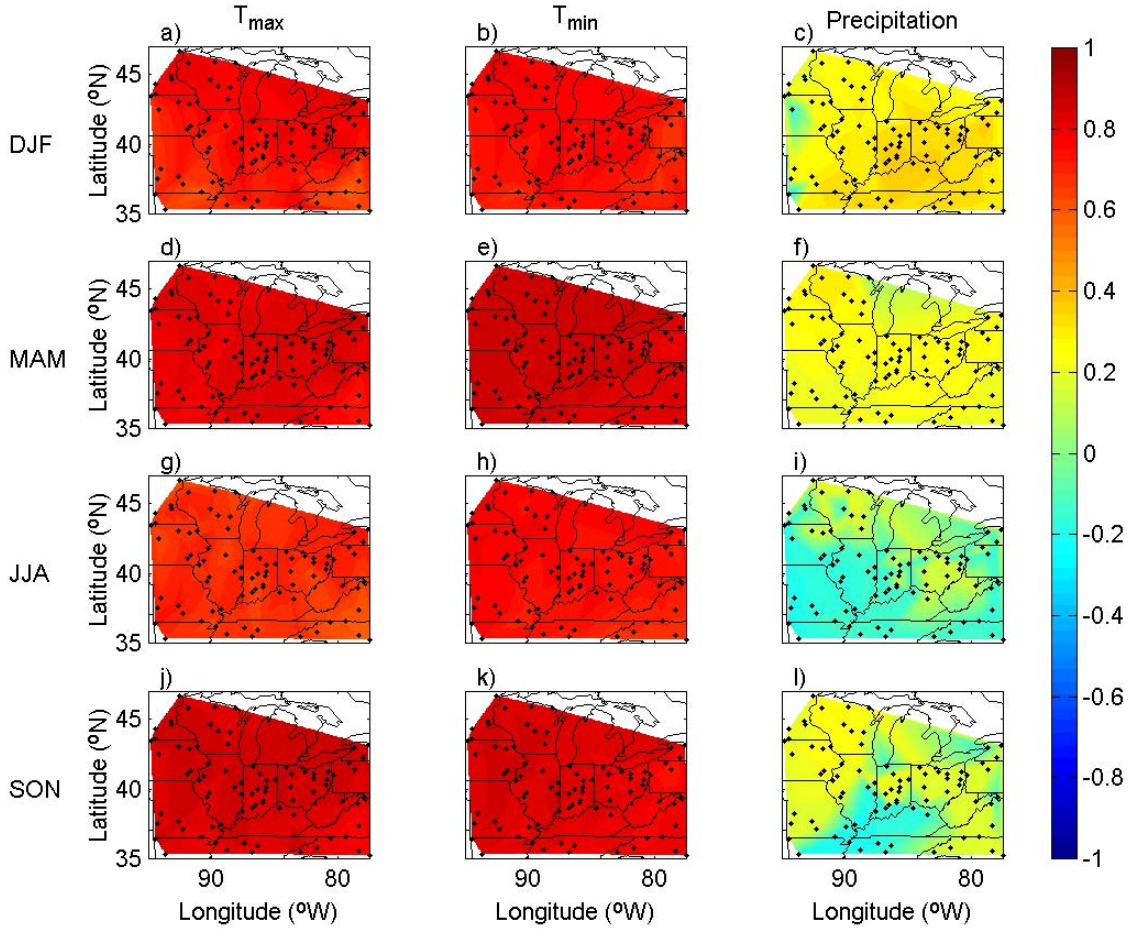

Figure 5.7 Maximum correlation (of positive or negative sign) between the daily 850-500 hPa layer thickness field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.1.4 850-hPa specific humidity (Q_{850})

Specific humidity at 850 hPa is strongly correlated with both T_{max} and T_{min} during all seasons. As with the other free atmosphere variables, the strongest correlations with T_{max} and T_{min} occur during spring and autumn and are statistically significant over much larger areas than during winter and summer (Figures 5.8 and 5.9). For T_{max}, the

maximum correlations range from 0.50 to 0.76, 0.59 to 0.79, 0.46 to 0.66, and 0.67 to 0.80 during winter, spring, summer, and autumn, respectively. The correlations for T_{min} are slightly higher, with maximum values ranging from 0.60 to 0.81, 0.77 to 0.86, 0.65 to 0.83, and 0.80 to 0.91, for winter, spring, summer, and autumn, respectively. These correlations likely result from concurrent changes in surface temperature and Q_{850} associated with synoptic scale systems (e.g., the warm sector of a mid-latitude cyclone passing over the station location resulting in warm temperatures and moisture advection).

The strongest correlations between Q_{850} and station precipitation are positive for all stations during all seasons, although correlations are as low as 0.10 for southern stations during the summer. During winter, stations in the central part of the study area have correlations with Q_{850} as high as 0.55. As anticipated, the correlations between Q_{850} and station precipitation are stronger than the correlations associated with the previously discussed predictor variables. The correlations presented here account for precipitation variability associated with moisture availability, while other predictors may better account for the vertical motions which are also needed for precipitation development.

The correlations between Q_{850} and the three surface variables decrease with increases in the length of the temporal lag. Correlations computed with differenced values of Q_{850} are weaker than their lag-0 counterparts, but are non-trivial for many stations. For example, during summer 24 and 48-hour differences (which may, for example, represent moisture differences across fronts) are correlated with T_{min} with correlations as high as 0.50 for many stations in the northern part of the domain.

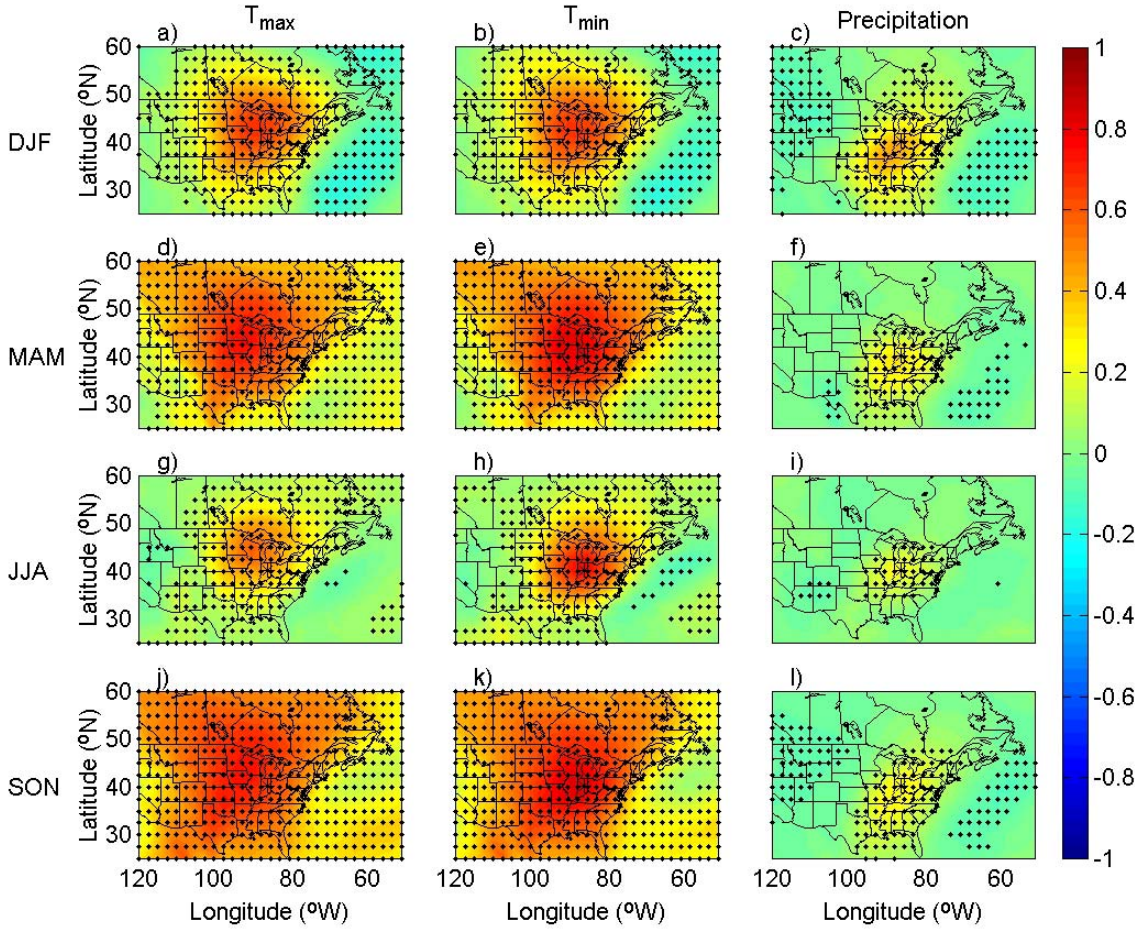

Figure 5.8 Correlation between the daily 850-hPa specific humidity field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

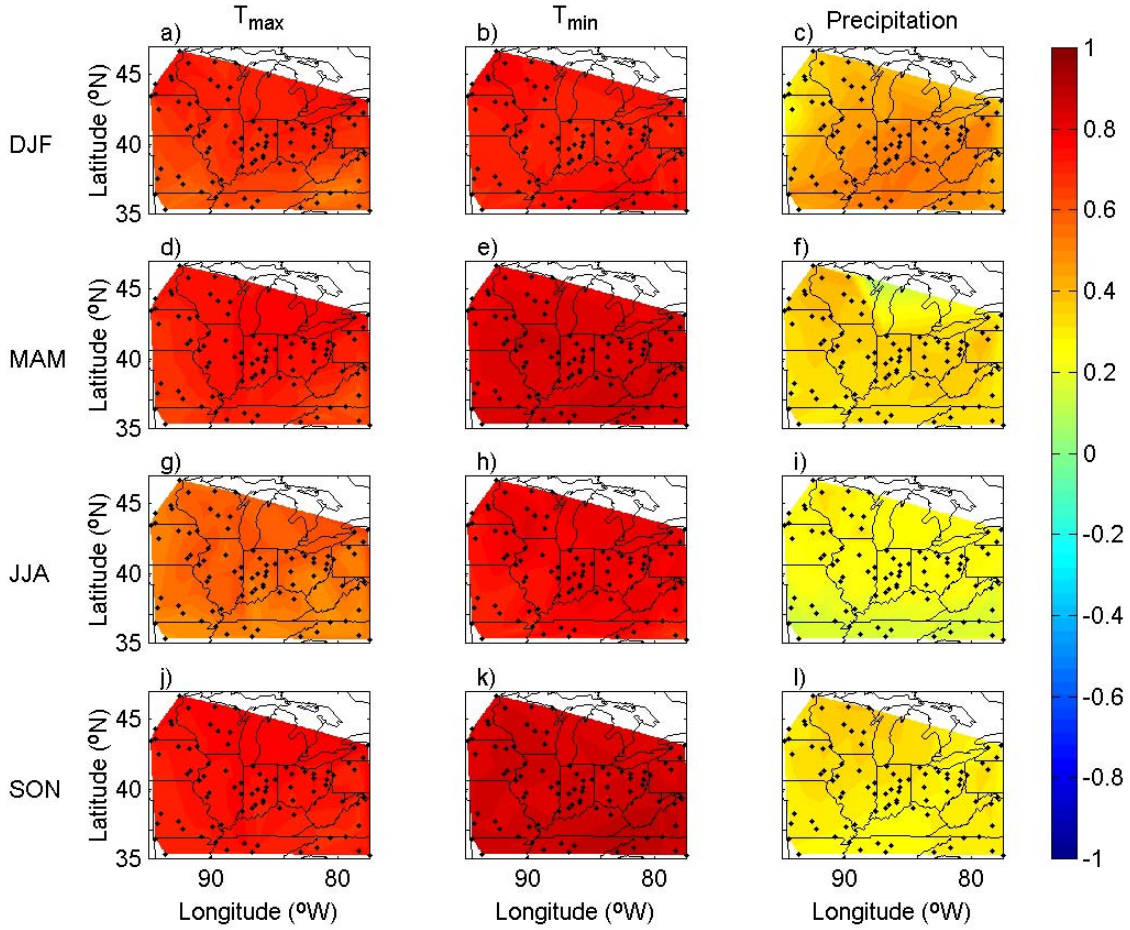

Figure 5.9 Maximum correlation (of positive or negative sign) between the daily 850-hPa specific humidity field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.1.5 850-hPa relative humidity (RH_{850})

Correlations between daily grid-point RH_{850} and surface values of T_{max}, T_{min}, and precipitation are much lower than those for the previously discussed variables, although maximum correlations for each variable at each station are still significant at the 95% level. For all three variables, the grid point correlations are both positive and negative

(Figure 5.10), although with the exception of an off-shore area of negative correlations, the spatial patterns are not consistent. For some stations, the strongest correlation is positive, while for others, the strongest correlation is negative (Figure 5.11). For T_{max}, the maximum correlations range from -0.48 to 0.37, -0.47 to 0.34, -0.52 to -0.20, and -0.52 to 0.56 for winter, spring, summer, and autumn, respectively. For T_{min}, the correlations range from -0.38 to 0.55, -0.42 to 0.46, -0.40 to 0.33 and 0.43 to 0.64 for winter, spring, summer, and autumn, respectively. With the exception of one station during spring, the strongest correlations between RH_{850} and station precipitation are positive, with values ranging from around 0.20 to 0.40.

During autumn, both T_{max} and T_{min} exhibit a region of high correlation with remote RH_{850} grid points in northern Mexico, resulting in positive correlations for most stations during that season (Figures 5.10 and 5.11). These correlations, when taken in concert with similar anomalies in the Q_{850} correlation field (Figure 5.8), suggest that there may a link between increases in southwest US moisture in late summer and autumn associated with the onset of the North American monsoon and surface temperatures in the study area. Enhancement of the western US ridge during these circumstances increases the likelihood of warm air advection into the study region through modification of the longwave pattern.

Due to the temperature dependence of RH_{850}, upper-air observations which are not concurrent with surface precipitation observations are expected to be weaker than those at lag-0. Indeed, the strength of the correlations is always lower at lag-1 and generally less than 0.10 at lag-2. Correlations computed with differenced values of RH_{850} are weaker than the lag-0 values for all three surface variables.

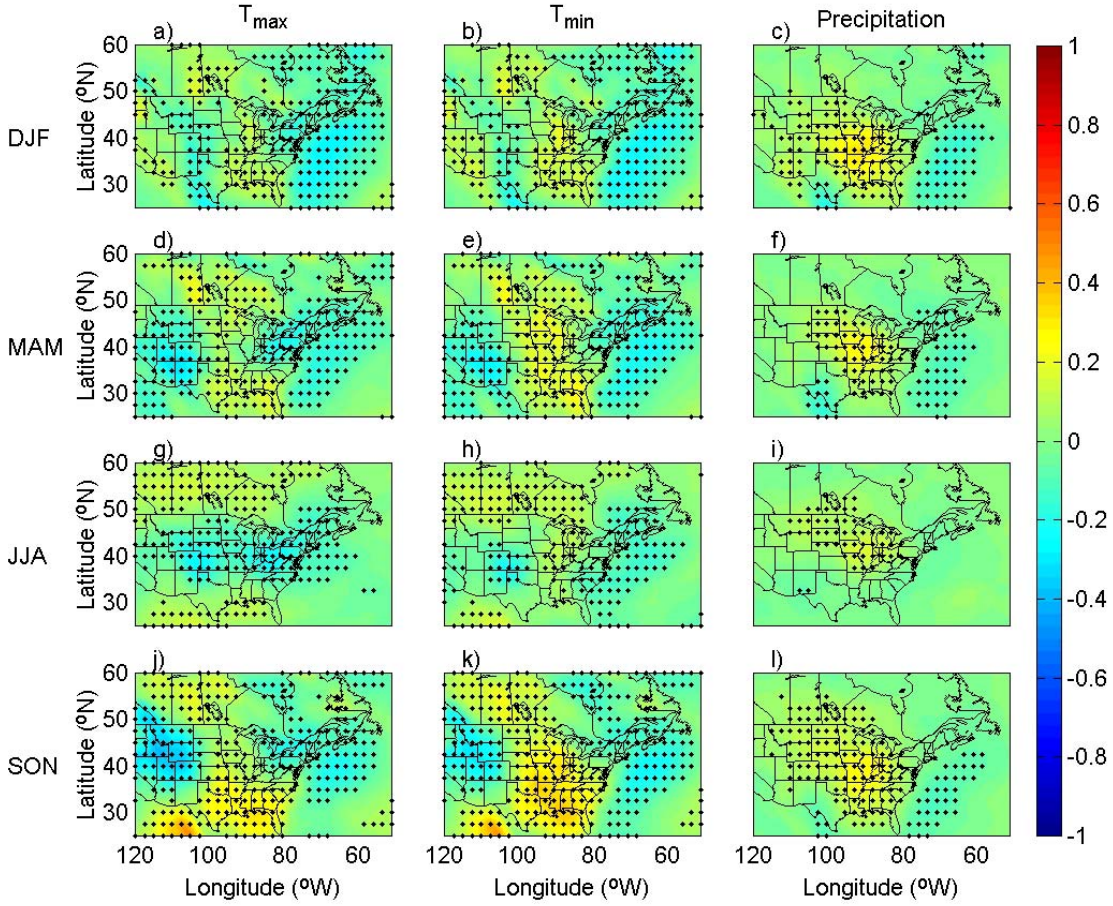

Figure 5.10 Correlation between the daily 850-hPa relative humidity field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

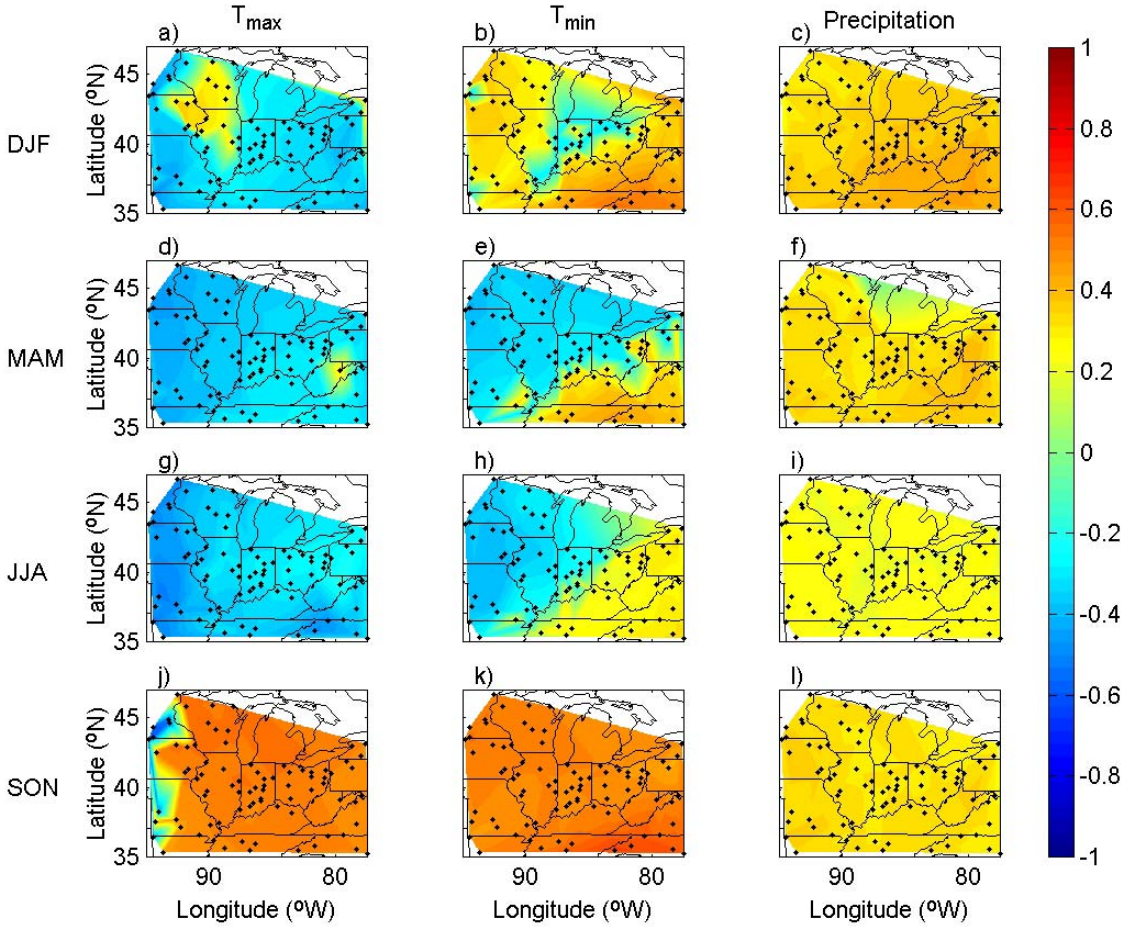

Figure 5.11 Maximum correlation (of positive or negative sign) between the daily 850-hPa relative humidity field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.2 Surface variables

5.5.2.1 Sea-level pressure (SLP)

Patterns of the correlation between sea-level pressure and the temperature variables exhibit a dipole pattern, with negative correlations covering a large region west of the station location and positive (and generally weaker) correlations covering a large

area east of the station (Figure 5.12). As Figure 5.12 shows, the correlations are weakest during summer and cover the largest area during autumn. For precipitation, the correlation pattern is similar, but substantially weaker. The negative correlations with temperature are remote from the station location and therefore thought to be associated with warm air advection when a low surface pressure occurs west and southwest of the station location rather than with increased radiative effects associated with cloud cover increases in the vicinity of passing cyclones, while the positive correlations are associated with warm air advection associated with high surface pressure to the east of the station location. For precipitation, maximum grid point correlations are negative and occur near the station location, consistent with the increases in vertical motion associated with decreases in sea-level pressure. Positive correlations with precipitation are likely to be associated with increased moisture advection associated with anticyclonic circulation east of the station location (Figure 5.12).

Figure 5.13 shows the maximum correlation between station values of T_{max}, T_{min}, and precipitation and SLP field and indicates that, for the temperature variables, the strongest correlations are negative during winter, spring, and autumn, while some stations exhibit positive correlations during the summer. The strongest correlations occur during autumn, when all stations exhibit correlations between -0.50 and -0.60 for T_{max} and between -0.47 and -0.62 for T_{min}. For T_{min}, the correlations are also consistently high during winter, while those for T_{max} are weaker for stations in the western part of the study area. The strongest correlations with precipitation are negative for all stations and all seasons, ranging from -0.40 to -0.17, -0.39 to -0.09, -0.31 to -0.14, and -0.41 to -0.22 for winter, spring, summer, and autumn, respectively. The weaker relationship during

summer is directly related to the convective, rather than frontal, nature of summer precipitation.

Lag-1 and lag-2 correlations with SLP are similar to those at lag-0 for the temperature variables and slightly weaker for precipitation. Correlations computed with differenced values are weaker than those at lag-0, although still statistically significant for most stations during each season.

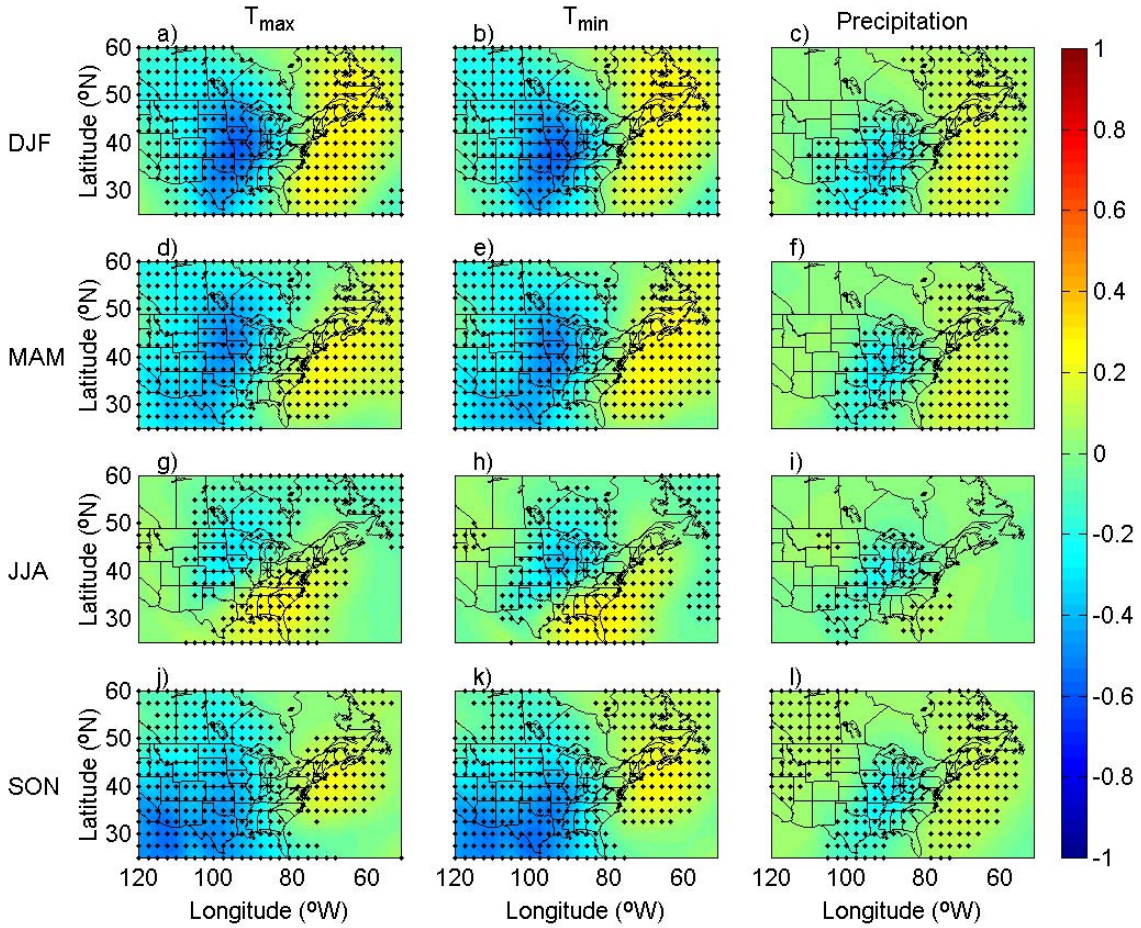

Figure 5.12 Correlation between the daily sea level pressure field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

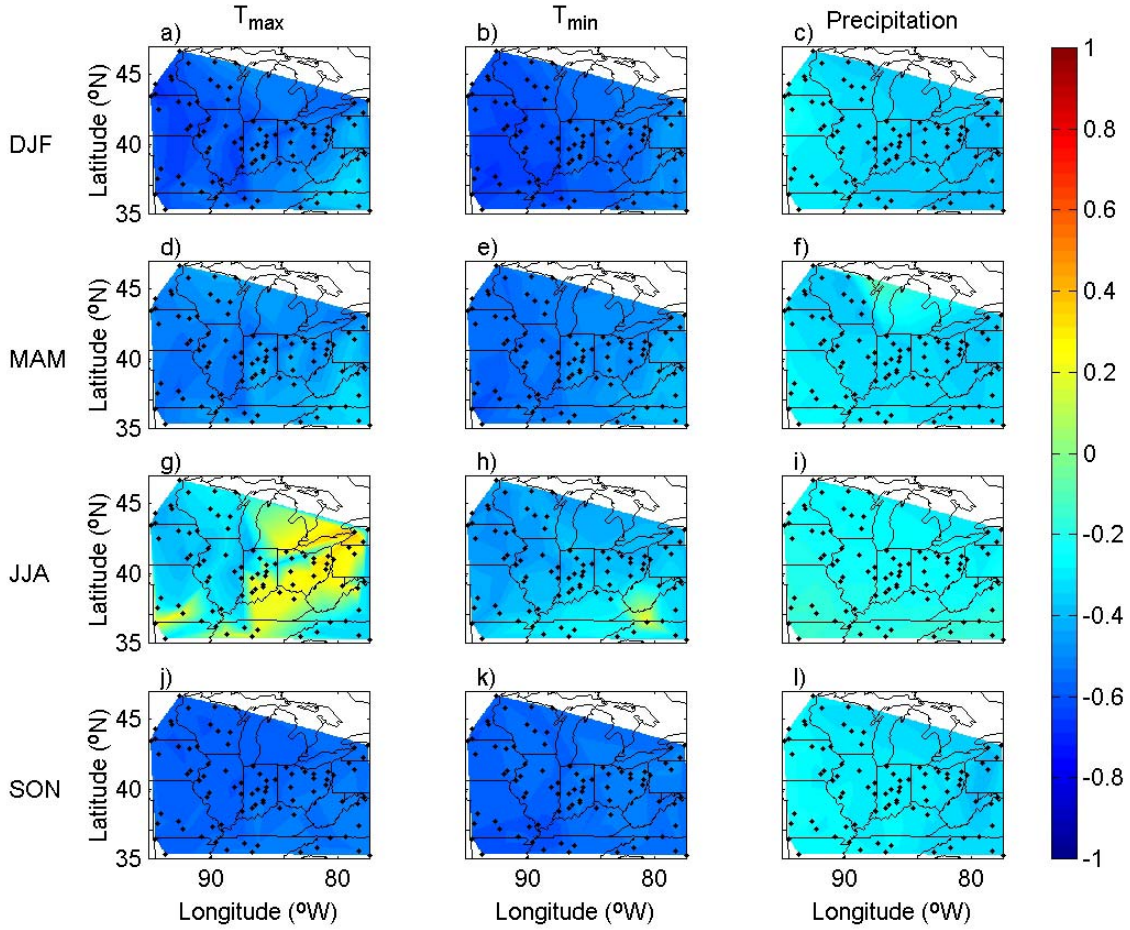

Figure 5.13 Maximum correlation (of positive or negative sign) between the daily sea level pressure field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.2.2 Zonal component of geostrophic flow (GEOW)

Correlations between the zonal component of geostrophic flow (GEOW) and the surface temperature variables are generally positive at locations near and just south and southwest of the station location, with areas of negative correlation north and southeast of the station location (Figure 5.14). During autumn, the station location is typically surrounded by small negative correlations, with positive correlations located well to the

south and west of the station location. These patterns are consistent with warm air advection when grid cells south of the station location exhibit strong westerly flow and cold air advection when grid cells north of the station location exhibit strong westerly flow. For precipitation, the correlations are lower, but display a similar pattern with positive values south of the station location and negative values north and southeast of the station location, consistent with the advection of moist and dry air, respectively.

For T_{max}, the strongest correlations with GEOW range from -0.40 to 0.64, -0.38 to 0.51, 0.27 to 0.57, and -0.36 to 0.45, for winter, spring, summer, and autumn, respectively, while those for Tmin are slightly stronger, ranging from -0.46 to 0.64, -0.37 to 0.57, 0.35 to 0.61, and 0.26 to 0.57 (Figure 5.15). The strongest correlations between GEOW and station precipitation range from -0.36 to 0.40, -0.31 to -0.39, -0.21 to 0.31, and -0.29 to 0.36 for winter, spring, summer, and autumn, respectively (Figure 5.15).

Lag-1 correlations are similar in magnitude and pattern to those at lag-0, while those for lag-2 are considerably lower. For the temperature variables, correlations computed with differenced values of GEOW are weaker than those at lag-0. For precipitation, correlations are of similar magnitude and pattern.

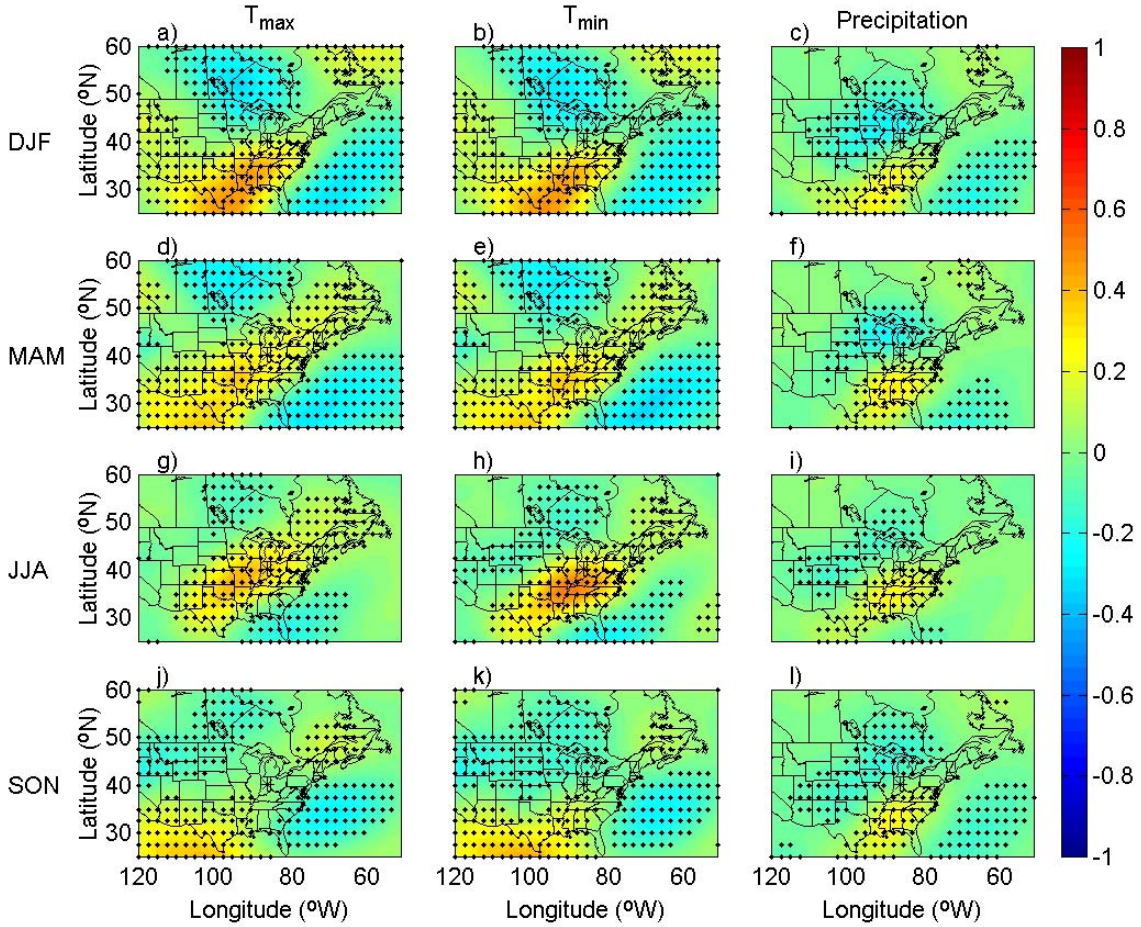

Figure 5.14 Correlation between the daily GEOW field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

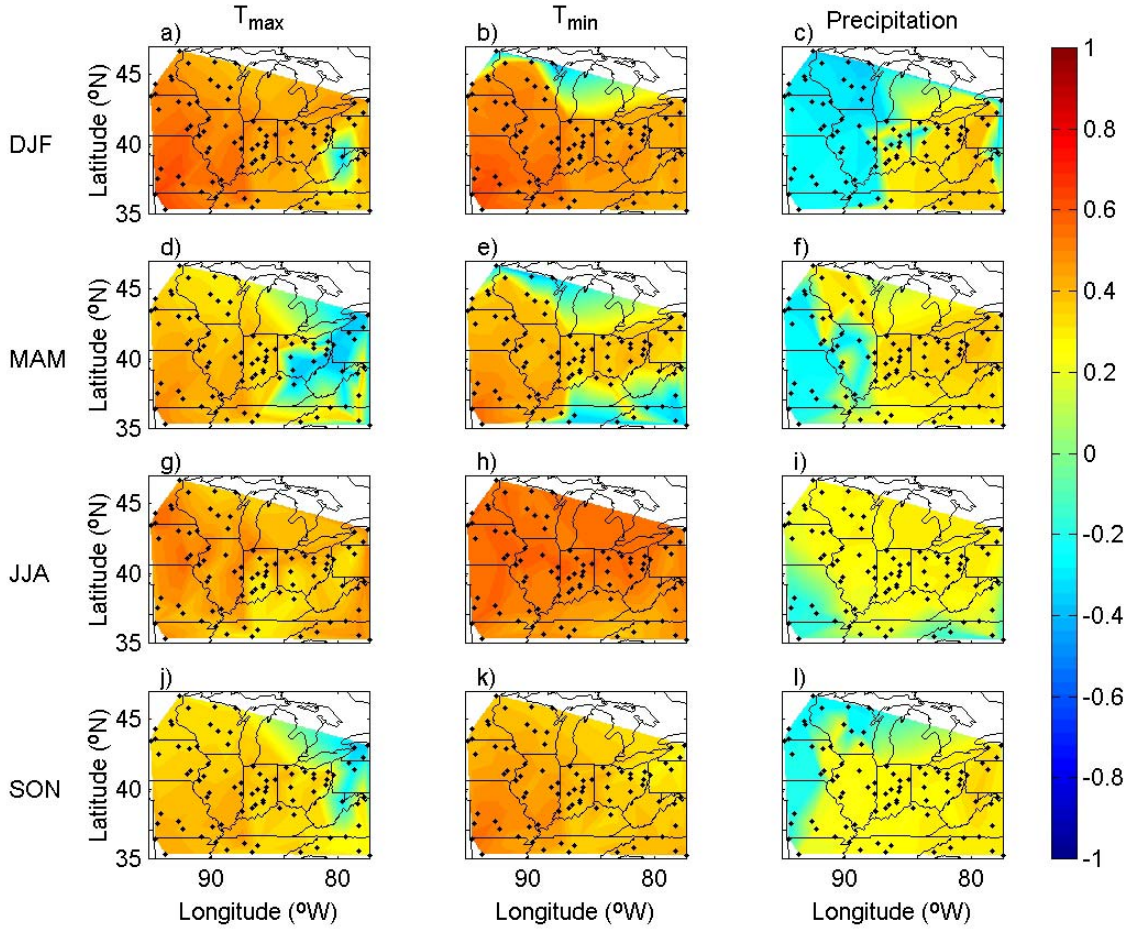

Figure 5.15 Maximum correlation (of positive or negative sign) between the daily GEOW field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.2.3 Meridional component of geostrophic flow (GEOS)

The correlations between the meridional component of geostrophic flow (GEOS) and the surface variables are shown in Figure 5.16. As expected, the temperature variables exhibit high correlations with grid points south of the station location consistent with warm air advection into the station vicinity, associated with flow from the south.

An area of negative correlations is also present west of the station location and is strongest during the winter. For precipitation, the correlation patterns are weaker, but a consistent pattern of positive correlations east of the station location and negative correlations west of the station location is present. This pattern indicates that precipitation amounts are greater with southerly flow east of the station location (consistent with moisture advection) and lesser when southerly flow occurs west of the station location (and the advected air is likely to be drier). The patterns for both temperature and precipitation are also consistent with a long-wave pattern that places the study area on the exit sector of a trough.

As shown in Figure 5.17, for some stations the strongest correlations with temperature variables are associated with the negative correlations to the west of the station location. For temperature, the strongest correlations are positive for all stations during winter, spring, and autumn and negative for some stations during summer.

At lag-1, the correlations are of similar magnitude to those at lag-0. However, during winter, the temperature variables exhibit consistently positive correlations (across stations). For precipitation, the lag-1 correlations are smaller in magnitude that those at lag-0. Lag-2 correlations are lower than those at lag-0 for all three variables. Correlations between differenced GEOW and the station variables are negative for all seasons, stations, and variables, but weaker in magnitude than the lag-0 values, although statistically significant for each variable, station, season.

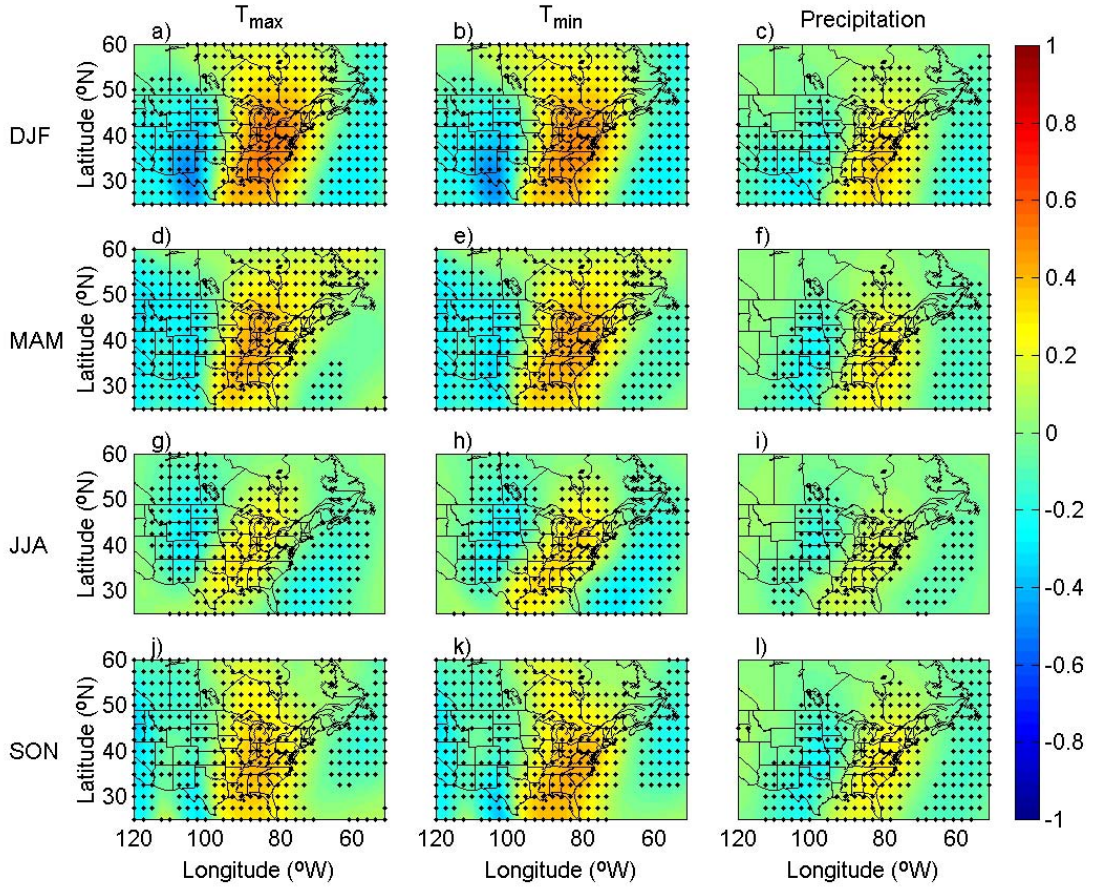

Figure 5.16 Correlation between the daily GEOS field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

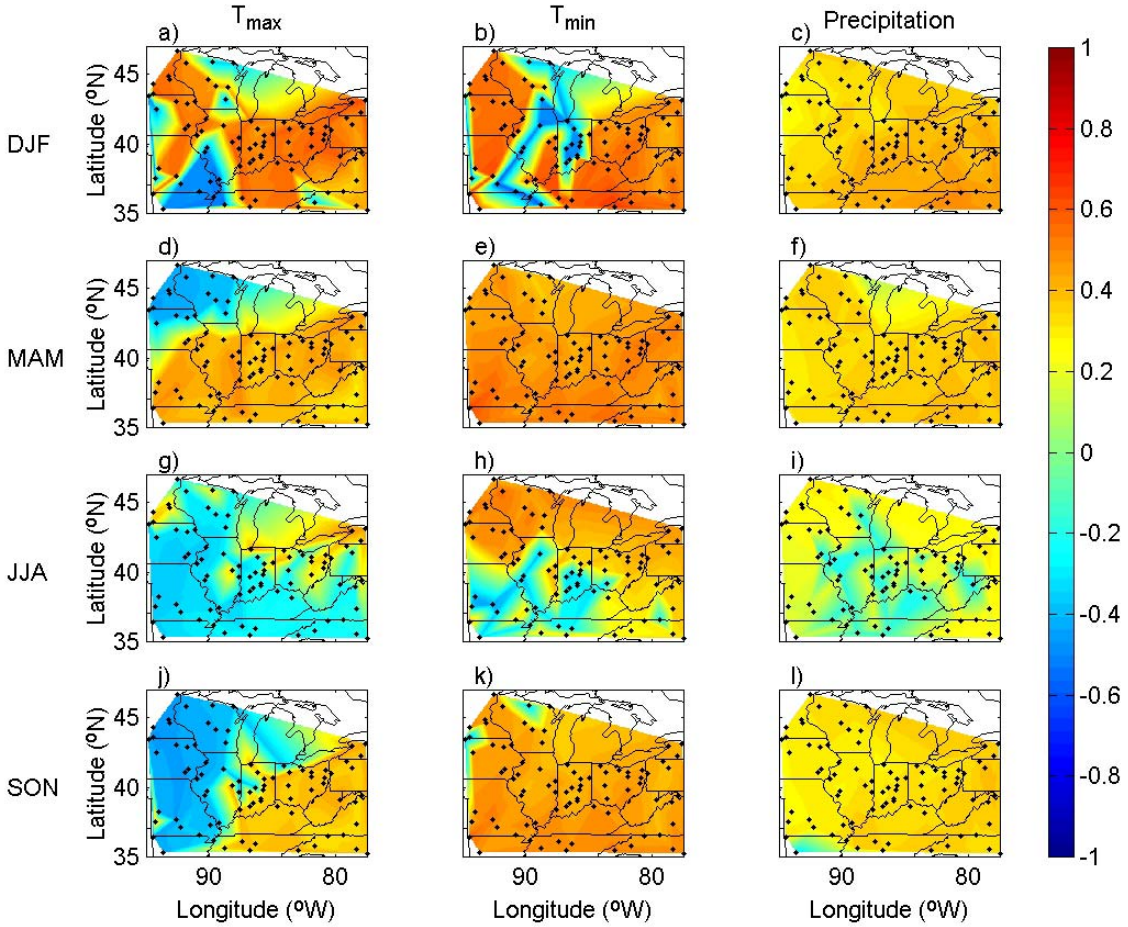

Figure 5.17 Maximum correlation (of positive or negative sign) between the daily GEOS field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.2.4 Resultant geostrophic flow (GEOWS)

Correlations between the strength of the resultant geostrophic flow (GEOWS) and the surface variables shown in Figure 5.18 are generally lower than those for the other predictors. For the temperature variables, the area of maximum correlation is usually remote from the station location (northern Mexico), although the correlation pattern is not

coherent from station to station or between seasons. For precipitation, the maximum correlations are positive and located near and to the south of the station location, consistent with a strong pressure gradient in that region. Given the geographic location of the study area, it is not surprising that the surface variables exhibit higher correlations with directional flow variables (e.g., GEOS and GEOW which capture advection of warm, moist air from the south, or cool dry air from the west) than with resultant flow.

The strongest correlations are presented in Figure 5.19 and show that for T_{max} and T_{min}, the strongest correlations are consistently positive during winter, spring, and autumn, while some correlations are positive during the summer. Correlations with precipitation are lower, with the strongest correlations ranging from -0.16 to 0.27, 0.06 to 0.28, 0.10 to 0.28, and -0.20 to 0.21 for winter, spring, summer, and autumn, respectively.

Lag-1 and lag-2 correlations between GEOWS and the temperature variables are similar to those at lag-0 during the winter, spring, and autumn, and slightly stronger during the summer. For precipitation, lag-1 and lag-2 correlations are weaker than those at lag-0. For T_{max} and T_{min}, the correlations computed with differenced values of GEOWS are lower than those at lag-0, while those for precipitation are of similar magnitude.

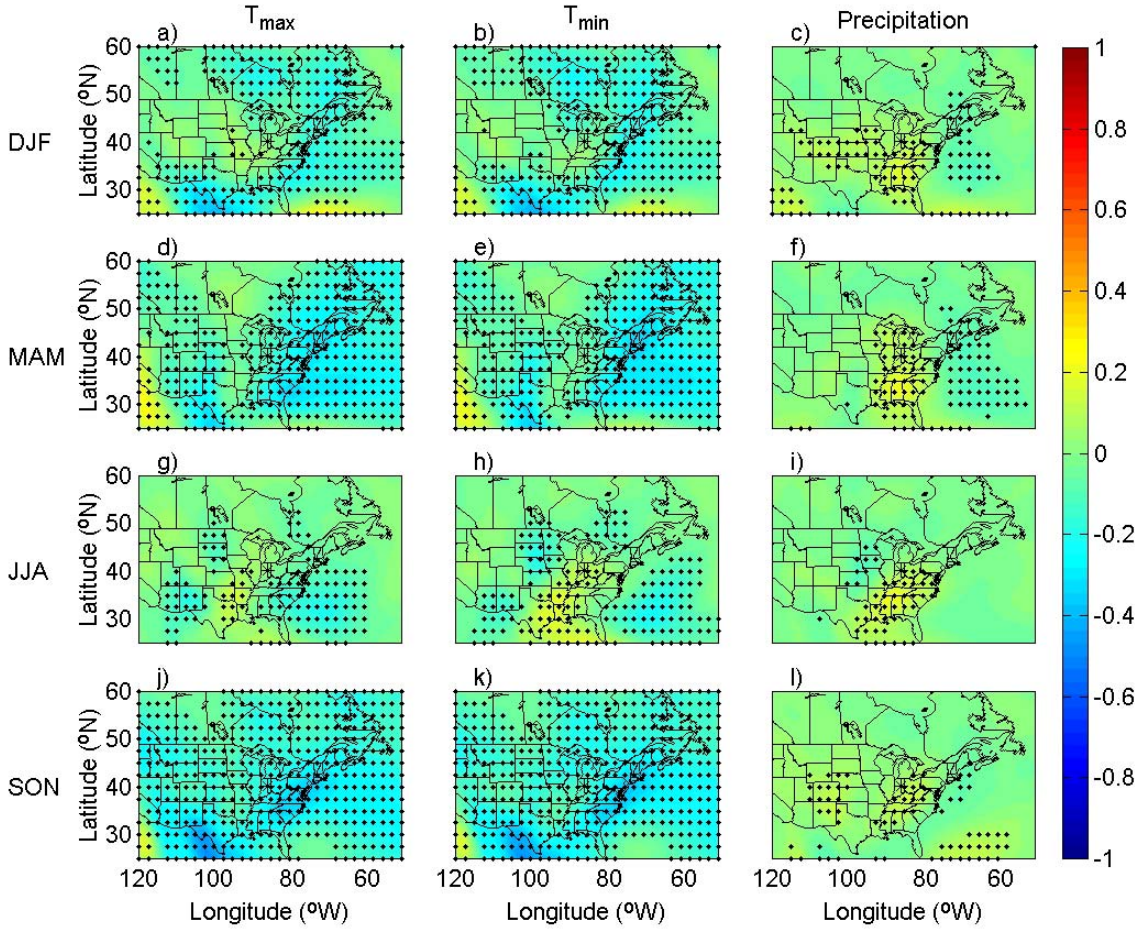

Figure 5.18 Correlation between the daily GEOWS field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

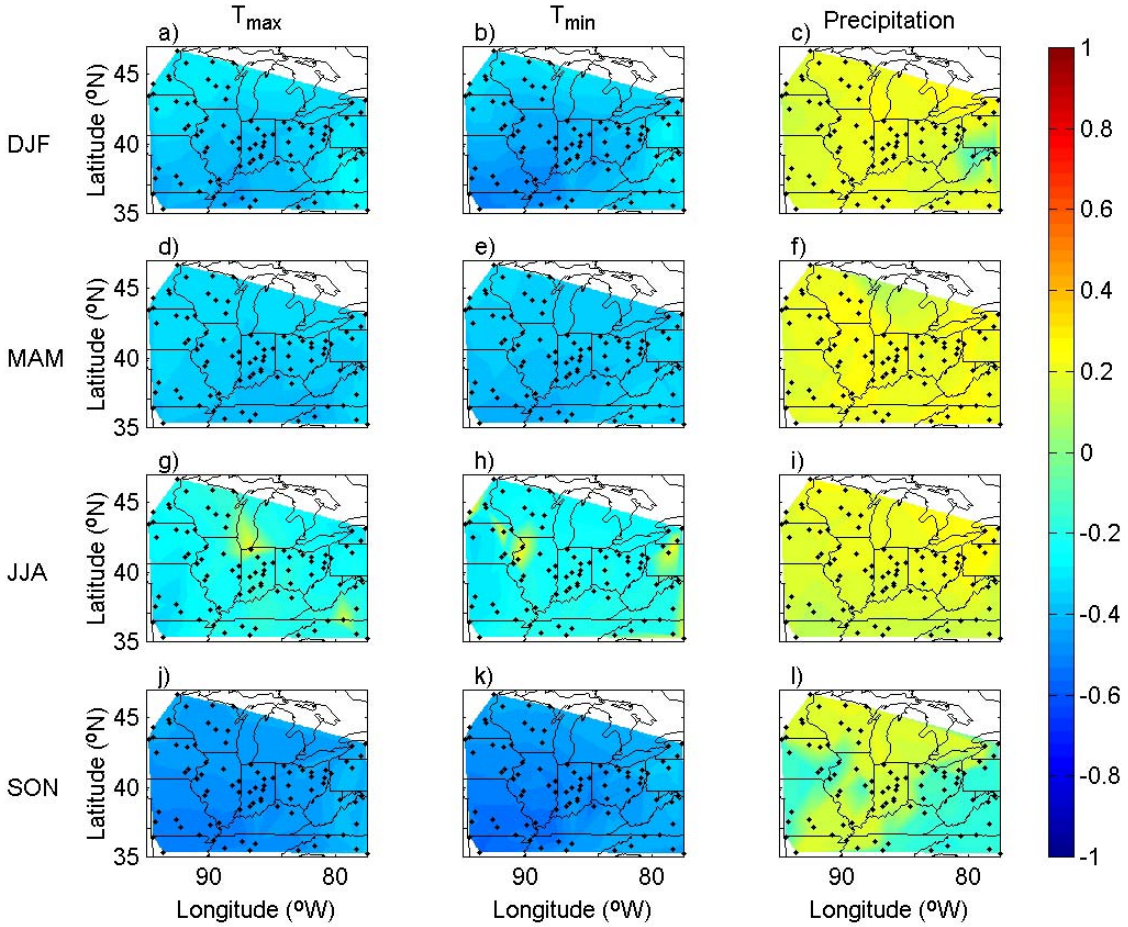

Figure 5.19 Maximum correlation (of positive or negative sign) between the daily GEOWS field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.2.5 Westerly geostrophic vorticity (GEOZW)

The correlation between GEOZW and the surface temperature variables is positive at grid points near or slightly north of the station location and negative at locations south of the station location. Positive values of GEOZW occur when the westerly component of geostrophic flow south of the grid point exceeds that north of the

grid point. These areas of positive correlation are therefore associated with warm air advection into the study region, consistent with the station being located in the exit sector of either a long-wave trough or short-wave trough (i.e., at the trough axis the flow has a large westerly component, while in the exit sector the geostrophic flow has a larger southerly component). Similarly, when the trough axis is located south of the station location, the station is itself is subjected to cold air advection resulting in the negative correlations depicted in Figure 5.20. As Figure 5.20 shows, the region of positive correlations is strongest during winter and summer, while the negative correlations are most prevalent during summer. For precipitation, positive correlations are found near the station location, with a region of weaker negative correlations both north and south of the station location. These correlations result from the same physical mechanism described above (i.e., enhanced vertical motion when the station location lies within the exit sector of a long or short-wave trough).

As shown in Figure 5.21, the strongest correlations between GEOZW and T_{max} range from -0.44 to 0.46, -0.42 to 0.34, -0.42 to 0.40, and -0.40 to 0.41 for winter, spring, summer, and autumn, respectively. For T_{min}, correlations are slightly stronger, ranging from -0.35 to 0.51, -0.38 to 0.34, -0.49 to 0.49 and -0.30 to 0.44. The strongest correlations with precipitation are positive for all stations in each season and range from 0.19 to 0.48, 0.07 to 0.47, 0.19 to 0.35, and 0.21 to 0.40 for winter, spring, summer, and autumn, respectively. Correlations between GEOZW and the surface variables computed with lagged and differenced values are substantially lower than lag-0 values, with insignificant values at many stations for the differenced variables.

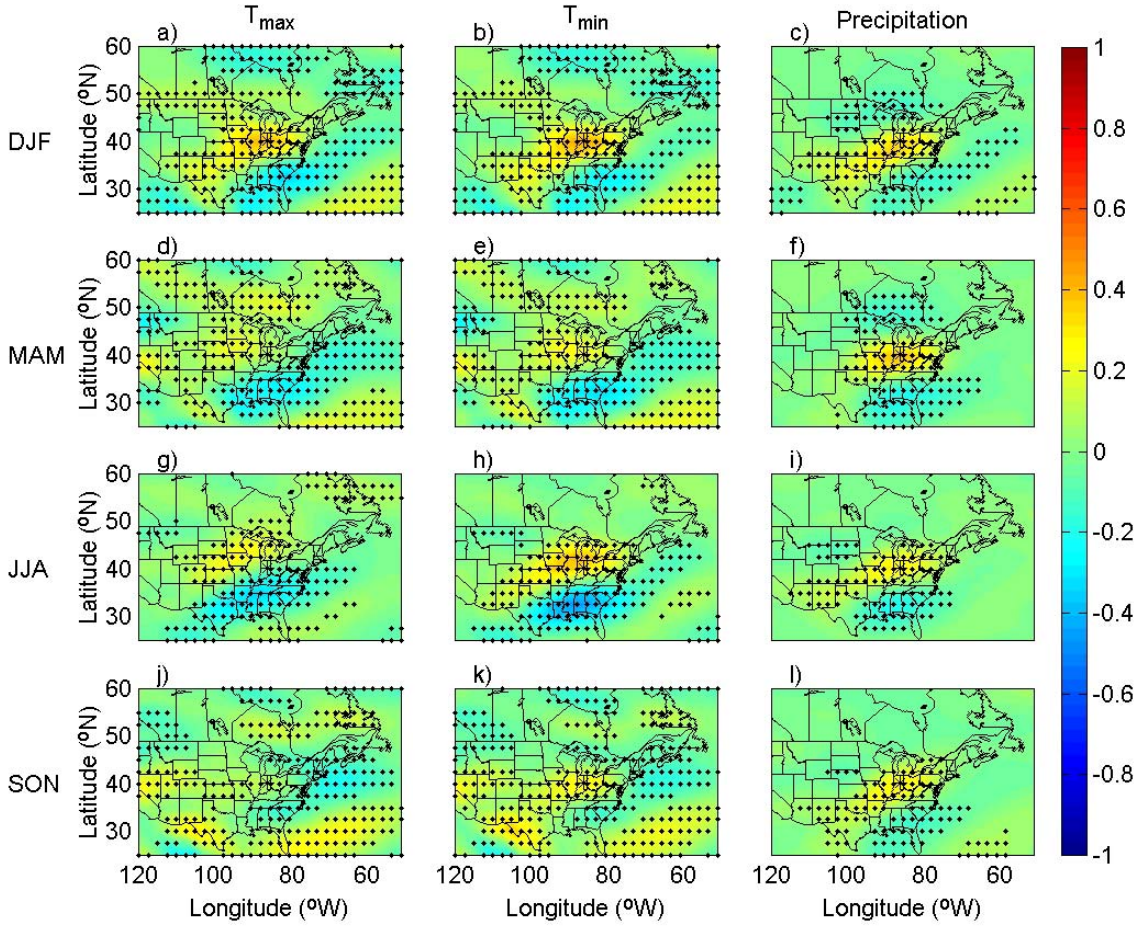

Figure 5.20 Correlation between the daily GEOZW field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

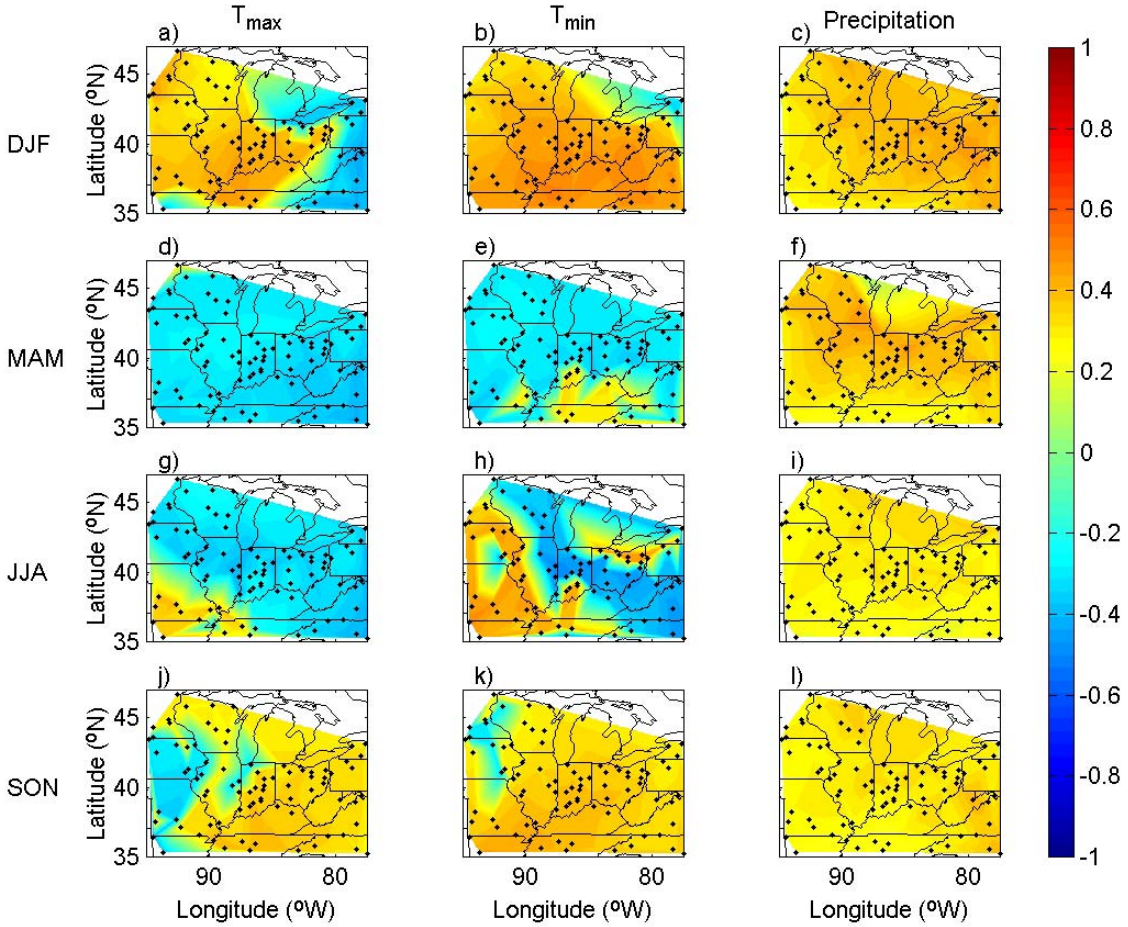

Figure 5.21 Maximum correlation (of positive or negative sign) between the daily GEOZW field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.2.6 Southerly geostrophic vorticity (GEOZS)

Correlations between the southerly geostrophic vorticity (GEOZS) and the surface temperature variables exhibit an interesting spatial pattern with alternating bands of positive and negative correlations located remotely from the station location (Figure 5.21), consistent with temperature advection associated with the longwave pattern

associated with strength of GEOZS in those areas. As Figure 5.22 shows, the pattern is strongest during winter and autumn and slightly weaker during the spring and summer. For precipitation, positive correlations are found directly to the west of the station location, banded by weaker negative correlations (Figure 5.22). This pattern is consistent with both moisture advection and enhanced vertical motion in the presence of positive GEOZS values in the area of positive correlations.

Figure 5.23 shows the strongest correlation between GEOZS and the surface variables for each station and indicates a high degree of spatial variability. For some stations, the relationship is quite weak with near zero correlations for at least a few stations for each variable in each season. For example, the maximum correlation between GEOZS and precipitation during spring ranges from 0.00 to 0.43 (Figure 5.23).

For the temperature variables, lag-1 correlations are similar to those at lag-1, while for precipitation lag-1 values are lower. For all three surface variables, correlations at lag-2, and those computed with differenced GEOZS values are weaker than those at lag-0.

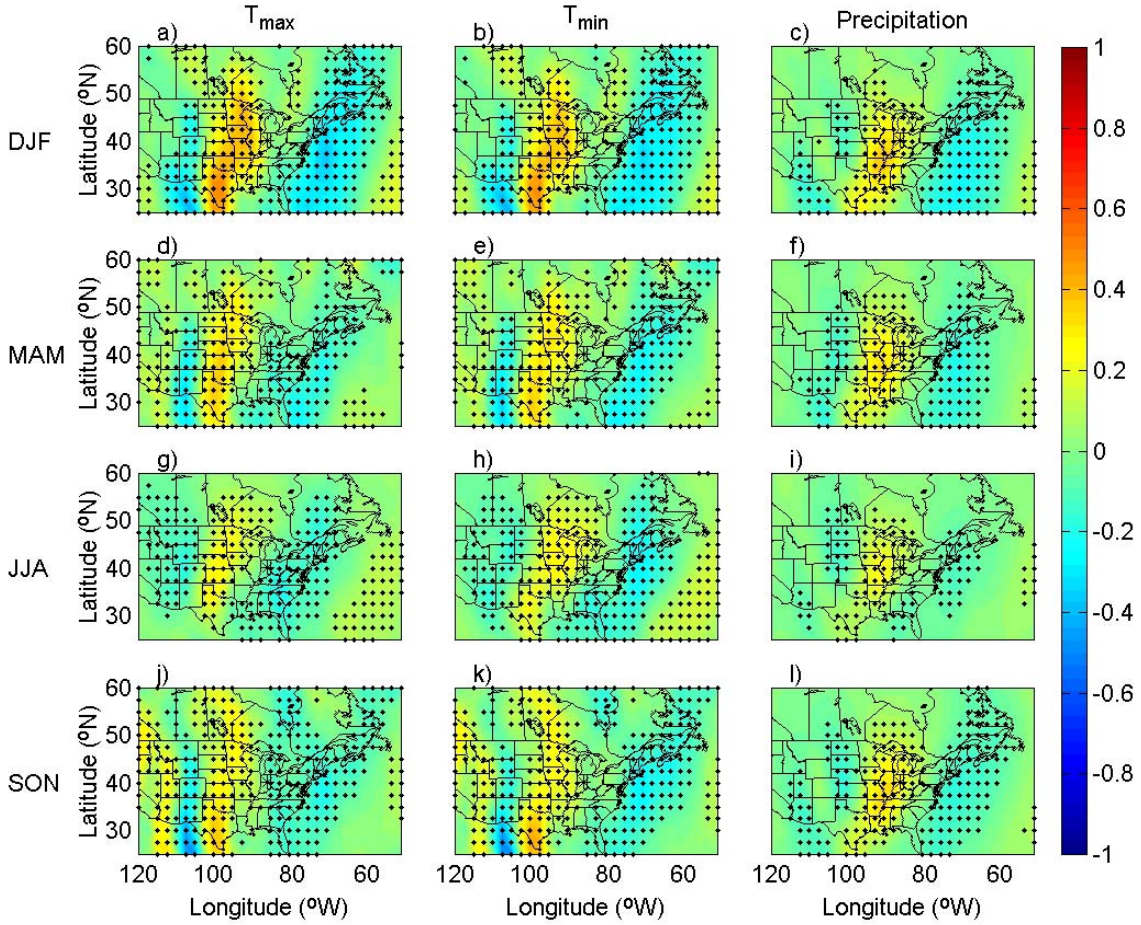

Figure 5.22 Correlation between the daily GEOZS field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

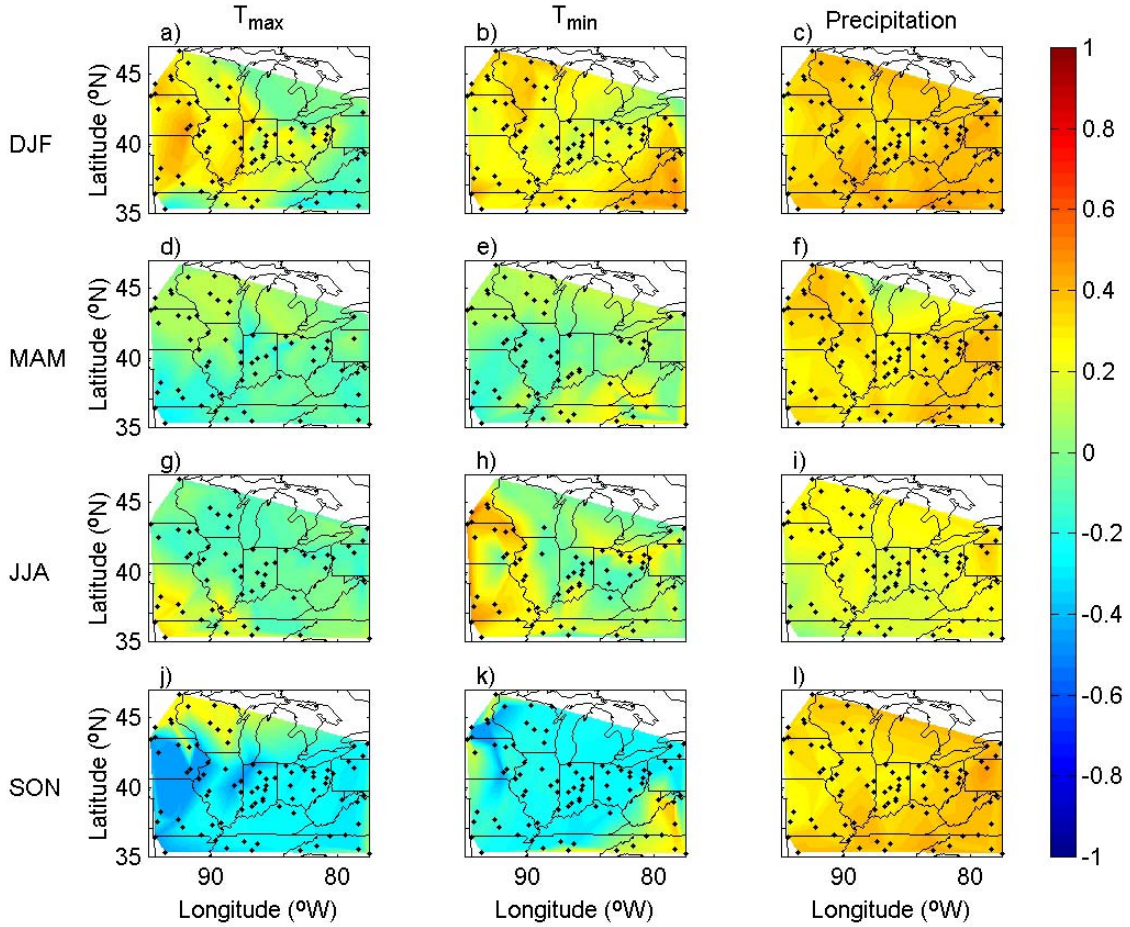

Figure 5.23 Maximum correlation (of positive or negative sign) between the daily GEOZS field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.5.2.7 Total geostrophic vorticity (GEOZT)

The correlations between total geostrophic vorticity (GEOZT) and the surface variables are similar to those for both GEOZW and GEOZS described above. For all three surface variables, the general pattern of correlations is characterized by a region of positive correlations located near the station location and stretching to the south and west accompanied by an area of negative correlations south of the station location (Figure

5.24). These patterns likely reflect the station location relative to cyclonic circulation, and hence, the highest correlations are found in areas that would place the station in the warm sector of the cyclone.

As shown in Figure 5.25, the sign of the strongest correlations between GEOZT and the surface air temperature variables varies by season and location. Correlations between GEOZT and T_{max} range from -0.43 to 0.54, -0.41 to 0.40, -0.41 to 0.45, and -0.55 to -0.31 for winter, spring, summer, and autumn, respectively, while those for Tmin range from 0.31 to 0.58, -0.35 to 0.49, -0.47 to -0.57, and -0.52 to 0.54 for winter, spring, summer, and autumn, respectively. For precipitation, correlations are consistently positive across seasons and stations. These correlations are among the highest observed for precipitation, ranging from 0.23 to 0.54, 0.07 to 0.51, 0.21 to 0.39, and 0.28 to 0.48 for winter, spring, summer, and autumn, respectively. These results further emphasize the importance of vorticity advection in enhancing vertical motions resulting in precipitation occurrence.

Lag-1 and lag-2 values of the correlation between GEOZT and the surface variables are generally weaker than those at lag-0, although those for the temperature variables at lag-1 are similar to lag-0 values at most locations and times of year. Differenced values of GEOZT produce correlations that are lower than lag-0 values and statistically insignificant for some stations and seasons.

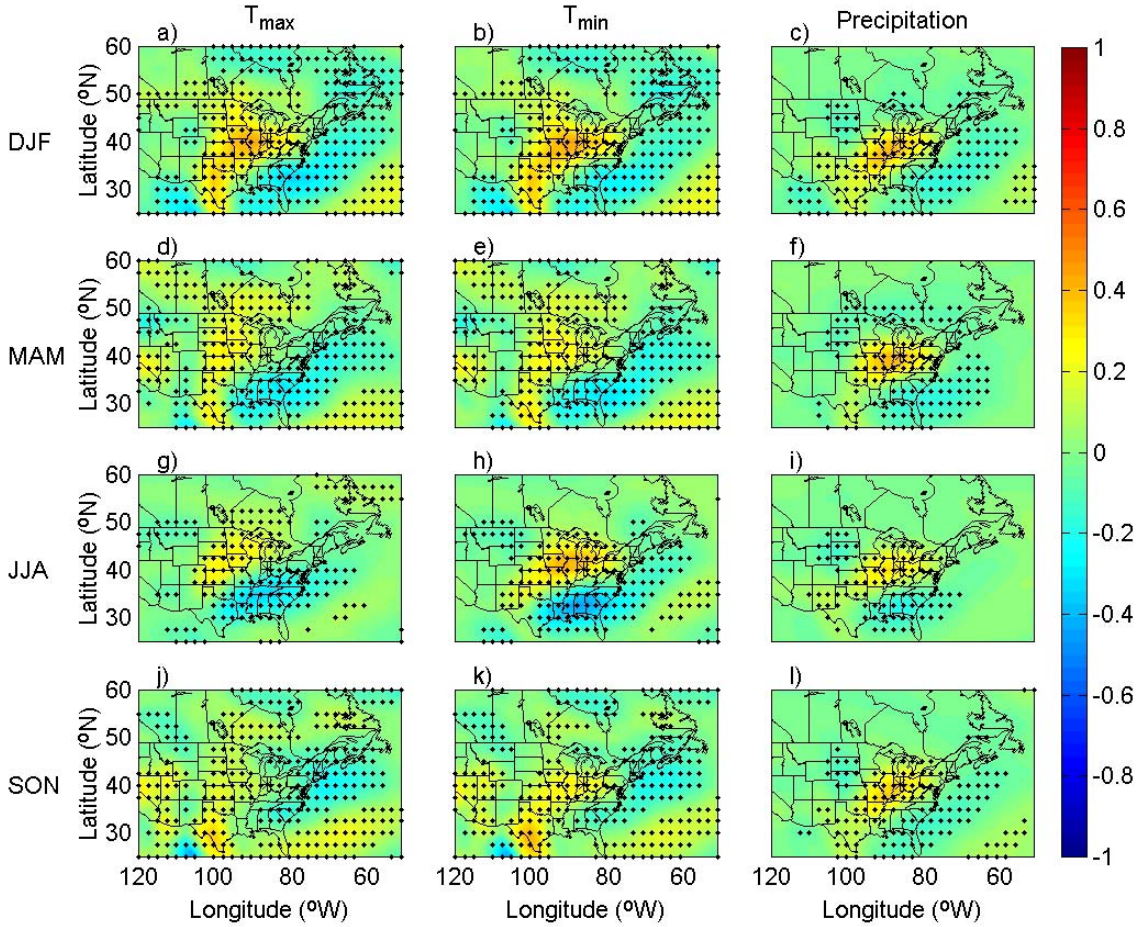

Figure 5.24 Correlation between the daily GEOZT field and surface variables at Anderson, IN (∗): a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation. Reanalysis grid points are shown from grid point locations which are significantly correlated with station T_{max}, T_{min} and precipitation at the 95% level.

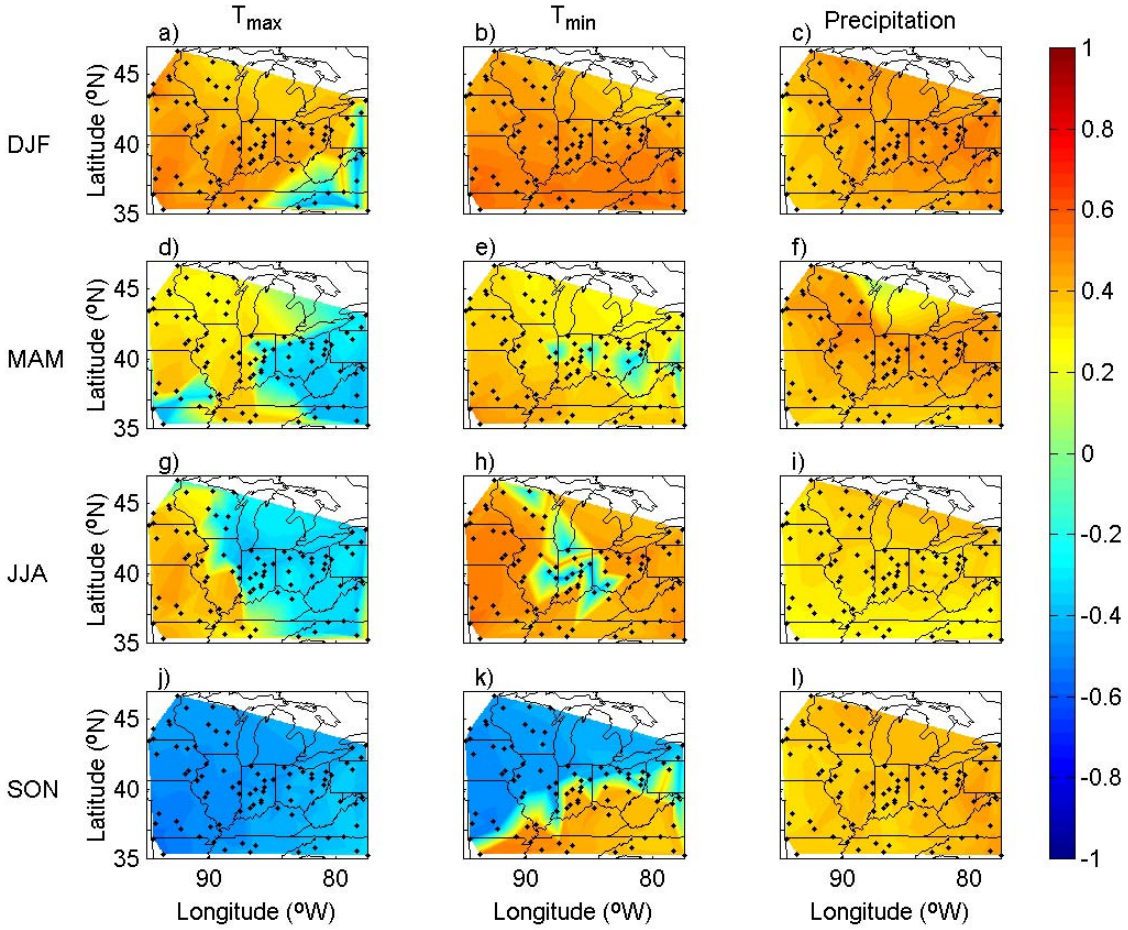

Figure 5.25 Maximum correlation (of positive or negative sign) between the daily GEOZT field and surface variables at the USHCN/D stations: a) winter T_{max}, b) winter T_{min}, c) winter precipitation, d) spring T_{max}, e) spring T_{min}, f) spring precipitation, g) summer T_{max}, h) summer T_{min}, i) summer precipitation, j) autumn T_{max}, k) autumn T_{min}, and l) autumn precipitation.

5.6 Summary

The relationships described above suggest that the predictor variables chosen in this study are closely related to the surface variables, although the nature and intensity of the associations vary substantially both geographically and seasonally. In most cases, the correlations can be explained in terms of the physical causes which are related to either thermal or moisture advection or enhanced vertical motion.

To provide a hierarchy of potential predictor variables, the twenty most highly correlated predictor variables for T_{max}, T_{min}, and precipitation at a central station (Anderson, IN) are tabulated in Tables 5.2, 5.3, and 5.4, respectively. As shown in the tables, and described in Section 5.5, several of the predictors exhibit strong correlations with T_{max} and T_{min} at the USHCN/D stations. The variables exhibiting the strongest relationships with T_{max} and T_{min} (in all seasons) are 850-500 hPa thickness (THICK), 500-hPa geopotential height (Z_{500}), and 850-hPa specific humidity (Q_{850}), sea-level pressure (SLP), the southerly and westerly components of the geostrophic flow (GEOS and GEOW), and 850-hPa geopotential height (Z_{850}). Each of the other variables exhibit strong correlation with T_{max} and T_{min} during some seasons at some stations. Correlations are generally higher and significant over larger areas during spring and autumn, consistent with higher spatial autocorrelation of the surface variables during those seasons.

As shown in Table 5.4, the correlations between the predictor variables and station precipitation are weaker than those for the temperature variables, but several predictor variables have strong relationships with station precipitation. At most stations, the most highly correlated predictors are the geostrophic vorticity variables (GEOZS, GEOZW, and GEOZT), 850-hPa specific humidity (Q_{850}), the southerly and westerly components of the geostrophic flow (GEOS and GEOW), and sea-level pressure (SLP). Several other variables, including 850-hPa relative humidity (RH_{850}) and the resultant geostrophic flow, exhibit high correlations with some stations during some seasons.

Table 5.2. Predictor variable hierarchy for T_{max} at a central station (Anderson, IN). The table shows the twenty variables that exhibit the strongest correlation with T_{max}. Correlations are shown in parentheses. L1 indicates lag-1, L2 indicates lag-2, D1 indicates 24-hour difference, and D2 indicates 48-hour difference.

	DJF	MAM	JJA	SON
1	THICK (0.77)	THICK (0.83)	THICK (0.69)	THICK (0.86)
2	THICK L1 (0.70)	Z_{500} (0.80)	Z_{500} (0.66)	Z_{500} (0.83)
3	Q_{850} (0.68)	THICK L1 (0.79)	THICK L1 (0.63)	THICK L1 (0.83)
4	Z_{500} (0.68)	Z500 L1 (0.77)	Z_{500} L1 (0.60)	Z_{500} L1 (0.82)
5	Z_{500} L1 (0.64)	Q_{850} (0.76)	Q_{850} (0.59)	Q_{850} (0.75)
6	Q_{850} L1 (0.58)	THICK L2 (0.71)	Q_{850} L1 (0.52)	Z500 L2 (0.75)
7	SLP (-0.58)	Q_{850} L1 (0.70)	THICK L2 (0.51)	THICK L2 (0.75)
8	SLP L1 (-0.57)	Z_{500} L2 (0.70)	Z_{850} (0.49)	Q_{850} L1 (0.72)
9	GEOS L1 (0.57)	Q_{850} L2 (0.64)	Z_{500} L2 (0.48)	Q_{850} L2 (0.67)
10	THICK L2 (0.55)	Z_{850} (0.57)	GEOW L1 (0.48)	Z_{850} L1 (0.61)
11	GEOS (0.54)	Z_{850} L1 (0.56)	Z_{850} L1 (0.47)	SLP L1 (-0.60)
12	SLP L2 (-0.51)	SLP L1 (-0.54)	GEOS L1 (0.42)	Z_{850} (0.60)
13	Z_{500} L2 (0.51)	SLP (-0.50)	GEOW (0.42)	SLP L2 (-0.59)
14	Z_{850} L2 (-0.49)	SLP L2 (-0.50)	GEOW L2 (0.42)	SLP (-0.57)
15	GEOW (0.48)	Z_{850} L2 (0.50)	SLP L1 (-0.39)	Z_{850} L2 (0.55)
16	GEOW L1 (0.47)	GEOS L1 (0.47)	Q_{850} L2 (0.39)	RH_{850} L2 (0.53)
17	GEOZT (0.45)	GEOS L2 (-0.44)	GEOS L2 (0.37)	RH_{850} L1 (0.53)
18	Q_{850} L2 (0.45)	GEOS (0.43)	Z_{850} L2 (0.37)	RH_{850} (0.51)
19	Z_{850} (0.44)	RH_{850} L2 (-0.41)	SLP L2 (-0.37)	GEOS L1 (-0.49)
20	Z_{850} L1 (-0.44)	GEOW L1 (0.39)	GEOZT L1 (0.36)	GEOWS (-0.48)

Table 5.3. Predictor variable hierarchy for T_{min} at a central station (Anderson, IN). The table shows the twenty variables which exhibit the strongest correlation with T_{min}. Correlations are shown in parentheses. L1 indicates lag-1, L2 indicates lag-2, D1 indicates 24-hour difference, and D2 indicates 48-hour difference.

	DJF	MAM	JJA	SON
1	THICK (0.76)	THICK (0.84)	Q_{850} (0.76)	Q_{850} (0.83)
2	THICK L1 (0.73)	Q_{850} (0.82)	THICK (0.74)	THICK (0.82)
3	Z_{500} (0.69)	Z_{500} (0.81)	THICK L1 (0.66)	THICK L1 (0.80)
4	Q_{850} (0.68)	THICK L1 (0.80)	Z_{500} (0.65)	Z_{500} (0.79)
5	Z_{500} L1 (0.68)	Z_{500} L1 (0.78)	Q_{850} L1 (0.64)	Z_{500} L1 (0.78)
6	Q_{850} L1 (0.61)	Q_{850} L1 (0.76)	Z_{500} L1 (0.63)	Q_{850} L1 (0.77)
7	THICK L2 (0.60)	THICK L2 (0.72)	GEOW L1 (0.53)	THICK L2 (0.71)
8	SLP (-0.57)	Z_{500} L2 (0.71)	GEOS L1 (0.53)	Z_{500} L2 (0.71)
9	Z_{500} L2 (0.57)	SLP L1 (0.67)	GEOW (0.53)	Q_{850} L2 (0.68)
10	SLP L1 (-0.55)	Z_{850} L1 (0.57)	THICK L2 (0.51)	SLP L1 (-0.59)
11	GEOS L1 (0.54)	Z_{850} (0.57)	Z_{500} L2 (0.49)	SLP L2 (-0.58)
12	SLP L2 (-0.51)	SLP L1 (-0.55)	Z_{850} L1 (0.48)	Z_{850} L1 (0.58)
13	Q_{850} L2 (0.50)	SLP L2 (-0.51)	Q_{850} L2 (0.47)	SLP (-0.56)
14	GEOS (-0.50)	Z_{850} L2 (0.51)	GEOZW (-0.46)	Z_{850} (0.56)
15	Z_{850} L2 (-0.48)	SLP (-0.51)	SLP L1 (-0.46)	Z_{850} L2 (0.52)
16	GEOW (0.48)	GEOS L1 (0.50)	GEOZT (-0.46)	RH_{850} L2 (0.50)
17	GEOZT (0.45)	GEOS L2 (-0.45)	GEOW L2 (0.45)	RH_{850} L1 (0.50)
18	GEOW L1 (0.44)	GEOS (0.44)	GEOS L2 (0.45)	GEOWS (-0.49)
19	GEOZW (0.43)	GEOW L2 (0.42)	Z_{850} (0.44)	RH_{850} (0.49)
20	GEOS L2 (0.43)	GEOW (0.38)	Q_{850} D2 (0.43)	GEOS L2 (-0.48)

Table 5.4. Predictor variable hierarchy for precipitation at a central station (Anderson, IN). The table shows the twenty variables which exhibit the strongest correlation with precipitation. Correlations are shown in parentheses. L1 indicates lag-1, L2 indicates lag-2, D1 indicates 24-hour difference, and D2 indicates 48-hour difference.

	DJF	MAM	JJA	SON
1	Q_{850} (0.45)	GEOZT (0.42)	GEOZT (0.34)	GEOZT (0.37)
2	GEOZT (0.42)	GEOZW (0.39)	GEOZW (0.30)	GEOS (0.31)
3	GEOS (0.37)	SLP (-0.33)	GEOW (0.28)	GEOZW (0.30)
4	GEOZW (0.36)	Q_{850} (0.32)	SLP (-0.25)	SLP (-0.29)
5	SLP (-0.34)	GEOS (0.32)	Q_{850} (0.25)	Q_{850} (0.28)
6	GEOS L1 (0.34)	GEOW (0.31)	GEOZT D2 (0.22)	RH_{850} (0.28)
7	GEOZS (0.33)	RH_{850} (0.29)	GEOWS (0.22)	GEOS L1 (0.28)
8	RH_{850} (0.33)	GEOS L1 (0.29)	GEOZT L1 (0.21)	GEOS D1 (-0.27)
9	Q_{850} L1 (0.32)	GEOZT L1 (0.29)	GEOS D1 (-0.21)	GEOW (0.27)
10	THICK (0.30)	GEOZT D2 (0.28)	SLP D1 (-0.21)	GEOS D2 (-0.26)
11	SLP L1 (-0.30)	GEOZS (0.27)	GEOS L1 (0.20)	GEOZT L1 (0.25)
12	Z_{850} (-0.29)	Z_{850} D1 (-0.27)	Z_{850} (-0.20)	SLP D1 (-0.25)
13	Z_{850} L1 (-0.29)	GEOS D1 (-0.27)	GEOS (-0.20)	Z_{850} L1 (-0.24)
14	Z_{500} (0.29)	GEOS D2 (-0.27)	GEOZS (0.19)	Z_{850} (-0.24)
15	RH_{850} D2 (0.28)	SLP D2 (-0.26)	GEOZT D1 (0.19)	GEOZT D2 (0.23)
16	GEOS D2 (-0.28)	Z_{850} D2 (-0.26)	SLP D2 (-0.19)	SLP D2 (-0.23)
17	GEOW L1 (0.28)	Z_{850} (-0.26)	Z_{850} D1 (-0.19)	GEOZS L1 (0.23)
18	GEOZT L1 (0.28)	SLP D1 (-0.26)	GEOW D1 (0.19)	SLP L1 (-0.23)
19	GEOZT D2 (0.28)	SLP L1 (-0.26)	GEOS D2 (-0.19)	Z_{850} L1 (-0.23)
20	GEOS D1 (-0.27)	Z_{500} D2 (0.25)	GEOZW D2 (0.19)	GEOZS (0.22)

For many of the predictor variables listed in Table 5.1, strong correlations with the surface variables occur at both lag-0, lag-1, and in some cases lag-2. However, the lag-1 and lag-2 values provide little, if any, improvement over concurrent predictors. In most cases, the differenced variables exhibit relatively lower correlation with station values of T_{max} and T_{min}. For precipitation, the differenced variables exhibit a similar level of correlation, but still provide only small improvement over concurrent values. These results indicate that lagged and differenced predictors should not be used in the downscaling work presented in Chapter 8.

Throughout this chapter, it has been shown that the highest correlations typically occur at non-local grid points (i.e., away from the station location). These findings provide support for the use of remote grid points for downscaling T_{max}, T_{min} and precipitation. Hence, the downscaling analyses presented in Chapter 8 will include the use of remote grid points.

The results presented in this chapter provide a measure of the statistical associations between the large-scale predictor variables and the surface parameters. The strongest associations for T_{max} and T_{min} are associated with 850-500 hPa thickness, 500-hPa geopotential height, and 850-hPa specific humidity. For precipitation, the strongest associations are with vorticity variables, sea-level pressure, and 850-hPa specific humidity. As shown in Table 5.5, these large-scale variables exhibit some common variability (most notably thickness and geopotential height), but each contributes unique explanation of the variance of the surface predictands.

Table 5.5. Correlations between daily values of the predictor variables (computed using data at all grid points). The values shown in each box are for winter (DJF), spring (MAM), summer (JJA), and fall (SON).

	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	RH_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
Z_{500}	1	0.89	0.96	0.37	0.69	-0.27	0.34	-0.23	-0.48	-0.02	-0.17	-0.15
	1	0.84	0.95	0.22	0.66	-0.32	0.32	-0.03	-0.45	-0.01	-0.15	-0.13
	1	0.85	0.92	0.52	0.60	-0.41	0.28	-0.09	-0.48	-0.06	-0.24	-0.22
	1	0.87	0.95	0.36	0.68	-0.40	0.30	-0.27	-0.54	-0.03	-0.21	-0.20
Z_{850}		1	0.72	0.73	0.49	-0.31	0.32	-0.25	-0.49	-0.23	-0.37	-0.39
		1	0.63	0.69	0.41	-0.34	0.30	-0.07	-0.45	-0.26	-0.39	-0.42
		1	0.58	0.86	0.43	-0.29	0.26	-0.10	-0.43	-0.27	-0.44	-0.47
		1	0.67	0.75	0.45	-0.39	0.25	-0.28	-0.51	-0.27	-0.44	-0.46
THICK			1	0.10	0.74	-0.22	0.32	-0.19	-0.42	0.11	-0.02	0.02
			1	-0.07	0.70	-0.26	0.29	0.00	-0.39	0.14	0.01	0.06
			1	0.17	0.61	-0.42	0.24	-0.07	-0.42	0.11	-0.04	-0.00
			1	0.07	0.72	-0.35	0.29	-0.23	-0.48	0.11	-0.05	-0.01
SLP				1	-0.06	-0.28	0.13	-0.22	-0.30	-0.50	-0.58	-0.66
				1	-0.16	-0.25	0.08	-0.15	-0.20	-0.52	-0.58	-0.67
				1	0.10	-0.18	0.12	-0.14	-0.29	-0.46	-0.58	-0.66
				1	-0.08	-0.28	0.06	-0.21	-0.27	-0.51	-0.60	-0.69
Q_{850}					1	0.19	0.24	-0.08	-0.29	0.19	0.08	0.13
					1	0.23	0.26	0.03	-0.26	0.21	0.13	0.18
					1	0.21	0.29	-0.01	-0.27	0.12	0.10	0.13
					1	0.16	0.22	-0.17	-0.35	0.19	0.09	0.14
RH_{850}						1	-0.06	0.02	0.20	0.19	0.26	0.29
						1	-0.07	-0.07	0.23	0.19	0.27	0.30
						1	-0.01	0.03	0.23	0.04	0.24	0.22
						1	-0.12	0.04	0.24	0.13	0.27	0.28
GEOS							1	-0.11	-0.27	0.03	-0.03	-0.01
							1	0.00	-0.14	0.03	-0.02	-0.01
							1	0.02	-0.00	-0.02	-0.01	-0.01
							1	-0.04	-0.14	0.03	-0.01	0.00
GEOW								1	0.21	0.03	0.00	0..01
								1	-0.03	0.04	-0.02	-0.01
								1	0.30	0.05	-0.04	-0.01
								1	0.28	0.04	-0.03	-0.01
GEOWS									1	0.06	0.14	0.14
									1	0.04	0.12	0.12
									1	0.04	0.12	0.11
									1	0.05	0.15	0.14
GEOZS										1	0.28	0.56
										1	0.26	0.56
										1	0.25	0.53
										1	0.27	0.56
GEOZW											1	0.95
											1	0.94
											1	0.95
											1	0.95
GEOZT												1
												1
												1
												1

CHAPTER 6: EVALUATION OF GCM SIMULATIONS

6.1 Introduction

One of the primary assumptions behind GCM downscaling is that the GCM can satisfactorily simulate the predictor variables of interest. All regional climate scenarios implicitly or explicitly rely upon this assumption (Mitchell and Hulme 1999). Hence it important to consider the accuracy with which GCMs represent the predictor variables to be used as input to the downscaling procedures. This evaluation allows an assessment of the GCMs suitability for future climate prognoses and may preclude certain variables from inclusion in empirical downscaling models. In this dissertation, two GCMs (see Chapter 2) are evaluated for use in empirical downscaling studies using multiple methodologies. The variables used for the evaluation are those examined in Chapter 5.

First, each of the simulated variables is compared to the reanalysis data (ECMWF for Q_{850}, NCEP/NCAR otherwise) in terms of the means and variances of the fields during the period in which the reanalysis data and GCM simulations overlap (1990-2001) (Section 6.2). Secondly, GCM data from individual grid points in the study region are directly compared to the reanalysis data (Section 6.3). The final evaluation, included in Chapter 7 in the form of a manuscript, is an evaluation of the GCMs in terms of their ability to reproduce fundamental hemispheric phenomena (modes of variability), such as the Pacific/North American teleconnection (PNA) and North Atlantic Oscillation (NAO) and their links to atmospheric circulation in the study region.

6.2 Evaluation of GCM simulation mean fields (1990-2001)

To evaluate and compare the mean fields of the observed (reanalysis) and GCM-simulated data during the period of concurrent data availability (1990-2001), Taylor diagrams (see Section 3.5) are constructed on a seasonal basis for each of the variables described in Chapter 5.

The similarity between the observed and GCM-simulated mean fields is characterized in Figure 6.1. As the figure shows, the mean Z_{500} fields from both models are very similar to the mean fields from the reanalysis data, resulting in a high pattern correlation (>0.95 in each season for both models). The variability in the mean Z_{500} field is slightly underestimated by HadCM3 during winter, spring, and autumn and CGCM2 during winter, summer, and autumn, resulting in slight displacement of the points from the origin (o) (i.e., normalized standard deviation <1, Figure 6.1). Geopotential heights closer to the surface (Z_{850}) are also simulated well, although pattern correlations are slightly lower than those for Z_{500} (still >0.95 in each season for both models). The mean HadCM3 Z_{850} pattern exhibits less spatial variability than observed in all seasons, while the mean CGCM2 Z_{850} pattern slightly overestimates the spatial variability during spring. Mean 850-500 hPa layer thickness fields from both models are highly correlated with the observed mean field, although the variability in the field in underestimated (during all seasons by CGCM2 and during winter by HadCM3).

The variability in the mean Q_{850} field from CGCM2 matches observed variability well. However, while pattern correlations between simulated and observed mean fields are generally high (r>0.90 during winter r>0.80 during spring and autumn), during the summer the correlation between the two fields is only 0.57 (Figure 6.1c). Seasonal

correlations for the HadCM3 humidity variable, RH_{850}, are lowest during winter (0.83) and highest during summer (0.93), with slight overestimation of the spatial variability in each season.

Correlations are much lower for the mean SLP field, especially for HadCM3 during autumn (r=0.41; Figure 6.1d). For CGCM2, correlations are high (>0.90) during summer and autumn, but lower during winter (0.72) and spring (0.65), although the ratio of modeled to observed standard deviation in HadCM3 is close to one in all seasons (Figure 6.1). CGCM2 overestimates the variability in the mean pattern during all seasons, while HadCM3 overestimates the variability during spring and summer, but underestimates the variability during winter. Although the analysis of HadCM3 included in Covey et al. (2003) and included in the IPCC TAR (IPCC 2001) encompasses global data over a longer simulation period, the HadCM3 SLP evaluation summarized in Figure 6.1 strongly resembles the results from these previous global analyses.

The geostrophic flow variables generally show similar agreement as the SLP fields from which they are derived. The correlations between the mean fields of the southerly component of the geostrophic flow (GEOS), are lower for both models during winter and autumn. The variability in the mean GEOS field is overestimated by HadCM3 in all seasons, with the largest overestimation in summer and autumn (Figure 6.1c, d). CGCM2 generally captures the spatial variability in the mean SLP field better than HadCM3, although the spatial variability is also largely overestimated during summer (Figure 6.1c). The mean fields of the westerly component of the geostrophic flow (GEOW) from CGCM2 exhibit low correlations with the observed field during spring (r=0.47), but greater than 0.70 during the other seasons. During all seasons,

CGCM2 overestimates the spatial variability in the mean field (Figure 6.1). HadCM3 exhibits much higher correlations (>0.80 during winter, spring, and autumn and >0.65 during summer) and the variability in the GEOW field is better simulated, but still overestimated during summer (Figure 6.1c). The mean field of the strength of the resultant geostrophic flow (GEOWS) is well simulated by both models with high correlations (>0.90 during winter, spring, and autumn, and greater than 0.75 during summer) and only slight differences in variability in each season.

The poorest agreement between observed and simulated mean fields is associated with the geostrophic vorticity variables (GEOZS, GEOZW, and GEOZTOT). The mean southerly shear vorticity field (GEOZS) from both models exhibits relatively low correlation with the observed field (r ranges from 0.24 to 0.44 for CGCM2 and from 0.39 to 0.55 for HadCM3). CGCM2 underestimates the variability in the field during each season (Figure 6.1), and by nearly 50% during spring. Conversely, HadCM3 overestimates the spatial variability in the mean GEOZS field by approximately 100%. Correlations between the mean HadCM3 GEOZW field and the observed mean GEOZW field are generally in the 0.60-0.70 range, but are slightly higher during spring (0.77). However, during all seasons, the variability in this field is overestimated. During spring and summer, the mean GEOZW field is more than twice as variable as the observed field. Correlations between the modeled and observed fields are slightly lower for CGCM2. While CGCM2 also overestimates the spatial variability in the field during spring and summer, the magnitude of the differences is smaller, resulting in a smaller RMS error (i.e., shorter distance to the origin). Pattern correlations between modeled and observed GEOZT fields are slightly higher for HadCM3 than CGCM2, although values for both

models are low relative to most of the other variables tested. However, HadCM3 also largely overestimates the spatial variability in the mean pattern during spring and summer, with more than twice the level of spatial variability as observed. CGCM2 exhibits similar behavior, although the level of overestimation is much lower.

The results presented here have broad implications for the downscaling work presented in Chapter 8. The Taylor diagrams presented in this section imply:

1) For both models, the variables exhibiting the closest agreement with reanalysis data (in terms of the 1990-2001 mean fields) are the large-scale circulation data (i.e., the geopotential height variables) and 850-500 hPa layer thickness.

2) The mean field of 850-hPa specific humidity is substantially more accurate than the 850-hPa relative humidity field (excepting summer)

3) The derived geostrophic flow and vorticity variables exhibit the lowest level of agreement with observations, with particularly poor correspondence during the summer.

Coupled with the statistical associations presented in Chapter 5, these results suggest that difficulties associated with precipitation downscaling will likely be exacerbated by poor agreement between GCM simulations and observations of the key predictor variables. Conversely, the variables exhibiting the strongest associations with the surface temperature variables are among the most well-simulated in terms of the 1990-2001 large-scale mean fields.

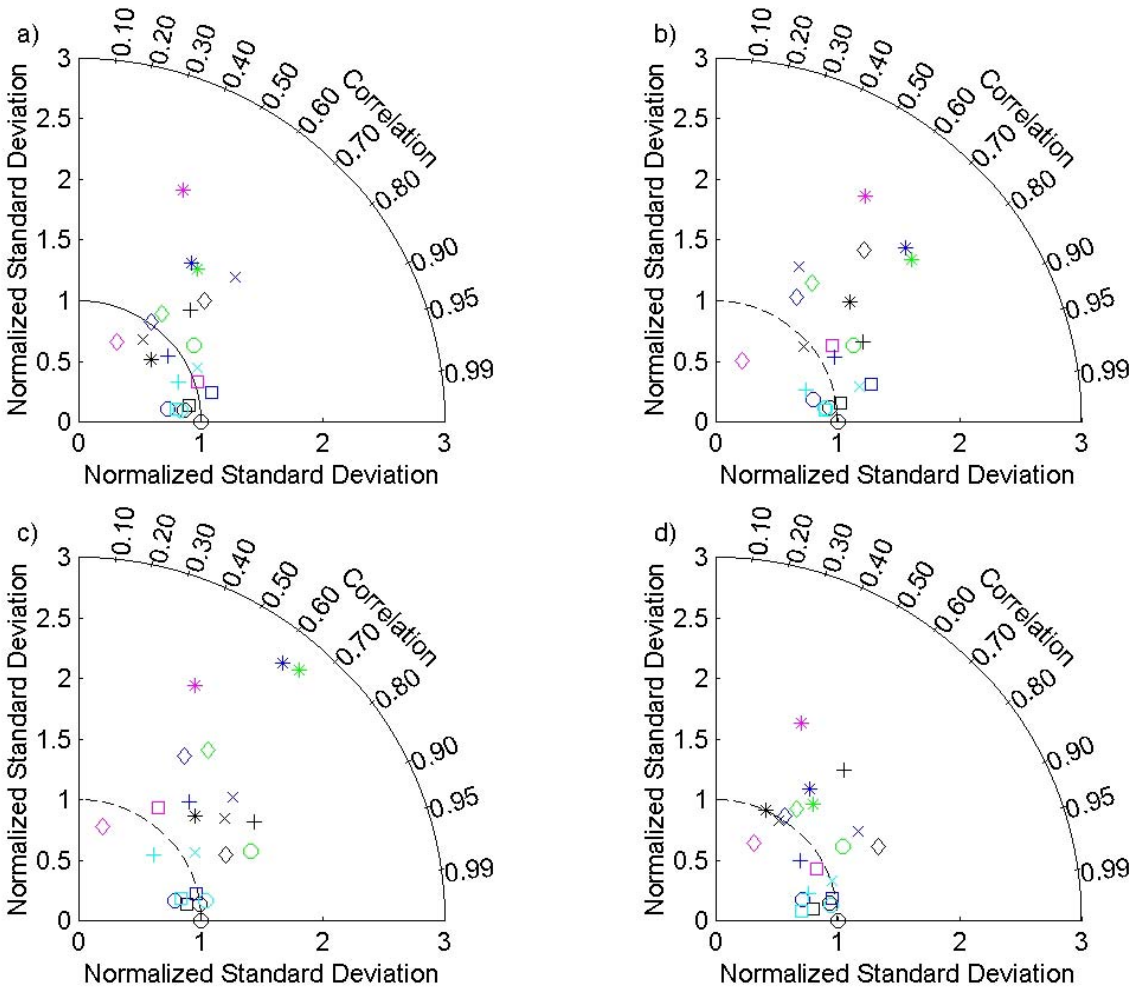

Figure 6.1. Taylor diagram for 1990-2001 seasonal means of GCM-simulated variables relative to reanalysis data: a) winter (DJF), b) spring (MAM), c) summer (JJA), and d) autumn (SON). Symbols are as follows: HadCM3 Z_{500} (o), CGCM2 Z_{500} (□), HadCM3 Z_{850} (o), CGCM Z_{850} (□), HadCM3 THICK (o), CGCM2 THICK (□), HadCM3 RH_{850} (o), CGCM2 Q_{850} (□), HadCM3 SLP (∗), CGCM2 SLP (◊), HadCM3 GEOS (+), CGCM2 GEOS (x), HadCM3 GEOW (+), CGCM2 GEOW (x), HadCM3 GEOWS (+), CGCM2 GEOWS (x), HadCM3 GEOZS (∗), CGCM2 GEOZS (◊), HadCM3 GEOZW (∗), CGCM2 GEOZW (◊), HadCM3 GEOZT (∗), and CGCM2 GEOZT (◊).

6.3 Evaluation of GCM grid point data

The method proposed by Chervin (1981) and later used and refined by Portman et al. (1992) (see Section 3.4.1) is employed to compare observed and GCM simulated variables. Following their approach, 100 expanded datasets were constructed from the

reanalysis data and compared to the GCM output using a bootstrap resampling technique (Section 3.4). First, 12 years were randomly chosen from the reanalysis (ECMWF for Q_{850}, otherwise NCEP/NCAR) time series (1958-2001). For each GCM and variable (Z_{500}, Z_{850}, THICK, Q_{850}, RH_{850}, SLP, and the derived geostrophic flow and vorticity variables), two test variates, r_1 and r_2 (reflecting agreement between observed and simulated means and variances, respectively (see Section 3.4.1)), were computed for each grid box in the study area for each of the 100 random samples of 12-years from the reanalysis. The null hypotheses that the reanalysis data (bootstrap samples) and GCMs (first 12 years) have the same means and variances can be rejected if a sufficient number of r_1 and r_2 values lie outside of the specified confidence intervals. To examine a range of temporal scales, separate analyses are conducted at annual, seasonal, and monthly time steps. It should be noted that 12 years represents a relatively small amount of data for model evaluation and precludes analysis of grid point GCM data at higher (e.g., daily) time steps. This 12-year evaluation period is dictated by the short period of overlap between the transient GCM and reanalysis data (1990-2001).

Comparison of the 1990-2001 GCM simulations with the 1958-2001 reanalysis data is based on the assumption that the 1958-2001 reanalysis data are also representative of the 1990-2001 period. To test this assumption, 12 years were randomly chosen (100 times) from the 1958-2001 reanalysis data and compared to the 1990-2001 data using the r_1 and r_2 statistics. For each variable, r_1 and r_2 values were within the confidence intervals for each random sample and annual, seasonal, and monthly timescales.

6.3.1 GCM evaluation at annual timescales

When evaluated at the annual timescale, the means and variances of the GCM output from both models are not found to be significantly different from the reanalysis data at most grid points for most variables (excepting the geostrophic vorticity variables as simulated by HadCM3). Table 6.1 shows the number of grid points (out of 435) for which more than 5% (5 out of 100) of the bootstrap samples produce values of the two test variates, r_1 and r_2, outside of the specified confidence interval (-0.64 to 0.64 for r_1 and 0.59 to 1.68 for r_2, given a sample size of 12).

As the table shows, both r_1 and r_2 values for HadCM3 Z_{500} and Z_{850} are within the confidence bounds for more than 95% of bootstrap samples at every grid point. While CGCM2 produces similar results for Z_{850}, the annual means of Z_{500} in the extreme southeastern part of the study are low relative to the reanalysis data, as evidenced by r_1 lying below the lower confidence bound for all 100 bootstrap samples (Figure 6.2a). The results for 850-500 hPa thickness (THICK) are nearly identical to those for Z_{500} due to the high degree of co-variance between the two variables (see Table 5.5).

Table 6.1. Number of grid points (out of 435) for which less than 95% of r_1 and r_2 values lie within the specified confidence interval at the annual timescale.

	CGCM2										
	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
r_1	48	0	57	0	17	1	0	0	12	2	3
r_2	1	0	0	0	34	0	8	0	0	11	3
	HadCM3										
	Z_{500}	Z_{850}	THICK	SLP	RH_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
r_1	0	0	0	23	14	15	0	9	22	21	22
r_2	0	0	0	0	2	0	0	0	416	414	409

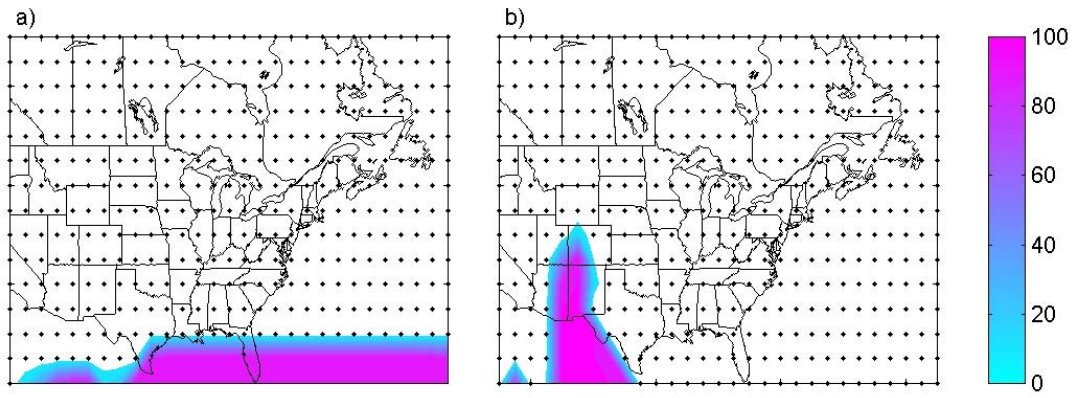

Figure 6.2. Number of cases (out of 100) for which r_1 lies outside of the specified confidence interval: a) CGCM2 Z_{500} and b) HadCM3 SLP. Grid point locations are shown for points which at which more than 95% of r_1 values lie within the confidence bounds.

Annual means of SLP from CGCM2 are within the confidence bounds for all grid points for at least 95% of the bootstrap samples, while those from HadCM3 exhibit a sufficiently low bias relative to the reanalysis to result in 23 grid points which have r_1 values outside of the confidence interval for more than 5% of the samples (Table 6.1, Figure 6.2b). As shown in Figure 6.2b, these grid points are located in the southwest corner of the domain. Both GCMs produce SLP r_2 values within the confidence interval for all grid points and at least 95% of the bootstrap samples. This implies that SLP values at the 23 HadCM3 points mentioned above exhibit the correct amount of interannual variability, but that the annual means are significantly different from those observed.

The geostrophic flow and vorticity variables derived from the SLP fields exhibit varying degrees of agreement with reanalysis data. For CGCM2, only a small number of grid point values differ significantly from the reanalysis in terms of means and standard deviations. These grid points are located in the extreme southwest corner of the domain

away from the primary study region (see Figure 2.1). For HadCM3, the geostrophic flow variables exhibit a similar level of agreement with reanalysis data as CGCM2. However, while the mean grid-point vorticity values are similar to those for SLP, the variances are statistically different over nearly the entire domain (Table 6.1). Examination of the individual grid point variances confirms that HadCM3 overestimates the variability in the geostrophic vorticity variables over nearly the entire domain, consistent with the information conveyed in the Taylor diagrams in the previous section (Figure 6.1). At the annual timescale, the variability at individual grid points is often twice as large as observed.

For the humidity variables (Q_{850} for CGCM2 and RH_{850} for HadCM3) the GCM and observed means differ sufficiently to result in r_1 values consistently outside of the confidence interval for multiple grid points in the western part of the domain (Table 6.1, Figure 6.3a, b). In CGCM2, multiple grid points consistently produce r_2 values outside of the confidence bounds, coinciding spatially with those for r_1 (Figure 6.3c). These results are in slight contrast with the results for the mean fields (Section 6.2), which indicate slightly superior agreement with observations for CGCM2 Q_{850} relative to HadCM3 RH_{850}. Therefore, while CGCM2 Q_{850} simulated more realistically in terms of the mean large-scale field, HadCM3 RH_{850} has a better correspondence with reanalysis data at the grid-point level.

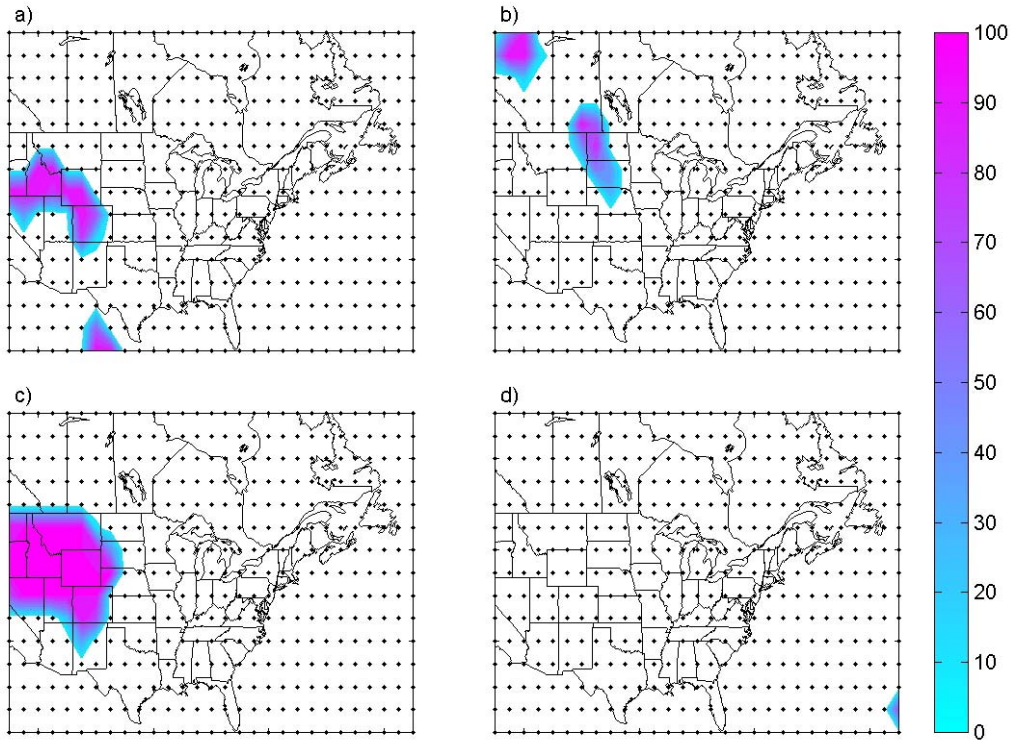

Figure 6.3. Number of cases (out of 100) for which test variates lie outside of the specified confidence interval: a) CGCM2 Q_{850} r_1, b) HadCM3 RH_{850} r_1, c) CGCM2 Q_{850} r_2, and d) HadCM3 RH_{850} r_2. Grid point locations are shown for points which at which more than 95% of r_1 or r_2 values lie within the confidence bounds.

6.3.2 GCM evaluation at seasonal timescales

The performance of HadCM3 Z_{500} at annual timescales is mirrored in each season, with all grid points exhibiting values of r_1 and r_2 within the confidence bounds for at least 95% of the bootstrap samples (Table 6.2). However, as expected, the agreement between the reanalysis data and GCM simulations typically declines as shorter timescales are considered. The disagreement between annual means of CGCM2 Z_{500} in the extreme southeast corner of the domain also translates to the seasonal timescale, with the greatest spatial extent during summer (Figure 6.4). While CGCM2 Z_{500} r_2 values were within the confidence bounds on annual timescales, seasonal analyses show that during winter (and

to a lesser extent, spring), CGCM2 Z_{500} r_2 values are outside of the specified confidence bounds for large areas of the western Atlantic Ocean (Figure 6.5). Examination of the individual r_2 values shows that CGCM2 underestimates the variance of Z_{500} in that region during these seasons.

Table 6.2. Number of grid points (out of 435) for which less than 95% of r_1 and r_2 values lie within the specified confidence interval at the seasonal timescale.

CGCM2 r_1											
	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	67	4	87	12	22	2	14	0	10	15	15
MAM	75	3	97	21	52	6	4	3	14	10	10
JJA	153	36	151	35	109	52	57	78	41	50	44
SON	93	25	87	0	48	2	5	0	16	7	9
CGCM2 r_2											
	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	105	9	65	0	5	0	0	0	3	0	2
MAM	53	0	5	0	22	0	0	0	0	1	0
JJA	0	0	0	7	47	3	13	6	22	54	42
SON	13	0	10	3	51	0	3	0	3	10	0
HadCM3 r_1											
	Z_{500}	Z_{850}	THICK	SLP	R_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	0	3	0	29	48	17	0	3	40	27	24
MAM	0	0	0	24	22	10	0	17	21	51	48
JJA	0	28	78	120	107	80	68	79	84	80	80
SON	0	3	0	47	43	24	5	4	26	13	15
HadCM3 r_2											
	Z_{500}	Z_{850}	THICK	SLP	R_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	0	0	0	0	64	0	6	0	399	317	349
MAM	0	0	0	1	5	0	0	0	424	430	427
JJA	0	0	0	0	21	0	0	9	435	435	435
SON	0	0	0	0	21	0	0	0	402	413	405

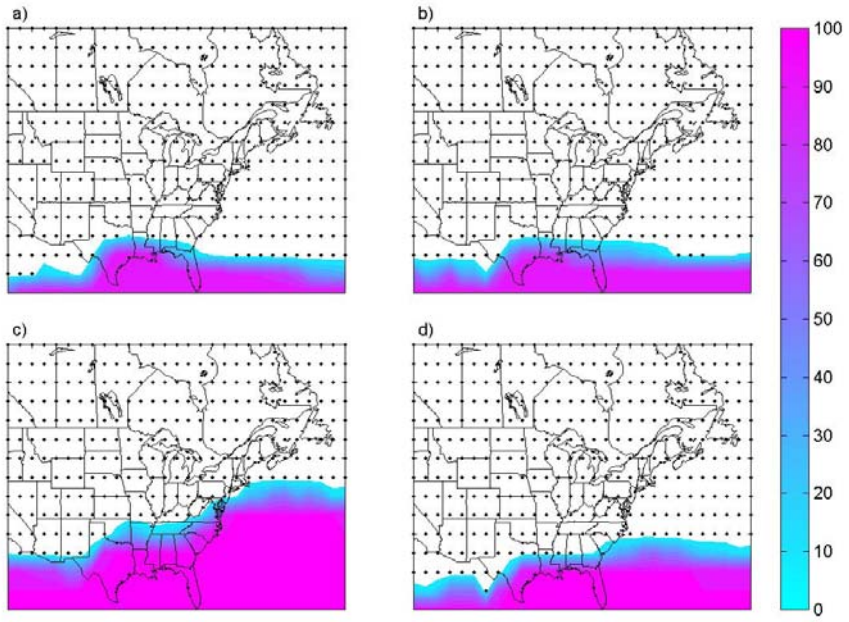

Figure 6.4. Number of cases (out of 100) for which the test variate r_1 lies outside of the specified confidence interval: a) CGCM2 Z_{500} winter, b) CGCM2 Z_{500} spring, c) CGCM2 Z_{500} summer, and d) CGCM2 Z_{500} autumn. Grid point locations are shown for points which at which more than 95% of r_1 values lie within the confidence bounds.

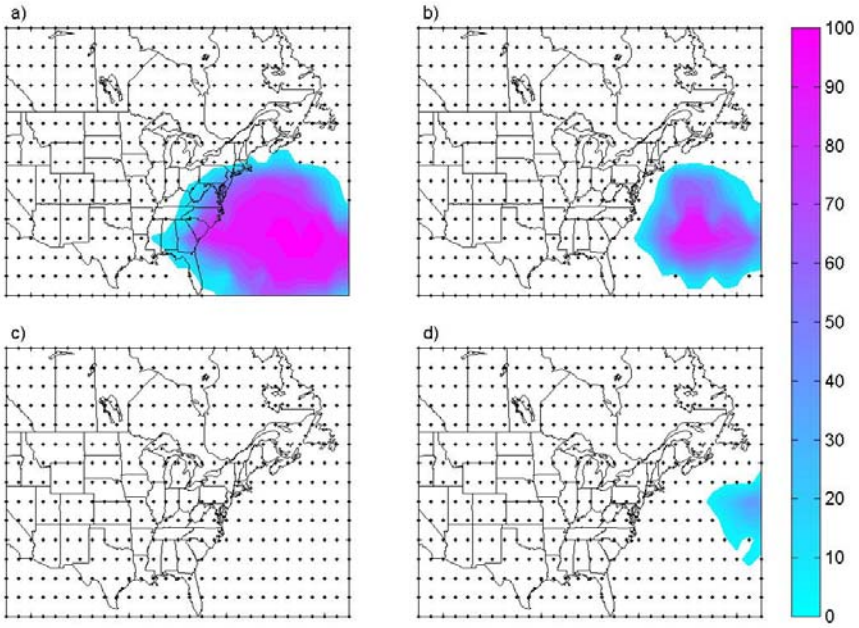

Figure 6.5. Number of cases (out of 100) for which the test variate r_2 lies outside of the specified confidence interval: a) CGCM2 Z_{500} winter, b) CGCM2 Z_{500} spring, c) CGCM2 Z_{500} summer, and d) CGCM2 Z_{500} autumn. Grid point locations are shown for points which at which more than 95% of r_2 values lie within the confidence bounds.

For both GCMs, a small number of grid points have Z_{850} r_1 values that are outside of the confidence interval, particularly during the summer (Table 6.2). For CGCM2, these grid points are in the southeast corner of the domain. For HadCM3, these grid points are located in the southwest corner of the domain. With the exception of a small number of grid points during winter in CGCM2, the data and statistical tests used here do not suggest a difference in the variances of observed and GCM simulated Z_{850} values on seasonal timescales (Table 6.2).

Results for 850-500 hPa thickness are similar to those for Z_{500} and Z_{850}. For CGCM2, the results are nearly identical to those for Z_{500}. For HadCM3, however, summer values of r_1 are outside of the confidence bounds for a large proportion of the bootstrap samples for a large area covering much of the southwest corner of the domain, consistent with the rejection Z_{850} grid points.

The performance of HadCM3 with respect to the mean of SLP (r_1) on the annual time scale is carried over to each season. The area of large differences in observed and GCM mean behavior has the greatest extent during summer (Figure 6.6). The seasonal analyses performed here show that during summer, CGCM2 also has difficulty with respect to mean SLP in the same region. However, HadCM3 mean SLP values in the region of grid-points for which the null hypothesis of equal means was rejected are lower than those from the reanalysis, while CGCM2 mean SLP values at such grid points are higher than those from the reanalysis, suggesting disagreement between HadCM3 and CGCM2 SLP at the grid-point level. Both GCMs simulate the variance of SLP well on seasonal timescales, with nearly all grid points producing values of r_2 within the

confidence bounds for all 100 bootstrap samples in each season (exceptions are a few grid points in the southwest during summer).

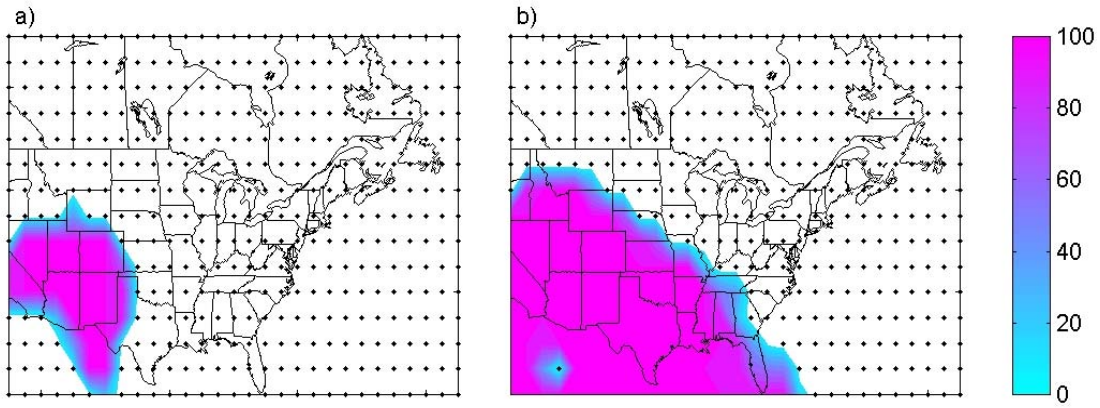

Figure 6.6. Number of cases (out of 100) for which r_1 lies outside of the specified confidence interval during summer: a) CGCM2 SLP and b) HadCM3 SLP. Grid point locations are shown for points which at which more than 95% of r_1 values lie within the confidence bounds.

The seasonal results for the geostrophic flow and vorticity variables are similar to those at the annual timescale. Differences between observed and simulated grid point means and variances of the geostrophic flow variables (GEOS, GEOW, and GEOWS) are not statistically significant at most grid points, with the exception of mean values during the summer. During this season, each GCM has grid points for which the mean values of the flow variables are significantly different from observed mean values (Table 6.2). These points coincide with those for the SLP field from which the flow indices are derived (Figure 6.6). The mean values of the geostrophic vorticity variables are also well simulated by both GCMs, with significant differences at only a small percentage of grid points during each season. Similar results were found for seasonal variances of the vorticity variables as simulated by CGCM2. As with the annual results, the seasonal

variances of HadCM3-derived geostrophic vorticity variables lack agreement with the observed variances (Table 6.2). Taken in concert with the Taylor diagrams presented in Figure 6.1 (showing consistently low pattern correlations and overestimation of the spatial variability in the mean vorticity fields), the results presented here suggest that HadCM3 vorticity values are more variable (both spatially and temporally) than those observed, but that the mean values generally agree on seasonal timescales.

For each season both GCMs have grid points at which humidity simulations (Q_{850} for CGCM2 and RH_{850} for HadCM3) are significantly different from the reanalysis in terms of both the mean and variance (Table 6.2). For both GCMs, the greatest number of poorly simulated grid points occurs during summer (Figure 6.7). As shown in Figure 6.7, GCM2 grid point Q850 r_1 values lie outside of the confidence interval for large areas of the Western United States and over the southeast corner of the domain. The grid points at which HadCM3 means are statistically different from the reanalysis are located in the same general areas, although the spatial patterns are slightly different (Figure 6.7). Regions in which the variances are found to be significantly different from the reanalysis generally coincide with those for which the means are different, although they are less extensive (Figure 6.7).

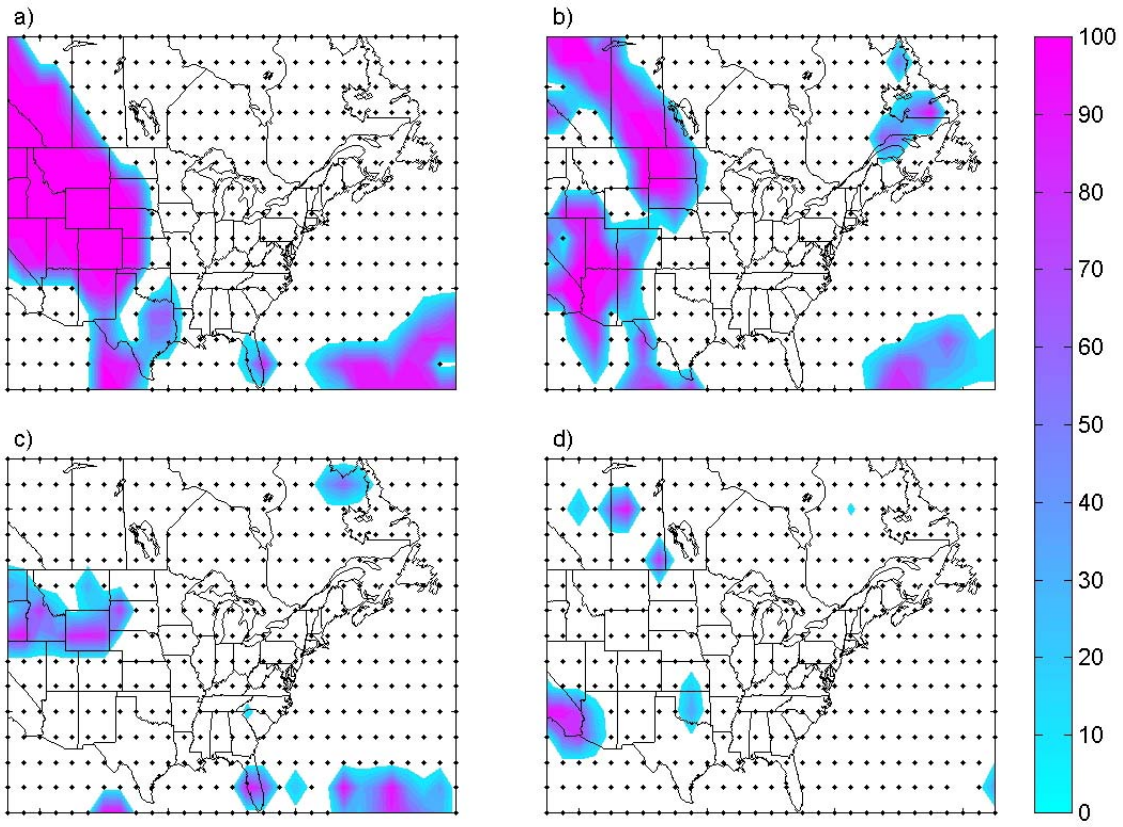

Figure 6.7. Number of cases (out of 100) for which test variates lie outside of the specified confidence interval during the summer (JJA): a) CGCM2 Q_{850} r_1, b) HadCM3 RH_{850} r_1, c) CGCM2 Q_{850} r_2, and d) HadCM3 RH_{850} r_2. Grid point locations are shown for points which at which more than 95% of r_1 or r_2 values lie within the confidence bounds.

6.3.3 GCM evaluation at monthly timescales

Evaluation at the monthly timescale provides greater detail relative to the seasonal and annual analyses presented above. As shown in Table 6.3, larger numbers of grid points have means and variances that are significantly different than those observed when monthly values are examined. As the table shows, the performance of HadCM3 is superior to that of CGCM2 for the geopotential height variables (Z_{500} and Z_{850}) and hence, for 850-500 hPa thickness, as evidenced by much smaller numbers of significantly different grid points. Mean values of CGCM2 Z_{500} are different than observed values for

more than half of the grid points during the months of April through July (Figure 6.8). Interestingly, during the period when mean values of CGCM2 Z_{500} (and hence, THICK) are best simulated, values of the variance are found to be different at a large numbers of grid points (Table 6.3).

With respect to mean SLP values, CGCM2 has slightly fewer significantly different grid points than HadCM3. During July and August, a large proportion of the grid points consistently produce r_1 values outside of the confidence intervals for HadCM3 (and CGCM2 to a lesser extent), with patterns similar to the seasonal patterns shown in Figure 6.6. The results of the statistical tests performed here do not result in rejection of the null hypothesis of equal variances of SLP at most of the grid points (Table 6.3).

Monthly means and variances of CGCM2 simulated geostrophic flow and vorticity variables are within the confidence limits for more than 95% of the bootstrap samples for most of the grid points tested. For each of these variables, the grid points at which means and variances are found to be significantly different are located in the areas where SLP means and variances are significantly different at seasonal timescales (Figure 6.6). HadCM3 means and variances are also within the confidence bounds for most points when the geostrophic flow variables are considered. However, the variances of the geostrophic vorticity variables are too high for nearly all grid points throughout the year (Table 6.3), consistent with the annual and seasonal results presented above.

Monthly results for GCM-simulated humidity variables are similar to those for the seasonal analyses presented in Section 6.2.2. For both GCMs, the largest number of significant differences occurs during the summer months, with spatial patterns similar to those in Figure 6.7.

Table 6.3. Number of grid points (out of 435) for which less than 95% of r_1 and r_2 values lie within the specified confidence interval at the monthly timescale.

CGCM2 r_1											
	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
JAN	79	5	139	21	17	3	19	11	20	20	15
FEB	76	5	95	47	26	7	19	0	16	14	18
MAR	76	1	114	42	55	6	39	2	13	3	7
APR	219	48	150	55	74	13	32	9	18	19	21
MAY	301	59	222	13	82	8	9	13	24	27	22
JUN	326	98	242	38	104	46	48	77	36	39	32
JUL	268	131	204	63	137	75	97	103	57	75	76
AUG	181	43	196	42	186	60	81	92	52	71	67
SEP	165	43	206	38	122	35	50	21	37	36	35
OCT	116	58	164	1	106	3	7	0	20	13	22
NOV	99	31	174	40	108	13	65	2	11	30	30
DEC	79	8	177	92	107	20	71	1	13	36	29
CGCM2 r_2											
	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
JAN	178	70	176	28	15	5	7	2	9	0	1
FEB	121	17	147	8	5	0	0	1	8	9	10
MAR	176	38	177	6	13	2	0	0	3	0	1
APR	96	3	114	0	15	2	0	0	4	0	0
MAY	0	0	2	4	19	0	3	0	16	36	16
JUN	4	0	3	2	36	11	11	10	37	88	51
JUL	1	5	16	22	63	25	36	36	30	80	78
AUG	94	0	122	28	71	48	31	42	70	78	78
SEP	54	0	134	3	81	1	4	0	4	30	25
OCT	71	0	213	0	66	0	0	0	25	17	16
NOV	133	9	221	0	46	0	0	0	30	14	26
DEC	178	34	213	3	32	0	1	0	22	16	13
HadCM3 r_1											
	Z_{500}	Z_{850}	THICK	SLP	RH_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
JAN	73	35	17	66	68	34	4	11	51	39	40
FEB	90	35	3	53	68	18	31	6	39	34	37
MAR	5	0	7	12	34	3	0	9	16	39	40
APR	0	2	0	35	54	10	11	25	27	72	61
MAY	14	8	18	59	110	34	7	39	46	58	59
JUN	0	40	57	96	126	57	52	90	82	64	65
JUL	25	116	118	171	166	121	129	114	119	124	115
AUG	55	39	119	133	149	94	81	84	79	87	86
SEP	4	0	18	84	99	43	39	20	43	40	43
OCT	0	22	0	74	64	30	12	9	27	11	15
NOV	50	46	10	64	64	30	4	8	37	11	16
DEC	20	7	1	41	58	34	1	6	54	25	26
HadCM3 r_2											
	Z_{500}	Z_{850}	THICK	SLP	RH_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
JAN	0	0	0	0	79	0	14	4	394	335	358
FEB	0	0	0	0	81	0	4	4	401	318	355
MAR	0	0	0	0	16	0	0	0	426	422	423
APR	0	0	0	0	17	0	0	2	330	378	375
MAY	0	0	0	0	14	0	0	4	420	430	432
JUN	0	0	0	1	43	7	0	11	434	425	428
JUL	0	0	4	0	82	7	14	31	432	435	435
AUG	0	0	36	1	51	3	3	11	426	435	434
SEP	1	0	7	0	16	0	0	0	355	420	416
OCT	0	0	0	0	19	0	0	2	426	433	434
NOV	9	0	0	0	80	0	10	0	361	342	351
DEC	0	0	0	0	73	0	15	3	400	326	364

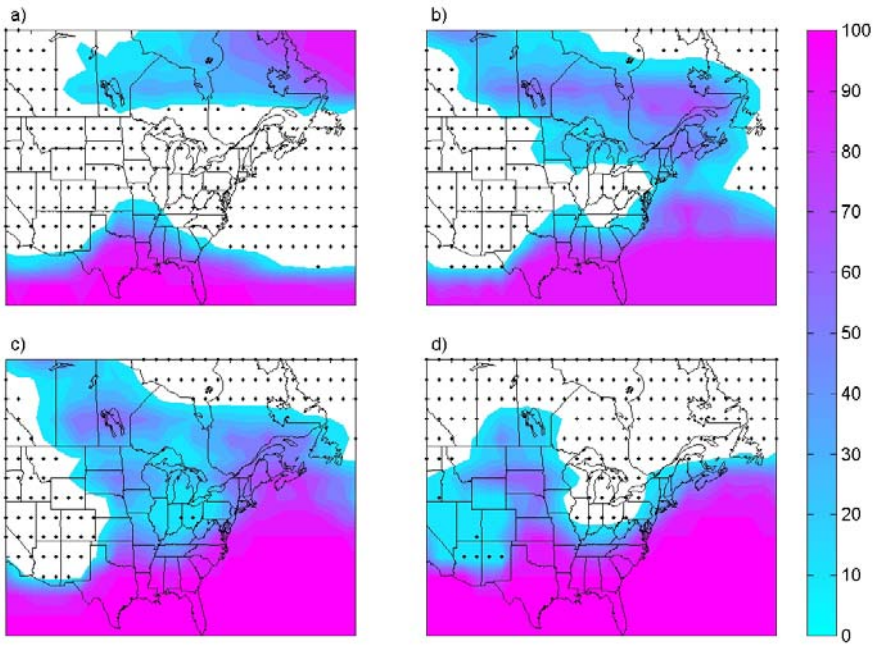

Figure 6.8. Number of cases (out of 100) for which r_1 lies outside of the specified confidence interval for CGCM2 Z_{500}: a) April, b) May, c) June, and d) July. Grid point locations are shown for points which at which more than 95% of r_1 values lie within the confidence bounds.

6.4 Summary

The annual mean fields for the period of overlapping reanalysis data and GCM simulations (1990-2001) have been shown to exhibit close agreement for Z_{500}, Z_{850}, THICK, and Q_{850} and, to a lesser extent, the geostrophic flow variables GEOS, GEOW, and GEOWS. For SLP and RH_{850}, correlations between observed and GCM simulated mean fields were lower, and the spatial variability in the mean pattern is substantially overestimated during some seasons. While the mean fields of the CGCM2 geostrophic vorticity variables exhibit low correlations relative to the other variables, the variability in the mean fields is well simulated. The mean fields of HadCM3 simulated geostrophic vorticity variables exhibit substantially greater spatial variability than observed.

In Section 6.3, it was shown that, for most variables (excepting HadCM3 geostrophic vorticity variables) and grid points, GCM simulated means and variances are not found to be different from reanalysis data on annual time scales. However, as the temporal resolution increases, more GCM grid point means and variances are statistically different from those observed. In particular, CGCM2 Z_{500} means are found to be significantly different for many grid points during the summer months, while variances are found to be significantly different during the winter months. These characteristics of Z_{500} also impact 850-500 hPa thickness (THICK), for which many grid point means and variances are significantly different from those observed during the same time periods. For HadCM3, Z_{500} means and variances are not found to be statistically different nearly as often as those for Z_{850}. Multiple grid points exhibit statistically different means for humidity variables in both GCMs (Q_{850} for CGCM2 and RH_{850} for HadCM3) during the summer. For both GCMs, SLP grid point means and variances are statistically different from reanalysis data less often than the other variables, with the exception of HadCM3 during the summer months. These results do not preclude the use of all grid point values of these variables. However, the results do suggest that some variables should not be used in some regions during particular times of year.

The results of this grid-point evaluation are critical for development of the downscaling tools considered in Chapter 8. HadCM3 generally exhibits closer agreement with reanalysis data than CGCM2 with respect the geopotential height and thickness data. However, CGCM2 performs much better with respect to the geostrophic shear vorticity variables (and specifically with respect to the grid point variances of vorticity). Because these predictors are differentially associated with the surface parameters (e.g.,

temperature is more strongly associated with geopotential height while precipitation is more strongly associated with vorticity), these results imply that grid-point based downscaling models of temperature variables should be more reliable when driven by output from HadCM3, while those for precipitation should be more reliable from CGCM2. To overcome this potential bias, an innovative MOS-based downscaling approach is employed (see Chapter 8 for details).

The analysis of simulations at the synoptic scale, presented in Chapter 7, indicates that circulation regimes at 500 hPa as simulated by HadCM3 and CGCM2 for 1900-2001 show a high degree of correspondence with those manifest in the NNR in terms of frequencies, persistence and progression. However, HadCM3 overestimates the frequency of the most prevalent map type during all seasons. Additional discrepancies between mean frequencies of the NNR and HadCM3 classifications result from different classification of days with weak anticyclonic characteristics. CGCM2 generally simulates the mean frequencies of the NNR map types better than HadCM3, a surprising result given the poor performance of CGCM2 Z_{500} on short time scales. Some, but not all, of the shortcomings of the GCM-based classifications can be explained by differences in the observed and simulated relationships between the map types and two teleconnection indices. Differences between the mean frequencies of observed and GCM map types are larger than changes in the map type frequencies as manifest in the historical reanalysis data or prognostic changes in the mean map type frequencies as manifest in the 2030-2041 GCM simulations.

Taken in concert, these results support earlier findings that GCMs exhibit superior agreement with observations over multiple grid points and longer time scales.

Additionally, both models are capable of reproducing grid point means and variances for some variables and grid points across a range of timescales, and synoptic-scale map patterns during most seasons. The results presented herein would appear to preclude the use of vorticity components in downscaling based on the output of HadCM3 even using MOS-based approaches due to the apparent overestimation in the variability of these derived variables. Potential reasons for the discrepancies between performance of the two GCMs in terms of grid averaged and grid cell specific analyses are many-fold, but include:

1) As described in Chapter 2, HadCM3 does not include flux adjustments while CGCM2 does. This term refers to adjustment terms applied to the atmosphere-ocean fluxes of heat and moisture before these fluxes are imposed on the modeled ocean and atmosphere. Flux adjustments have been employed to avoid the problem of coupled GCMs drifting into unrealistic climate states are particularly critical to the correct simulation of sea-ice. The most recent IPCC report has suggested that their use is likely to decrease in the future, due to the improved ability of non-flux adjusted models to simulate heat balances over the oceans. Nevertheless, the non-flux adjusted model HadCM3 exhibits excess warming of the northern North Atlantic (east of Iceland) that may be manifest as unrealistic simulation of the NAO and other teleconnection patterns. Conversely, the improved ability of HadCM3 relative to CGCM2 to represent humidity and geopotential height variables in the study region may reflect the fact that some flux-adjusted models have been demonstrated to exhibit unrealistic sea-surface temperature and circulation in the tropical oceans.

2) HadCM3 is a Cartesian model while CGCM2 is spectral. Hence, there may be differences in their ability to represent wave features in the atmosphere that are key to vorticity.

6.5 Discussion

The results presented in the previous and current chapters provide a framework for choosing downscaling predictors. In Chapter 5, it was concluded that the variables exhibiting the strongest relationships with T_{max} and T_{min} (in all seasons) are 850-500 hPa thickness (THICK), 500-hPa geopotential height (Z_{500}), and 850-hPa specific humidity (Q_{850}), sea-level pressure (SLP), the southerly and westerly components of the geostrophic flow (GEOS and GEOW), and 850-hPa geopotential height (Z_{850}). For precipitation, the most highly correlated predictors are the geostrophic vorticity variables (GEOZS, GEOZW, and GEOZT), 850-hPa specific humidity (Q_{850}), the southerly and westerly components of the geostrophic flow (GEOS and GEOW), and sea-level pressure (SLP).

The results presented in the current chapter suggest that, for many variables, grid point means and variances are not in agreement with observations and hence, better downscaling models will result from either the MOS-based downscaling approach or by using GCM information on larger scales. The evaluation of synoptic phenomena presented in Chapter 7 shows that the latter approach improves the level of agreement between observed and simulated Z_{500} values. The regional climate change scenarios developed in Chapter 8 will therefore employ synoptic circulation methods, which

incorporate GCM information on larger scales, and the MOS-based downscaling approach. Detailed descriptions of the downscaling models are presented in Chapter 8.

CHAPTER 7: EVALUATION OF GCM SIMULATIONS OF SYNOPTIC PHENOMENA

Schoof JT, Pryor SC. 2004. An evaluation of two GCMs: Simulations of North American teleconnections and synoptic phenomena. In preparation for submission to *Climate Research*

Abstract

We evaluate the ability of two coupled atmospheric-oceanic GCMs (the Hadley Center's 3rd generation coupled climate model (HadCM3) and the Canadian Center for Climate Modeling and Analysis 2nd generation coupled model (CGCM2)) to simulate the North Atlantic Oscillation (NAO), the Pacific North American teleconnection pattern (PNA), and synoptic scale variability in the Midwest region of the United States relative to NCEP/NCAR reanalysis (NNR) data. The 'observed' (NNR-derived) and modeled (GCM-derived) probability distributions of the daily NAO index show agreement over the reference period (1990-2001), while the probability distributions of the daily PNA indices are statistically different during winter, but not consistently different during the spring and fall. Both GCMs successfully reproduce the range of synoptic-scale map-patterns over the study region as manifest in 500 hPa height fields. During the reference period (1990-2001), observed and modeled map types are similar in terms of frequency, coherence, persistence, and progression, although the most common map type occurs too often in HadCM3 relative to NNR. Additional discrepancies between the NNR and HadCM3 classifications result from different classification of days with weak anticyclonic characteristics. Despite the relatively high degree of correspondence

between the observed and simulated teleconnection indices and synoptic types in the study area, differences between the GCM and NNR derived map type frequencies in the period 1990-2001 are greater than either (1) differences in NNR data between 1953-1964 and 1990-2001 or (2) differences in the GCM simulations between 1990-2001 and 2030-2041, indicating that changes in these phenomena over recent and approaching decades are of insufficient magnitude relative to model uncertainty to be definitively identified.

1 Introduction

1.1 Background

A key component of the IPCC Third Assessment Report (TAR; IPCC 2001) was evaluation of the suitability of coupled atmospheric-oceanic global climate models (AOGCMs, referred to herein as GCMs) for use in climate change projection and detection and attribution studies. Within this assessment, several GCMs were systematically inter-compared and evaluated against reanalysis data. The results suggest indicate that coupled models can provide credible simulations of the present annual mean climate and the climatological seasonal cycle over broad continental scales for most variables of interest for climate change and are suitable tools to provide climate projections for the future (McAveney et al. 2001). However, the IPCC summary and other GCM evaluation studies (e.g,. CMIP, Covey et al. 2003) are largely focused on spatial scales above sub-continental and temporal scales from seasonal to decadal. Due to the relatively coarse spatial resolution of GCMs, development of regional-scale climate prognoses typically requires downscaling tools (either Regional Climate Models (RCM) - e.g. Pan et al. (2001) or empirical models - e.g. Wilby et al. (1998)), which rely upon

accurate simulation of synoptic and larger scale climate phenomena within GCMs at temporal scales relevant to climate change impacts research (usually shorter than seasonal time scales). If these larger-scale features are not realistically simulated within GCMs at the required timescale, the derived local climates will likewise be invalid. Indeed, Pan et al. (2001) demonstrated in their analysis of dynamically downscaled hydrologic regimes that uncertainties resulting from biases in the GCM boundary conditions for the RCM simulations were comparable to those arising from variations between RCMs. Hence, there is a need to extend GCM evaluation of synoptic phenomena to the daily timescale in order to identify the degree of confidence with which we can view climate prognoses at the spatial scales needed for impacts research, and to identify parameters or processes that are not well simulated in order to infer aspects of the GCMs that may require improvement. While it must be acknowledged that accurate simulations of the mean or variability of a parameter by a GCM need not necessarily reflect an accurate model formulation, it is asserted that deviations from observations can be used to identify potential model weaknesses and to provide uncertainty bounds for use in assessing prognostic changes in those parameters relative to the current day.

In one of the first evaluations of synoptic scale phenomena in GCM simulations McKendry et al. (1995) used a Kirchhofer correlation-based classification to test the ability of the Canadian Climate Center's second-generation atmospheric GCM to reproduce the current climate of Western Canada. They found that the model reproduced the synoptic types, but that seasonal frequencies and their variability were not accurately simulated. More recent comparisons have shown greater robustness of GCMs with respect to the synoptic scale (e.g., Schubert 1998; Lapp et al. 2002), but relatively few

studies have focused on this scale and model evaluation results are regionally and seasonally variable. Hence, an additional detailed investigation of synoptic-scale variability and teleconnection patterns in observed and GCM-simulated climates is presented here and others are being undertaken under the Atmospheric Model Intercomparison Project (AMIP: Gates et al. 1999; Kang et al. 2002) and the Coupled Model Intercomparison Project (CMIP: Stephenson and Pavan, 2003).

1.2 Objectives

The primary objective of this study is to evaluate the simulation of synoptic-scale climate within two GCMs (the Hadley Center's 3rd generation coupled climate model: HadCM3 (Gordon et al. 2000; Pope et al. 2000) and the Canadian Center for Climate Modeling and Analysis 2nd generation coupled climate model: CGCM2 (Flato et al. 2000; Flato and Boer 2001)) relative to the NCEP/NCAR reanalysis (NNR) data set (Kalnay et al. 1996; Kistler et al. 2001) over the Midwest region of the United States. An additional objective is to examine prognostic climates generated by HadCM3 and CGCM2 for approaching decades relative to changes documented over the latter portion of the observed climate record. This comparison provides a mechanism for determining the confidence with which we can view prognostic climates for this region as derived from these GCMs.

Our diagnostics of HadCM3 and CGCM2 are performed at two spatial scales; (i) large-scale teleconnection indices and (ii) synoptic scale circulation patterns (map types) at 500 hPa over the study region. We also examine the linkages between these two spatial scales by evaluating the relative frequencies of the dominant map types during

different phases of the teleconnections. The evaluation of both GCMs is conducted for a 'reference period' (1990-2001) relative to the NNR, and observed discrepancies between the GCM derived and NNR data are considered in the context of changes over recent decades as manifest in the NNR and potential changes in a future time window as simulated by the GCMs.

1.3 Synoptic climate of the Midwestern USA and links to hemispheric teleconnection indices

Here we present analyses of the ability of two GCMs to reproduce the synoptic scale climate of the Midwestern USA (Figure 1). The synoptic scale climate of this region is characterized by high day-to-day variability associated with both the behavior of the polar jet stream and the relative intensity of several semi-permanent pressure systems, including the subtropical (Bermuda) high (often located within or to the southeast of the study area) and the sub-polar (Hudson Bay) low (often located within or to the northeast of the study area). Variations in the intensity of these pressure systems and the tracking and intensity of synoptic scale phenomena in the study region have been linked to larger pressure oscillations (or teleconnection patterns), specifically the North Atlantic Oscillation (NAO) and Pacific North American (PNA) pattern. Both the NAO and PNA are prominent hemispheric-scale modes of climate variability that are strongly linked to surface temperature and precipitation in the study region at least partly via synoptic scale circulation patterns. For example, the NAO accounts for 31% of the variance in mid-latitude winter mean surface air temperatures (Hurrell 1996) and the PNA is closely linked with both thermal and hydrologic regimes within the study area across a range of

timescales (e.g., Coleman and Rodgers 2003), including daily (Sheridan 2003), and accounts for as much as 40% of the variance in surface air temperature during winter season (Leathers et al. 1991).

The NAO and PNA teleconnection indices describe patterns of mean sea level pressure (MSLP) and tropospheric height variability, respectively, and hence represent the steering patterns of synoptic scale phenomena (Leathers et al. 1991; Hurrell 1995). They are used here both as an evaluation tool for GCMs and to explain differences in the synoptic scale climate in the study region across time or data sets. Both the NAO and PNA exhibit variability on a range of temporal scales. Stephenson and Pavan (2003) show that the NAO signal is dominated by short-term (i.e., year to year) variations and state that more than 70% of the NAO variance is explained by fluctuations with periods of less than a decade. Likewise, the PNA teleconnection varies on all timescales from days to decades (Blackmon et al. 1984; Yarnal and Leathers, 1988).

The NAO describes a redistribution of atmospheric mass between the Arctic and the subtropical Atlantic, typically characterized by winter sea-level pressure anomalies associated with the Azores high and the Icelandic low (Walker, 1924; Walker and Bliss, 1932). During the positive phase of the NAO, both of these pressure centers are stronger than average and shifted to the northeast (Glowienka-Hense 1990), resulting in stronger westerlies in the mid-latitude Atlantic Ocean. Although it is centered east of North America, the positive phase of the NAO is generally associated with higher temperatures and increased moisture content in the Midwestern USA in response to the strengthening of the flow around the subtropical Atlantic high generating southerly flow over the eastern USA (Dickson and Namias 1976; Yarnal and Leathers 1988; Yin 1994).

Coupled GCMs generally simulate the NAO quite well (IPCC 2001) and have been shown to reproduce the slight serial correlation present in observations at an annual time scale ($r < 0.3$ for successive winters) (Stephenson and Pavan 2003). However, several models have been shown to exhibit unrealistic monotonic trends in the NAO or overly strong correlations with the El Niño-Southern Oscillation (ENSO) (Stephenson and Pavan 2003). Previous analyses of NAO simulation by GCMs have focused primarily on the spatial representation of the phenomenon and temporal behavior over interannual and longer timescales (Osborn et al. 1999; Stephensen and Pavan 2003). Here we extend this research by focusing on shorter timescales.

The mean flow over the Pacific-North American sector is characterized by a trough in the east-central North Pacific, a ridge over the Rocky Mountains, and a trough over eastern North America (Leathers et al 1991). The PNA teleconnection index reflects deviations from this mean flow, suggesting more meridional (positive phase) or zonal (negative phase) flow over North America. During the positive phase, negative geopotential height anomalies are located south of Alaska and in the southeastern United States, while positive geopotential height anomalies are located near Hawaii and over the Rocky Mountains resulting in meridional flow over North America. During the negative PNA phase, the anomalies at the pressure centers are reversed and the flow over North America is more zonal.

Comparatively few studies have quantified the ability of GCMs to accurately simulate the PNA. Several atmospheric GCMs have been shown to underestimate the amplitude of the PNA pattern associated with the 1997-98 El Niño event (Kang et al. 2002), although Renshaw et al. (1998) showed that HadAM3, the atmospheric component

of HadCM3, correctly reproduces the changes in the frequency distribution of the PNA index associated with the phases of ENSO.

2 Data

2.1 NCEP/NCAR Reanalysis

The NCEP/NCAR reanalysis (NNR) project was designed to provide homogenized (gridded) records to support climate research. The system combines data from a range of sources in an assimilation scheme to produce temporally continuous "observed" atmospheric fields for 1953-2001. The atmospheric model used for the NCEP/NCAR reanalysis project is spectral (spherical harmonic basis functions) with transformation to a Gaussian grid for calculation of nonlinear quantities and physics. The horizontal resolution is spectral triangular 62 (T62, ~200 km). The vertical domain is divided into 28 unequally spaced sigma levels with enhanced resolution near the bottom and the top. The data used here are archived at a resolution of 2.5° × 2.5° (Figure 1). In this study, twice-daily 500 hPa geopotential height fields from the NNR are averaged to produce daily fields from which the PNA index and circulation climatology for the study region are derived. Daily mean sea-level pressure (MSLP) fields are used to compute the NAO index.

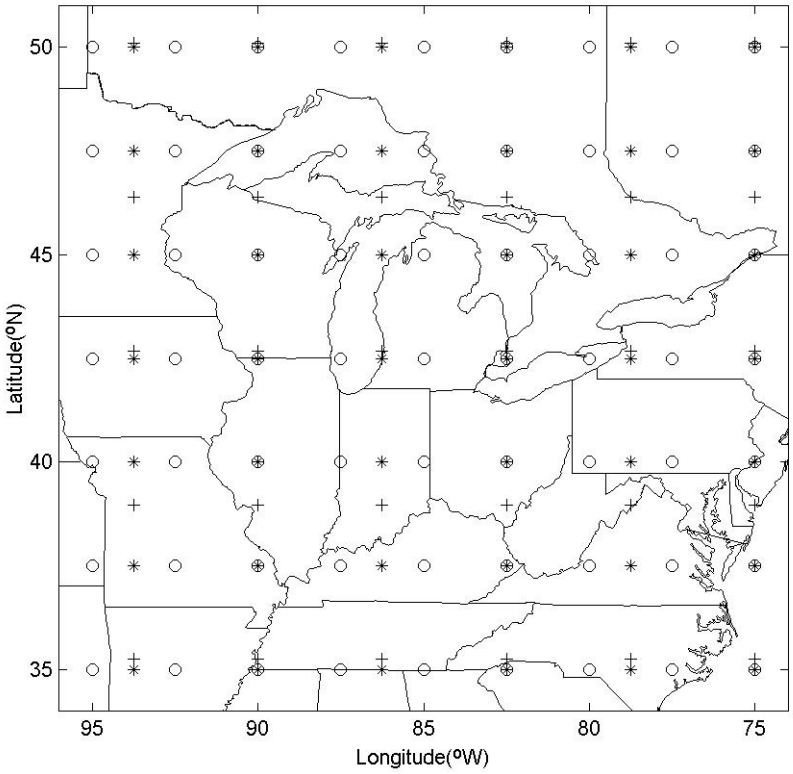

Figure 1. Map of the study region showing NNR (o), HadCM3 (∗) and CGCM2 (+) grid points.

2.2 GCM Simulations

This study uses daily output from transient simulations for the period 1990-2100 conducted using two GCMs; HadCM3 and CGCM2. HadCM3 is an example of a Cartesian GCM (Johns et al. 1997; Stratton 1999; Pope et al. 2000). The atmospheric component of HadCM3 has 19 levels with a horizontal resolution of approximately 2.5° latitude by 3.75° longitude, which produces a global grid of 96 × 73 points. This is equivalent to a surface resolution of about 417 km × 278 km at the equator, reducing to 295 km × 278 km at a latitude of 45° (comparable to a spectral resolution of T42). CGCM2 is the second generation spectral GCM from the Canadian Center for Climate Modeling and Analysis (McFarlane et al. 1992; Flato et al. 2000). The atmospheric

component of the model has ten vertical levels and employs a triangular spectral truncation having 32 longitudinal waves (T32/L10 or approximately 3.75° × 3.75°). Unlike HadCM3, CGCM2 employs flux adjustments as first described in Hansen et al. (1984). The output from HadCM3 used in this analysis was archived at a spatial resolution of 2.5° latitude × 3.75° longitude, while that from CGCM2 was archived at approximately 3.75° × 3.75°. The model grid points are shown in Figure 1. Prior to application of the synoptic map pattern classification, both GCM grids were interpolated to the 2.5° × 2.5° NNR grid (Figure 1) using an inverse distance based interpolation algorithm.

The GCM simulations used here were conducted using the SRES A2 emissions scenario (IPCC, 2000). The SRES scenarios, as developed by the IPCC, encompass a range of parameters that dictate emissions including demographic, technological and economic changes. The SRES A2 scenario group results in global carbon dioxide (CO_2) emissions from industry and energy in 2100 that are almost four times the 1900 value and emissions from land use change by 2100 are close to zero, leading to a global CO_2 emission in 2100 of almost 28 GtC yr^{-1}. This emissions scenario is used in this analysis to provide an upper bound on likely climate change and hence a high signal to noise ratio when comparing current to future teleconnection indices and synoptic climate.

3 Methodology

3.1 Teleconnection indices

3.1.1 The North Atlantic Oscillation (NAO) index

There is no unique way to define the spatial structure of the NAO (Hurrell et al. 2003). It has been defined in terms of pressure differences between point locations, area weighted pressure extremes or in terms of principal components of pressure fields (see Osborn et al. 1999; Hurrell et al. 2003). In this study, the NAO index is computed in terms of the former; as the difference in standardized SLP between grid points near the NAO "centers of action", at Ponta Delgada, Azores (37.7°N, 25.7°W) and Stykkisholmur, Iceland (65.1°N, 22.7°W). Results of compositing analyses (not shown) indicate that these locations represent accurate proxies for the NAO centers of action as manifest in the NNR and GCMs. Because NNR, HadCM3, and CGCM2 output are archived at different resolutions, output from all data sources was interpolated to these station coordinates. The seasonal cycle was removed from each data series prior to the analysis. Since the NAO is most pronounced during the winter (e.g., Rogers 1984), we examine the NAO index only during the winter season (defined here as DJF) and our NAO index is standardized using DJF means and standard deviations. This normalization procedure generates a data series that has a mean of zero and unit variance. Normalized indices from the GCMs and NNR (where each is normalized to the specific model values) are compared here because we wish to focus on the relative behavior of the centers of action and specifically the temporal variability of the pressure gradient rather than the absolute values at the 'station locations'. It is worthy of note that the mean difference in MSLP values between the two NAO station locations derived from NNR, HadCM3, and CGCM2 data for the reference period (1990-2001) are 24.3, 15.5, and 22.3 hPa, respectively, suggesting that HadCM3 may underestimate the mean pressure gradient between the two station locations.

3.1.2 The Pacific North American (PNA) index

The PNA index used in this study is computed using the equation of Wallace and Gutzler (1981):

$$PNA = \frac{1}{4}\left[Z(20^{\circ}N,160^{\circ}W) - Z(45^{\circ}N,165^{\circ}W) + Z(55^{\circ}N,115^{\circ}W) - Z(30^{\circ}N,85^{\circ}W)\right] \quad (1)$$

where Z are standardized (by season within the study period) 500 hPa geopotential height values. Like the NAO, results of compositing analyses (not shown) indicate that these locations are accurate representations of the PNA centers of action as manifest in the NNR and GCMs. As with the NAO, the input values for the PNA calculations are obtained by interpolating HadCM3, and CGCM2 output to the locations given in (1). The NNR grid points coincide with those specified in (1). All data series have the seasonal cycle removed prior to analysis. While the PNA is not a major mode of the Northern Hemisphere circulation during the summer (Barnston and Livezey 1987), Leathers et al. (1991) found statistically significant relationships between the PNA index and US temperatures during the winter, spring, and fall. Hence, we examine the PNA index during winter (DJF), spring (MAM), and fall (SON).

3.1.3 Evaluation of large-scale teleconnection indices

Differences in observed and GCM-simulated teleconnections will be manifest as differences in the probability distributions of the teleconnection indices. We therefore evaluate the teleconnection indices using two statistical methods to identify differences in probability distributions of the NAO and PNA: the Kolmogorov-Smirnov (K-S) and χ^2 goodness-of-fit tests (D'Agostino and Stephens 1986). The K-S test is applied to continuous data and is based on the largest absolute difference between empirical

cumulative distribution functions (ECDFs). The χ^2 test is applied to discrete data and, as employed here, is based on binned counts of the teleconnection indices (with bin sizes of 0.50 and 0.25 for the NAO and PNA, respectively). It identifies differences over the discretized bins and suggests rejection of the null hypothesis when the sum of squared differences between binned counts is sufficiently large. Both tests are conducted under the null hypothesis that the daily teleconnection indices from NNR and the GCMs are drawn from the same distribution.

The period of standardization plays a critical role in the calculation of teleconnection indices. Hence, when comparing NNR and GCM-simulated teleconnection indices for the period 1990-2001, we standardize the data using the 1990-2001 mean and standard deviation from the NNR or GCMs. To place the differences between observed and simulated teleconnection indices in context, we also examine their magnitude relative to recent and possible future changes in the teleconnection indices. In these cases, data are standardized using the 1953-2001 and 1990-2041 mean and standard deviation for the NNR and GCMs, respectively.

3.2 Synoptic-scale map-pattern analysis

Several methodologies can be applied to develop synoptic-scale circulation classifications (Yarnal 1993). In this study, we use correlation based methods because they are readily amenable to (i) development of targeted classifications for the GCM output relative to the NNR and (ii) comparison of the NNR and GCM classifications using bootstrapping techniques to develop confidence intervals for the mean seasonal map type frequencies. The former allows direct comparison of the NNR- and GCM-

derived classifications, while the latter facilitates application of quantitative statistical tools to assess differences in the classifications.

3.2.1 Correlation-based map typing

Correlation-based methods of synoptic-scale weather classification were first used by Lund (1963), who classified map patterns using the Pearson product-moment correlation of gridded fields. Kirchhofer (1973) improved on the correlation-based map typing technique by requiring that sub-map scale patterns also meet the correlation threshold (El-Kadi and Smithson, 1992). The Kirchhofer score of similarity, S, for two grid point maps, containing x and y, is given by:

$$S = \sum_{i=1}^{n} (z_{xi} - z_{yi})^2 \qquad (2)$$

In (2), z_{xi} and z_{yi} are the normalized (spatially) grid-point values of x and y, respectively, and n is the number of grid points. Normalization of the data ensures that the value of S indicates similarity in map patterns and is not contaminated by the magnitudes of the values being classified (Blair, 1998). The relationship between the Kirchhofer score, S, and the correlation coefficient, r (Wilmott 1987; Blair 1998) is given by:

$$S(n-1)^{-1} = 2[1 - r(x,y)] \ , \ 0 \leq S(n\text{-}1)^{-1} \leq 4 \qquad (3)$$

By setting a threshold for the required correlation, (3) can be used to determine the appropriate threshold for S. It is then necessary to calculate S for all combinations of pairs of observations being classified. Sub-grid scale similarity is ensured by also requiring each of the rows and columns of the grids to meet the similarity criteria. The observation with the maximum number of threshold exceedances is referred to as the first

key day. This key day and all 'similar' observations are then removed from the analysis. All days similar to those days are also removed. This process is then repeated until there are no more days left. Once all of the key days have been identified, each observation that has been removed is assigned to the key day for which it has the highest S value.

Different thresholds for grid and sub-grid similarity and minimum group size have previously been applied for different variables (Yarnal 1985; McKendry et al. 1995). In this study, the correlation threshold for the entire gird is set at 0.75, the individual row and column thresholds are set at half of the entire grid threshold (0.375), and the map-pattern groups are required have more than five members. These values are chosen to ensure a manageable number of weather types and a high percentage of classified days.

3.2.2 Evaluation of synoptic-scale map-pattern classification

To evaluate the ability of the two GCMs to represent the circulation regimes of the study area, we apply the Kirchhofer method to the NNR 500 hPa geopotential heights fields for the period of overlap with the GCMs (1990-2001). Viewing the 1990-2001 NNR results as one possible realization of the current climate, other potential realizations are extracted using a bootstrap resampling method (Efron, 1982) to generate a 'climatology' comprising 1000 random samples of this 12 year period, which are presented in the form of confidence intervals for the mean map type frequencies.

'Targeted' Kirchhofer classifications are subsequently applied to HadCM3 and CGCM2 output for 1990-2001. In each targeted classification, the key days identified in the 1990-2001 NNR are used as seeds. Days from the GCM simulations meeting the

similarity threshold with one of the established key days are removed from the analysis along with all days that are similar to them. Once all observations have been compared to all key days a second run through the data allows reassignment of previously classified observations. Use of the targeted classifications facilitates quantitative comparisons between observations (NNR) and simulations (GCMs) to determine whether differences in the synoptic regimes as manifest the NNR and GCM time series result from differences in the map types or differences in the frequencies with which the map types are observed. In addition to the seasonal individual map type frequencies, the analysis of the synoptic scale map types also includes an examination of the persistence of individual map patterns (defined as the average run length) and the progression from one map pattern to another (examined by tabulating the map types which most commonly precede each of the identified map types).

To place the differences between NNR and the GCMs in context, we examine their magnitude relative to recent changes in NNR and potential future changes as manifest in the GCMs. For the former, 1953-1964 NNR map type frequencies are compared to the NNR 1990-2001 bootstrap confidence intervals. For the latter, we construct bootstrap confidence intervals using the 1990-2001 GCM data and then compute the mean map type frequencies over the period 2030-2041.

3.3 Relationships between teleconnections and Midwest US weather types

Using the NAO and PNA indices described above, the frequency of each weather type defined using the Kirchhofer method is determined for each phase of the NAO and PNA. Each day is assigned as either a positive (more than one standard deviation greater

than the seasonal mean of the teleconnection), negative (more than one standard deviation less than the seasonal mean of the teleconnection), or neutral (within one standard deviation of the seasonal mean of the teleconnection). The strength of the link between the teleconnections and map types is then determined by comparing the frequency of the map type occurrence during positive and negative phases of the teleconnections.

To quantify variations in the links between the teleconnection indices and the weather types a two-sample difference of proportions test (Ott, 1993; Sheridan 2003) was applied to the proportions of positive and negative phase days for each teleconnection using a z-score:

$$z = \frac{\pi_1 - \pi_2}{\hat{\sigma}} \tag{4}$$

where π_1 and π_2 represent the proportions of positive and negative phase days for a particular season and particular teleconnection, and $\hat{\sigma}$ is:

$$\hat{\sigma} = \sqrt{\hat{\pi}(1-\hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \tag{5}$$

where n_1 and n_2 are the total number of days in the season of the positive and negative NAO or PNA phase, respectively, and $\hat{\pi}$ is

$$\hat{\pi} = \frac{n_1\pi_1 + n_2\pi_2}{n_1 + n_2} \tag{6}$$

Proportions that result in a large (absolute) z-score (i.e., not expected more than 5% of the time if the proportions are the same) suggest that a particular map type is more likely to occur in either the positive or negative phase of the teleconnection. The test is not

performed for weather types for which $n\hat{\pi}$ or $n(1-\hat{\pi})$ is less than 5 for either proportion because the normal approximation to the binomial does not hold for small samples.

4 Results

4.1 Teleconnection indices

4.1.1 North Atlantic Oscillation (NAO)

The K-S test and χ^2 statistics for the probability distributions of the winter NAO indices as manifest in NNR and HadCM3 and CGCM2 during the reference period (1990-2001) do not warrant rejection of the null hypotheses that the observed and simulated NAO indices are drawn from the same underlying distribution at the $\alpha = 0.05$ level (Table 1). Comparison of the NNR data and GCM output as interpolated to the NAO index station locations indicates the behavior of these 'centers of action' is well captured by the models both in terms of the absolute magnitude and common behavior of MSLP (mean MSLP values for NNR, HadCM3, and GCM2 over the reference period differ by less than 6 hPa at both stations while standard deviations of MSLP values differ by less than 5 hPa). Therefore, although HadCM3 underestimates the mean MSLP gradient in the North Atlantic region (Section 3.1.1), the variability in the normalized gradient is well captured by the model. In addition, each model correctly simulates the strong lag-1 autocorrelation (r_1=0.86) present in the daily NNR data (r_1=0.85 and 0.86 for HadCM3 and CGCM2, respectively). Hence, it is inferred that the GCMs are correctly simulating the probability distribution and persistence of the daily NAO index.

Table 1. Results of K-S and χ^2 tests performed on NNR, HadCM3 and CGCM2 NAO and PNA indices. For each season the table provides the K-S and χ^2 statistics and p-values (i.e., the probability of observing a test statistic as extreme or more extreme than that observed if the null hypothesis is true) associated with the test. Table entries are in bold if statistically significant at α=0.05.

NAO DJF	**K-S**	**χ^2**
NNR vs. HadCM3 1990-2001	0.052 (0.105)	22.48 (0.522)
NNR vs. CGCM2 1990-2001	0.020 (0.982)	18.64 (0.960)
NNR 1953-1964 vs. NNR 1990-2001	**0.141 (0.000)**	**80.73 (0.000)**
HadCM3 1990-2001 vs. HadCM3 2030-2041	0.036 (0.475)	16.39 (0.731)
CGCM2 1990-2001 vs. CGCM2 2030-2041	**0.123 (0.000)**	**50.80 (0.000)**
PNA DJF	**K-S**	**χ^2**
NNR vs. HadCM3 1990-2001	**0.067 (0.014)**	**44.52 (0.000)**
NNR vs. CGCM2 1990-2001	**0.065 (0.018)**	**49.80 (0.001)**
NNR 1953-1964 vs. NNR 1990-2001	**0.063 (0.025)**	31.32 (0.137)
HadCM3 1990-2001 vs. HadCM3 2030-2041	**0.108 (0.000)**	**57.82 (0.000)**
CGCM2 1990-2001 vs. CGCM2 2030-2041	0.026 (0.861)	12.27 (0.136)
PNA MAM	**K-S**	**χ^2**
NNR vs. HadCM3 1990-2001	0.025 (0.884)	**8.73 (0.018)**
NNR vs. CGCM2 1990-2001	0.029 (0.751)	12.37 (0.142)
NNR 1953-1964 vs. NNR 1990-2001	**0.224 (0.000)**	**121.87 (0.000)**
HadCM3 1990-2001 vs. HadCM3 2030-2041	**0.081 (0.002)**	32.71 (0.100)
CGCM2 1990-2001 vs. CGCM2 2030-2041	0.039 (0.359)	16.37 (0.503)
PNA SON	**K-S**	**χ^2**
NNR vs. HadCM3 1990-2001	0.049 (0.144)	31.33 (0.137)
NNR vs. CGCM2 1990-2001	**0.063 (0.026)**	33.48 (0.082)
NNR 1953-1964 vs. NNR 1990-2001	0.042 (0.300)	27.10 (0.335)
HadCM3 1990-2001 vs. HadCM3 2030-2041	0.034 (0.544)	16.79 (0.552)
CGCM2 1990-2001 vs. CGCM2 2030-2041	0.050 (0.122)	27.61 (0.303)

To examine temporal evolution of the NAO, the wintertime index was calculated for the observed record (1953-2001) and for the GCM simulation period (1990-2041). The 1953-2001 NNR NAO index exhibits a trend toward more positive index values, in accord with previous studies (e.g., Hurrell 1995) and consistent with the established underlying trend in the north-south pressure difference in the Atlantic in the latter portion of the twentieth century (Pryor and Barthelmie, 2003). Indeed, application of the K-S and χ^2 tests to the first and last 12 years of the NNR data (i.e., 1953-1964 and 1990-2001) results in rejection of the null hypothesis for both tests (Table 1), suggesting that differences between the observations and GCMs are of lesser magnitude than changes in

the NAO index observed in recent decades. Application of the same statistical tests to HadCM3 for comparison of the reference period (1990-2001) with 2030-2041 does not result in a rejection of the null hypothesis for either test (Table 1, Figure 2a). However, the intervening 12-year periods (2000-2011, 2010-2021 and 2020-2031) all exhibit significantly lower NAO than the reference period. In contrast, the CGCM2 derived NAO for 1990-2001 v. 2030-2041 differ according to both the K-S and χ^2 statistics. As shown in Figure 2b, CGCM2 simulations suggest a shift to slightly more negative NAO values in 2030-2041 relative to the reference period. Differences between observed and simulated probability distributions of the NAO index over the reference period (1990-2001) are thus determined to be smaller than both changes in the indices during recent decades and prognostic changes quantified using transient GCM output. This finding that the GCMs do not suggest continued evolution towards higher north-south gradients across the Atlantic (higher positive NAO) is in accord with analyses in Osborn et al. (1999) which indicate that the positive trend in observed winter NAO from 1963-1992 is 'highly unusual' and that the HadCM2 winter NAO exhibited a decline from the 1990 level over the following decades. However, in an ensemble GCM experiment with quadrupled CO_2 concentrations, Paeth et al. (1999) found that the observed positive NAO trend continues for the next several decades and is accompanied by a decrease in variability. This result suggests that GCM-simulated NAO indices may be quite sensitive to the emissions scenario employed.

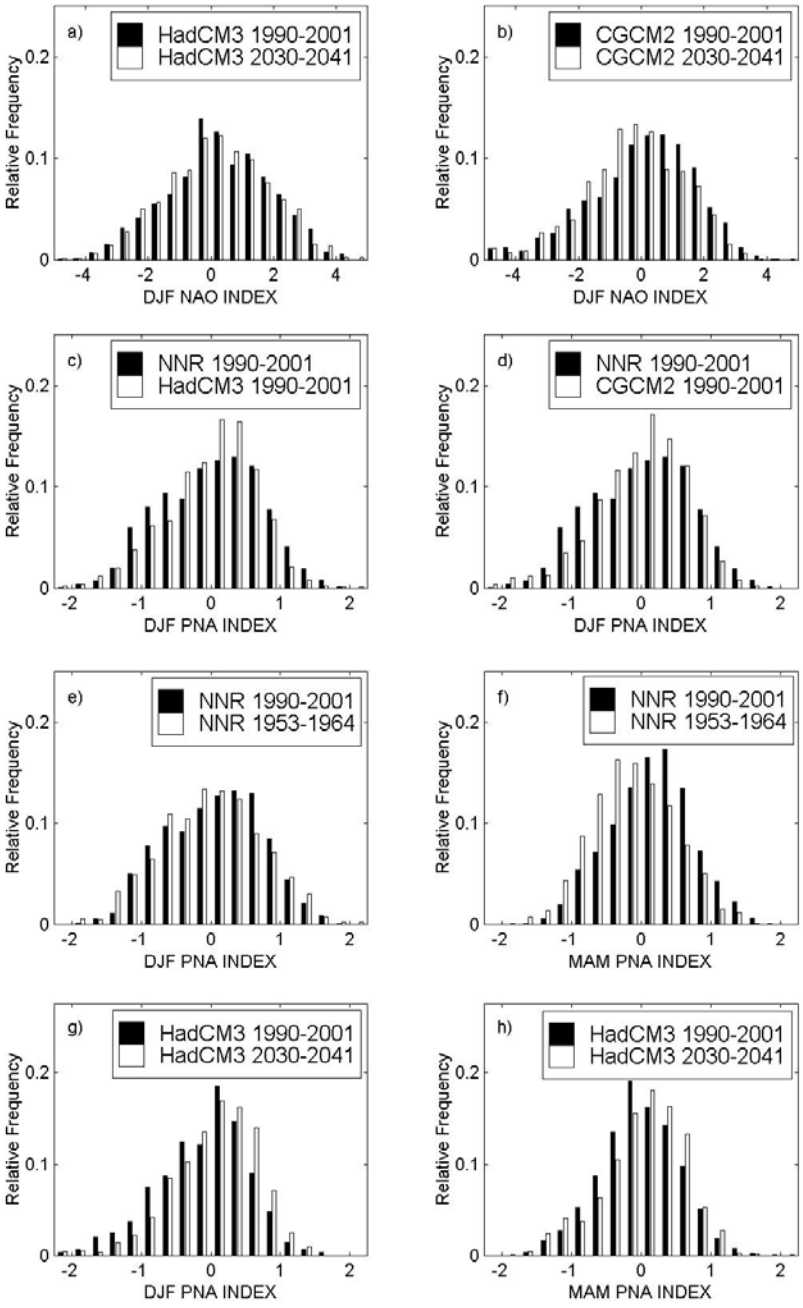

Figure 2. Histograms comparing the probability distributions of the observed (NNR) and GCM-simulated teleconnection indices: (a) winter HadCM3 NAO index (1990-2001 vs. 2030-2041), (b) winter CGCM2 NAO index (1990-2001 vs. 2030-2041), c) 1990-2001 winter PNA index (NNR vs. HadCM3), d) 1990-2001 winter PNA index (NNR vs. CGCM2), e) winter NNR PNA index (1953-1964 vs. 1990-2001), f) spring NNR PNA index (1953-1964 vs. 1990-2001), g) winter HadCM3 PNA index (1990-2001 vs. 2030-2041), and h) CGCM2 fall PNA index (1990-2001 vs. 2030-2041).

4.1.2 Pacific/North American (PNA) index

When the K-S and χ^2 tests are applied to the observed (NNR) and GCM-derived PNA indices for the reference period, the null hypothesis that the indices are drawn from the same underlying distribution is rejected (at α=0.05) for both tests and both GCMs for the winter season (Table 1). During the spring and fall, the distributions are significantly different only for one of two models and one of two tests (χ^2 for HadCM3 in spring, K-S for CGCM2 during fall, Table 1). Examination of the winter PNA probability distributions from the NNR and GCMs (Figure 2c and 2d) suggests that the indices differ in terms of both skewness and kurtosis. Both HadCM3 and CGCM2 simulate near-zero and small positive PNA values more often than NNR and negative PNA values less often than NNR. The probability distributions show better agreement near the tails, which are the areas of the distribution expected to be most influential to weather patterns in the study area (e.g., cooler temperatures in the study region resulting from increased meridionality during strong positive PNA events; Leathers et al. 1991). Each model correctly simulates the strong lag-1 autocorrelation (r_1=0.90, 0.85, and 0.85 for DJF, MAM, and SON, respectively) present in the NNR data (differences between observed and GCM-simulated lag-1 autocorrelation are less than 0.1 for both models in each season).

Investigation of the individual PNA centers of action suggests that the small positive PNA bias in HadCM3 during winter results from stronger low pressure in both the North Pacific and eastern USA (i.e., the 2nd and 4th terms in Equation 1) relative to observations (NNR). The differences between NNR and CGCM2 result from stronger high pressure over North America in conjunction with stronger low pressure over both

the North Pacific and eastern USA (i.e., the 2nd, 3rd and 4th terms in Equation 1) in the model relative to observations (NNR).

As with the NAO, we provide a context for the differences between observed and GCM-simulated PNA indices by examining the magnitude of the differences in indices from the observed record (1953-2001) and the GCMs (1990-2041). K-S and χ^2 tests applied to the first and last 12 years of the NNR PNA index (1953-1964 and 1990-2001) show a shift toward more positive phase days during spring (Table 1, Figure 2f). The K-S test also identifies a statistically significant difference in the winter PNA index between the two periods (Table 1, Figure 2e), although the differences were not manifest as rejection of the null hypothesis for the χ^2 test. Application of the statistical tests to the HadCM3 PNA indices for 1990-2001 and 2030-2041 suggests the largest changes in the probability distribution of the PNA will occur during the winter (Table 1, Figure 2g), with a shift toward more positive phase days. The difference in the spring PNA index between the two periods (Table 1, Figure 2h) is also significant according to the K-S test, but not according to the χ^2 test. These results, taken in concert with analyses by Leathers and Palecki (1992) who suggest that the 1957-1987 PNA index exhibited a bias toward positive values, may be indicative of a sustained trend towards more meridional flow over the continental US. In contrast to the results from HadCM3, application of the statistical tests to CGCM2 using data for the same period does not result in a rejection of the null hypothesis for any of the seasons tested, suggesting that the distribution of the PNA will be similar to present day conditions.

4.1.3 Summary

The results presented in sections 4.1.1 and 4.1.2 indicate that HadCM3 and CGCM2 manifestations of the NAO are in agreement with the NAO as derived from NNR over the 1990-2001 reference period. HadCM3 also properly simulate the probability distribution of the daily PNA index during spring (the season during which the largest changes are found to have occurred in the historical NNR record) and CGCM2 correctly simulates the index during fall. While the range of PNA values during the reference period is well captured by the GCMs, during the winter season small positive values of the PNA index are simulated too often by the GCMs, small negative values are simulated too seldom and the null hypothesis that observed and GCM-simulated indices are drawn from the same underlying distribution is rejected for both GCMs using two different statistical tests.

In brief, these analyses indicate differences between GCM and NNR realizations of the NAO for the reference period (1990-2001) are smaller than both changes in the observed historical record and prognostic changes for approaching decades. For the PNA index, differences in observed and GCM-simulated values over the reference period (1990-2001) are of comparable magnitude to historical changes in the PNA index.

HadCM3 indicates the probability distribution of the wintertime NAO in 2030-2041 will be statistically similar to that during the reference period (1990-2001), while CGCM2 indicates lower NAO values during the later period. The discrepancy between the model projections for the NAO may be due to differences in the model treatment of atmosphere-surface interactions. While these models contain similarly discretized oceanic models (HadCM3 has 20 vertical layers and a horizontal resolution of 1.25° ×

1.25°; CGCM2 has 29 vertical layers and horizontal resolution of 1.875° × 1.875°), CGCM2 uses monthly heat and water flux adjustments, which may moderate the atmospheric response to changing oceanic conditions. These characteristics of CGCM2 may be manifest in the NAO realizations because while the NAO is particularly responsive to atmospheric forcing on the annual timescale, it is most sensitive to changes in oceanic circulation and sea surface temperatures as well as snow cover within the North Atlantic sector and beyond on longer timescales (Kushnir 1994; Hurrell and van Loon 1997; Tourre et al. 1999; Bojariu and Gimeno 2003).

HadCM3 exhibits statistically higher PNA during the winter and spring in 2030-2041, while differences in the PNA in CGCM2 for the same period are not significant. These differences may also be explicable in terms of the treatment of atmosphere-surface exchange in the two models since up to 30% of the variability in winter 500-hPa geopotential height anomalies in the Pacific – North American sector is due to sea-surface temperature (SST) forcing (Hoerling and Kumar 2002).

4.2 Synoptic-scale map-pattern analysis

4.2.1 Observed map-patterns: 1990-2001

When applied to the 1990-2001 daily NNR 500-hPa geopotential height fields, the Kirchhofer method (Section 3.2.1) identifies 15 map patterns accounting for 91.37% of the observations. The "keydays" for these map-patterns are shown in Figure 3. Table 2 shows the percentage frequency for each map type along with the mean correlation coefficient (a measure of within-class variability computed as the mean correlation between each key day and members of the class associated with the key day), the

persistence and the map type progression (described by the most commonly preceding map types in descending order). The correlations shown in Table 2 are much higher than the threshold of 0.75, suggesting that this classification is robust (i.e. the 'types' are internally coherent).

The three most commonly observed Kirchhofer map patterns represent the "average" meridional (map type 1) and zonal (map types 8 and 11) conditions over the study area. The first pattern accounts for 17.6% of the observations and is characterized by a strong southwest-northeast height gradient and a trough located northeast of the study area. Weak north-south height gradients and slightly above average heights characterize the second and third most prevalent patterns identified by the analysis (map types 11 and 8), which account for approximately 15.0% and 14.4% of the observations, respectively. Each of the subsequent patterns accounts for less than 10% of the observations, although each represents a unique and meaningful synoptic-scale flow regime.

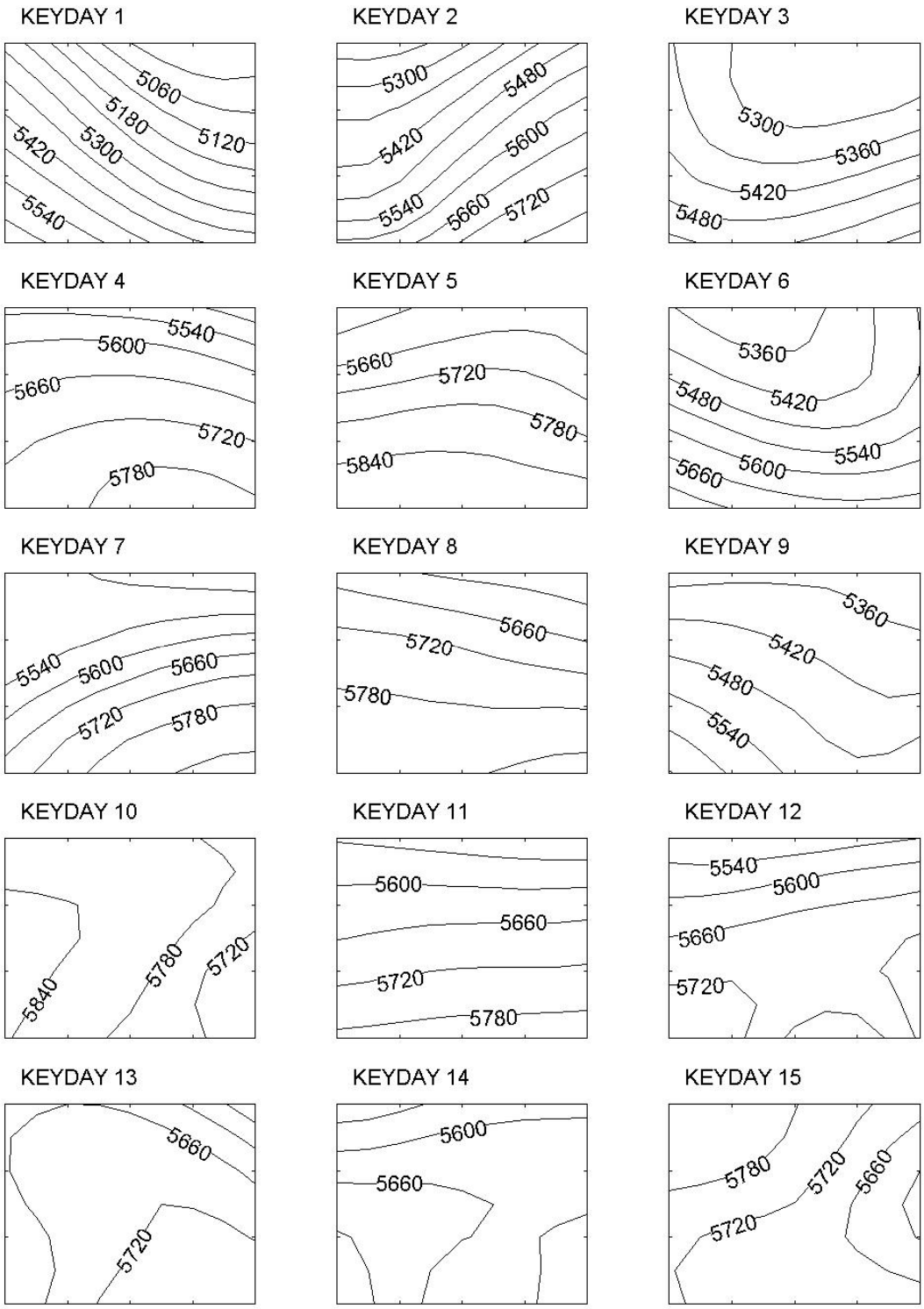

Figure 3. Key day maps for each of the 15 Kirchhofer map types associated with the 1990-2001 NNR 500-hPa geopotential height fields (in gpm). The domain is as shown in Figure 1.

Table 2. Percentage frequency, mean correlation coefficient, mean persistence, and most commonly preceding map types for each of the 15 Kirchhofer map types identified from the 1990-2001 NNR, 1990-2001 HadCM3, and 1990-2001 CGCM2 500-hPa geopotential height fields.

Map type	Percentage frequency (%)			Mean correlation coefficient			Mean persistence			Most commonly preceding types		
	NNR	HadCM3	CGCM2	NNR	HadCM3	CGCM2	NNR	HadCM3	CGCM2	NNR	HadCM3	CGCM2
1	17.59	26.57	16.58	0.927	0.918	0.941	1.78	2.25	1.90	1, 8, 3	1, 3, 8	1, 8, 11
2	7.60	8.54	4.25	0.920	0.915	0.933	1.46	1.65	1.49	2, 7, 5	2, 7, 4	2, 7, 5
3	6.68	5.90	7.92	0.934	0.929	0.948	1.28	1.18	1.40	3, 7, 11	7, 2, 11	3, 7, 11
4	5.80	9.72	4.54	0.948	0.928	0.950	1.24	1.41	1.24	8, 1, 4	4, 1, 8	8, 4, 1
5	7.73	3.50	7.99	0.948	0.925	0.954	1.20	1.09	1.25	1, 5, 11	1, 4, 9	5, 1, 8
6	3.74	2.13	2.60	0.944	0.931	0.951	1.15	1.16	1.18	2, 5, 6	2, 5, 8	11, 6, 2
7	6.50	6.41	6.78	0.952	0.937	0.951	1.27	1.22	1.30	11, 7, 2	11, 4, 7	11, 7, 4
8	14.37	9.44	15.27	0.951	0.931	0.952	1.50	1.42	1.49	8, 1, 11	8, 1, 11	8, 11, 1
9	4.15	4.21	4.41	0.929	0.915	0.939	1.17	1.24	1.21	1, 3, 9	1, 9, 3	1, 9, 3
10	0.39	0.44	0.02	0.877	0.795	0.803	1.70	1.27	1.00	10, 15, 1	1, 10, 15	-
11	14.99	9.07	18.81	0.972	0.955	0.967	1.53	1.26	1.54	11, 8, 5	11, 1, 2	11, 8, 5
12	0.66	0.86	0.57	0.903	0.864	0.899	1.26	1.19	1.25	12, 2, 5	4, 5, 12	5, 12, 2
13	0.59	0.74	0.09	0.881	0.868	0.901	1.24	1.28	1.00	4, 13, 1	4, 13, 1	1, 8
14	0.41	0.21	0.11	0.889	0.872	0.872	1.20	1.00	1.00	14, 1, 7	1	1, 9
15	0.14	0.07	0.02	0.848	0.819	0.852	1.00	1.00	1.00	1, 3	1	1

4.2.2 Comparison of map type frequencies in NNR and transient GCM simulations

The targeted Kirchhofer analysis of the HadCM3 and CGCM2 simulations for 1990-2001 led to classification of 87.80% and 89.98% of simulated days, respectively. As shown by Table 2, both GCM classifications reproduce the entire range of map types found in the NNR data. The HadCM3 and CGCM2 classifications also exhibit correlation coefficients for the members of each type which are of similar magnitude to those found in the NNR analysis, suggesting that within-type variability is of similar magnitude to that in the observed record.

The largest discrepancy between the NNR and HadCM3 map type frequencies is an overestimation of the first map type and subsequent underestimation of the second and third most prevalent map types (map types 8 and 11) both of which represent a dominant high pressure to the south and near zonal flow over the study region. The predominance of map type 1 in the HadCM3 simulation combined with an underestimation of the inter-annual variability in the relative frequency of this type may indicate HadCM3 under-represents the variability of the synoptic climate in this region. The map type frequencies from the targeted analysis of CGCM2 show better agreement with the NNR classification. The largest differences between NNR and CGCM2 classifications are an underestimation of map type 2 (a deep trough north/northwest of the study area) and overestimation of map type 11 by CGCM2.

Accompanying the overestimation of the most common map type by HadCM3 is an overestimation of the persistence associated with it. For the other map types, both GCMs simulate the persistence of the map types well. Although there are some differences in the progression of the map types (as manifest in the most commonly

preceding map types, Table 2), the first order progression of the most common map types is also well simulated by both GCMs.

Figure 4 shows that mean annual frequency of the map types by season relative to the bootstrapped confidence intervals derived from the NNR for the reference period. As shown, the first and most common NNR map type is simulated too often in all seasons by HadCM3, although the overestimation is small during fall. Most of the map types for which the largest differences between observed (NNR) and HadCM3 simulated map type frequencies exist are characterized by weak height gradients and above average heights in the study region consistent with anticyclonic conditions (map types 5, 8, and 11 during winter, map types 4 and 5 during spring, map types 4, 5, 8, and 11 during summer and fall).

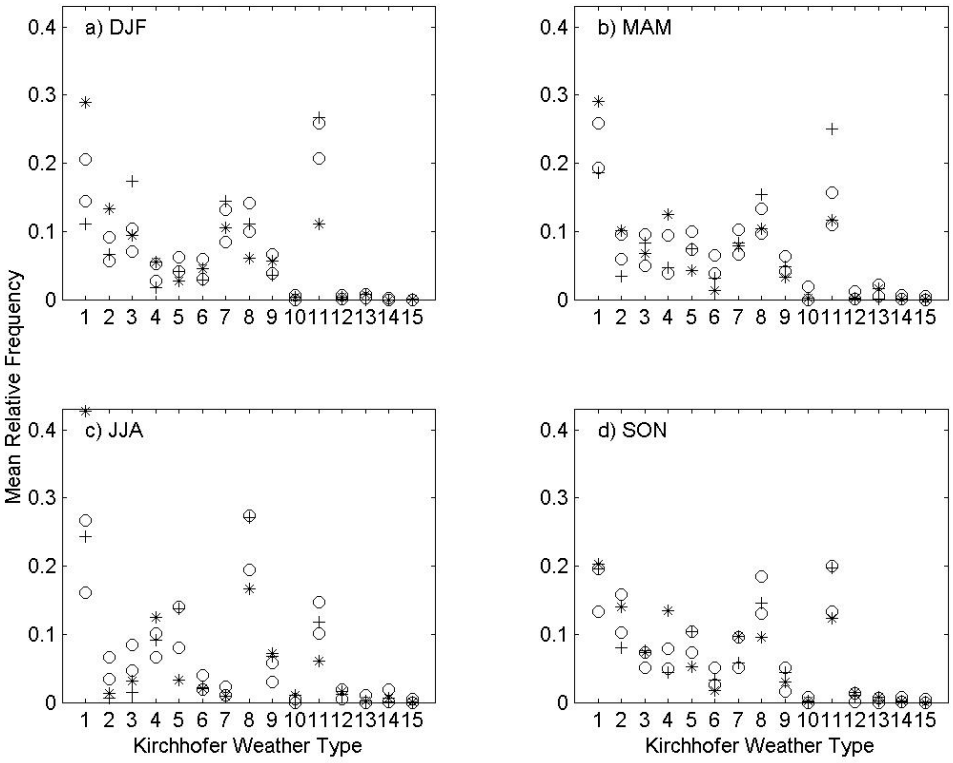

Figure 4. Mean relative frequency of 1990-2001 HadCM3 (*) and CGCM2 (+) Kirchhofer map types and bootstrap confidence intervals (o) computed with 1990-2001 NNR Kirchhofer map type frequencies: a) winter, b) spring, c) summer, and d) fall.

Unlike HadCM3, CGCM2 produces the most common map type too seldom during winter and spring, but relative frequencies are within the NNR confidence limits during the other seasons (Figure 4). For summer and fall, the mean frequencies of nearly all of the map types are within the NNR-derived confidence intervals, although frequencies for two of the cyclonic types (2 and 3 in summer, 2 in fall) are lower than those expected from NNR. The largest difference between NNR and CGCM2 during winter is associated with days assigned to map type 3 rather than map type 1. Both types are associated with low geopotential height anomalies to the northeast of the study region, although map type 1 exhibits much stronger pressure gradients (Figure 3). During spring, CGCM2 overestimates the mean frequency of the two weak anticyclonic map types (8 and 11) and underestimates the mean frequency of two of the cyclonic map types (2 and 6).

The differences between the map type frequencies from HadCM3 and CGCM2 during the spring are partially explicable by consideration of the seasonal cycles of geopotential height as simulated by the GCMs. As shown in Figure 4, CGCM2 overestimates the mean frequency of map type 11 during spring. This map type is more common in the NNR classification during the winter and hence the overestimation by CGCM2 may reflect the slower transition from low geopotential heights during the winter to higher geopotential heights during the summer as indicated by Figure 5. Equally, overestimation of map type 1 by HadCM3, which represents a long-wave trough over the eastern seaboard, may be due to the elevated 500 hPa heights in the western domain during spring (Figure 5).

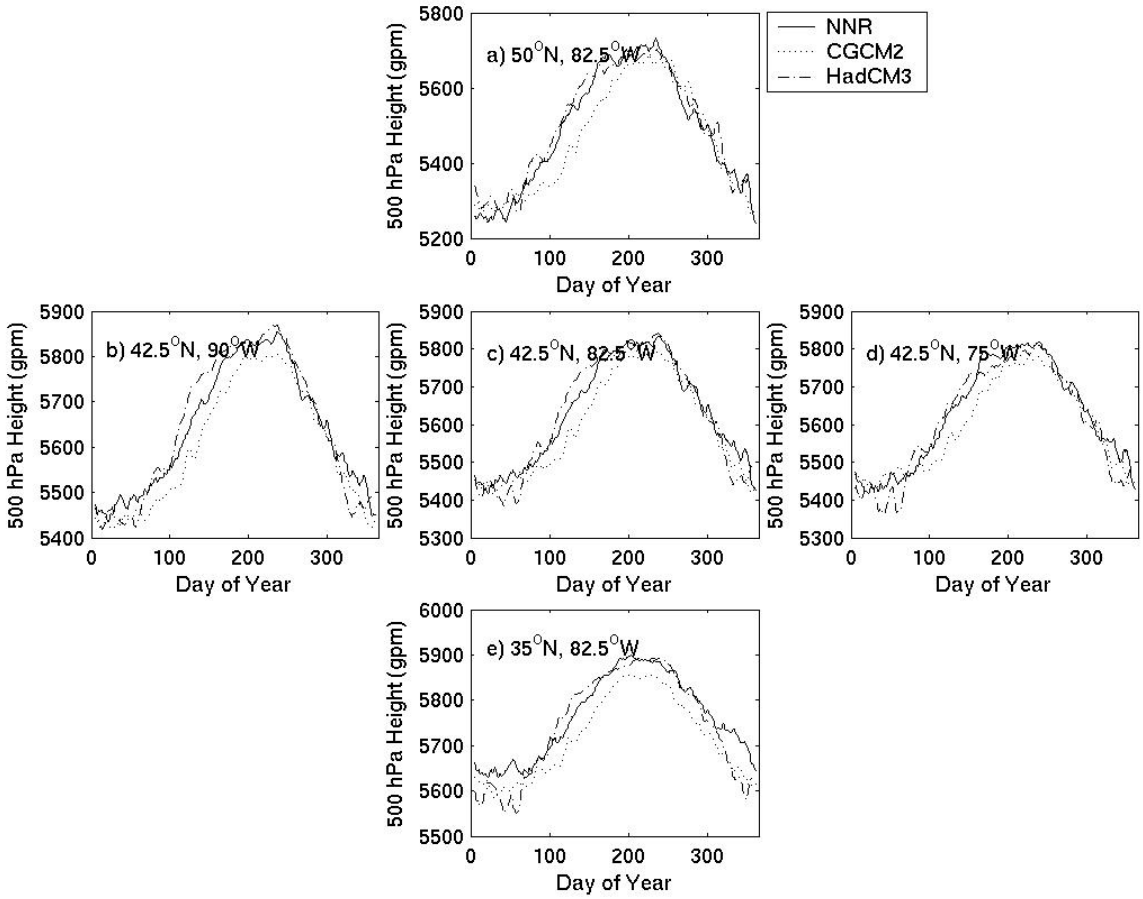

Figure 5. Mean annual cycle (1990-2001) of 500 hPa geopotential for a sample of grid points within the study area: a) 50°N, 82.5°W, b) 42.5°N, 90°W, c) 42.5°N, 82.5°W, d) 42.5°N, 75°W, and e) 35°N, 82.5°W°. An 8-point moving average is applied to each series to remove high frequency noise. Note that different scales are used in different frames.

Relative to differences in the observed (NNR) and simulated (HadCM3, CGCM2) mean map type frequencies, the mean map type frequencies exhibit only small changes over recent decades. Nearly all of the map type frequencies from the first 12 years of the NNR record (1953-1964) lie within the confidence bounds specified by the 1990-2001 NNR data. Exceptions are a slight decrease in the prevalence of map type 1 and a slight increase in the prevalence of map type 11 during winter during the later period.

Examination of the HadCM3 and CGCM2 map type frequencies for 2030-2041 relative to 1990-2001 suggests little evolution of map type frequencies relative to current observed interannual variability. During winter, only the mean frequency of map type 9 during 2030-2041 lies outside of the confidence interval specified by the bootstrapped 1990-2001 HadCM3 data. During summer and fall, there is some evidence of a shift to fewer cyclonic days (map type 3 in both seasons) and a small increase in anticyclonic days (map types 11 and 7 in summer and fall, respectively). Results for CGCM2 suggest fewer anticyclonic days (map types 4 and 8) in summer and more anticyclonic days during fall. These increases are consistent with small changes in frequencies of the cyclonic classes during these seasons (map types 9 and 3 during summer and fall, respectively).

4.2.3 Linking synoptic scale circulation and the teleconnection indices

Analysis of the relationships between the 1990-2001 NNR map types and the NAO index shows that none of the NNR map types are statistically associated with the NAO phase (Table 3). Similarly, the GCMs do not identify significant differences in the frequency of map types during different NAO phases.

Table 3. Results of the two-sample difference of proportions tests performed on NNR, HadCM3, and CGCM2 indicating preference to a particular phase of the NAO and PNA. For each season the table provides the p-value (i.e., the probability of observing differences as extreme or more extreme than that observed if the null hypothesis is true) associated with the test. Table entries are bold if statistically significant at α=0.05 for map types more likely under the positive teleconnection phase. Table entries are underlined if statistically significant at α=0.05 for map types more likely under the negative teleconnection phase A blank entry means that the test was not performed due to small sample size.

Map type	NAO DJF			PNA DJF			PNA MAM			PNA SON		
	NNR	HadCM3	CGCM2	NNR	HadCM3	CGCM2	NNR	HadCM3	CGCM2	NNR	HadCM3	CGCM2
1	0.340	0.113	0.130	**0.004**	0.078	**0.017**	0.239	**0.021**	**0.039**	**0.017**	**0.037**	**0.022**
2	0.301	0.073	0.241	0.211	0.081	0.286	0.242	0.099		0.105	0.107	0.058
3	0.198	0.500	0.085	**0.049**	0.626	**0.033**	0.259	0.267	0.134	0.110	0.139	0.062
4	0.400	0.092		0.198	0.288	0.500	0.154	0.184	0.169	0.217	0.151	0.093
5	0.391		0.174	0.170	0.150	0.049	0.062	0.180	0.041	0.173	0.086	0.107
6				0.454	0.420	0.447	0.269	0.613	0.263	0.274	0.299	0.589
7	0.589	0.213	0.177	0.037	0.103	0.160	0.285	0.138	0.311	0.217	0.234	0.245
8	0.614	0.613	0.690	0.532	0.288	0.604	0.558	0.463	0.451	0.265	0.305	0.222
9	0.616	0.302	0.453	0.303	0.612	0.305	0.574		0.295		0.500	0.500
10												
11	0.650	0.142	0.569	0.001	0.037	0.028	0.130	0.089	0.102	0.133	0.315	0.090
12												
13												
14												
15												

According to the NNR wintertime data, the frequencies of map types 1, 3, 7, and 11 show strong preference for either the positive (map types 1 and 3) or negative (map types 7 and 11) phase of the PNA. NNR spring map types do not show a preference for either PNA phase and during fall only map type 1 is significantly linked to the PNA index (more prevalent during positive phase). Examination of the links between the PNA and the map types as simulated by the GCMs provides insight into the differences between the classifications described in Section 4.2.2. Only one HadCM3 map type shows preference for a particular phase of the PNA within each season (map type 11 in winter (negative phase) and map type 1 in spring and fall (positive phase)). Hence the lack of agreement regarding the mean frequency of map type 1 during winter in HadCM3 relative to NNR may be due to the lack of a link between the PNA and map type 1 in HadCM3. CGCM2 mimics the NNR in terms of the connection between map types 1, 3, and 11, during the winter, although the association between map type 5 and negative PNA values (during the winter and spring) is not evident in the NNR data. Interestingly, the best agreement between NNR and GCM-based map type classifications occurs during fall, which is also the season in which only the first map type is linked to the PNA – a link that is reproduced by both GCMs.

4.2.4 Summary

Circulation regimes at 500 hPa as simulated by HadCM3 and CGCM2 for 1900-2001 show a high degree of correspondence with those manifest in the NNR in terms of frequencies, persistence and progression. However, HadCM3 overestimates the frequency of the most prevalent map type during all seasons. Additional discrepancies

between mean frequencies of the NNR and HadCM3 classifications result from different classification of days with weak anticyclonic characteristics. CGCM2 generally simulates the mean frequencies of the NNR map types better than HadCM3. Some, but not all, of the shortcomings of the GCM-based classifications can be explained by differences in the observed and simulated relationships between the map types and two teleconnection indices. Differences between the mean frequencies of NNR and GCM map types are larger than changes in the map type frequencies as manifest in the 1953-2001 NNR or prognostic changes in the mean map type frequencies as manifest in the 2030-2041 GCM simulations.

5 Concluding Remarks

Reliable regional climate projections for the mid-latitudes are critically dependent on the accuracy of the depiction of teleconnection indices and synoptic scale phenomena within GCMs. We have presented an evaluation of the synoptic-scale climate of the Midwest region of the United States as manifest in the transient HadCM3 and CGCM2 simulations for the A2 emission scenario relative to the NNR data set.

Probability distributions of daily NAO and PNA teleconnection indices from both GCMs and NNR during the reference period (1990-2001) show a high degree of similarity during the seasons in which the teleconnections are major modes of atmospheric circulation (NAO in winter, PNA in winter, spring, and fall). However, during the winter, both GCM-based PNA indices have probability distributions that are significantly different from the NNR PNA index, largely due to differences in the sign of near zero PNA values. These differences in the PNA realizations for the reference period

are larger or comparable to temporal changes in this index over the recent historical record or prognoses for the next 40 years as derived from HadCM3 or CGCM2. The NAO shows greater correspondence between the GCMs and NNR for the reference period and the prognoses indicate a reversal of the trend towards more positive NAO during the later part of the twentieth century over the coming decades.

In general, the analyses presented herein suggest CGCM2 simulates the mean frequencies of 500 hPa flow regimes over the Midwestern USA better than HadCM3. Both GCMs reproduce the range of synoptic-scale map-patterns over the study region as manifest in 500 hPa height fields and modeled map types are similar to those derived for the NNR data in terms of mean frequency, correlation coefficients, persistence, and progression. However, the most common map type occurs more often in HadCM3 than in the NNR, possibly because HadCM3 does not reproduce the observed dependence of map types 1 and 3 on the PNA phase during winter. Additional discrepancies between mean frequencies of the NNR and HadCM3 classifications result from different classification of days with high pressure characteristics.

Despite the relatively good correspondence between the NNR and GCMs with respect to the synoptic scale map types within the study region, differences between NNR and both HadCM3 and CGCM2 in the reference period (1990-2001) are demonstrated to be of comparable or greater magnitude than historical changes (differences between 1953-1964 and 1990-2001 NNR types) and prognostic changes HadCM3 and CGCM2 simulations (differences between 2030-2041 and 1990-2001). Hence, the inference that must be drawn from this analysis is that GCM-derived prognoses for the near-term synoptic climate of this region remain uncertain. This finding does not indicate a lack of

evolution in the map types, but the results do suggest that the evolution does not exceed the bounds of the interannual variability observed over the reference period. It is also important to note there are two major caveats to this finding. First, we examined 12-year time windows to ensure direct comparability with the overlap (reference) period for NNR, HadCM3, and CGCM2. Even with the implementation of the bootstrap resampling methodology, these time periods do not represent a full climatology. Second, it should be emphasized that the findings documented herein were specifically based on simulations conducted using a single emissions scenario (SRES A2). Just as the results could differ for other periods, they may also differ if other GCMs or emissions scenarios were used. Future work will address this issue by evaluating and comparing the performance of additional GCMs under various emissions scenarios.

Acknowledgements

This research was supported in part by the Office of Science (BER), U.S. 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. HadCM3 data has been supplied by the Climate Impacts LINK Project (DERFA Contract EPG 1/1/124) on behalf of the Hadley Centre and U.K. Meteorological Office. NCEP/NCAR mean sea-level pressure data were provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their website at http://www.cdc.noaa.gov.

References Cited

Barnston AG, Livezey RE (1987) Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. Mon Weather Rev 115: 1083-1126.

Blackmon ML, Lee Y-H, Wallace JM, Hsu H-H (1984) Time variation of 500-mb height fluctuations with long, intermediate, and short time scales as deduced from lag-correlation statistics. J Atmos Sci 41: 961-979.

Blair D (1998) The Kirchhofer technique of synoptic typing revisited. Int J Climatol 18: 1625-1635.

Bojariu R, Gimeno L (2003) The role of snow cover fluctuations in multiannual NAO persistence. Geophys Res Lett 30: 1156-1159.

Coleman JS, Rogers JC (2003) Ohio River Valley moisture conditions associated with the Pacific-North American teleconnection pattern. J Climate 16: 969-981.

Covey, C., AchutaRao, K. M., Cubasch, U., Jones, P., Lambert, S. J., Mann, M. E., Phillips, T. J. and Taylor, K. E. (2003). An overview of results from the Coupled Model Intercomparison Project. *Global and Planetary Change* **37**: 103-133.

D'Agostino RB, Stephens MA (eds) (1986) *Goodness-of-Fit Techniques*, Marcel Dekker, Inc. 576pp.

Dickson RR, Namias J (1976) North American influences on the circulation and climate of the North Atlantic sector. Mon Weather Rev 104: 1255-1265.

Efron B (1982) *The Jackknife, the Bootstrap, and Other Resampling Plans*, CBMS-NSF Regional Conference Series in Applied Mathematics, 38, Society for Industrial and Applied Mathematics. 104pp.

El-Kadi AKA, Smithson PA (1992) Atmospheric classifications and synoptic climatology. Prog Phys Geog 16: 432-455.

Flato GM, Boer, GJ, Lee WG, McFarlane NA, Ramsden D, Reader MC, Weaver AJ (2000) The Canadian Centre for Climate Modeling and Analysis coupled model and its climate. Clim Dynam 16: 451-467.

Flato GM, Boer GJ (2001) Warming asymmetry in climate change simulations. Geophys Res Lett 28: 195-198.

Gates W, Boyle JS, Covey C, Dease C, Doutriaux C, Drach R, Fiorino M, Glecker P, Hnilo J, Marlais S, Phillips T, Potter G, Santer B, Sperber K, Taylor K, Williams D

(1999) An overview of the results of the Atmospheric Model Intercomparison Project (AMIP 1). B Am Meteorol Soc 80: 29-55.

Glowienka-Hense R (1990) The North Atlantic Oscillation in the Atlantic-European SLP. Tellus 42A: 497-507.

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dynam 16: 147-168.

Hansen J, Lacis A, Rind D, Russell G (1984) Climate sensitivity: Analysis of feedback mechanisms. *Climate Processes and Climate Sensitivity, Geophysical Monographs*, American Geophysical Union, 29: 130-163.

Hoerling MP, Kumar A (2002) Atmospheric response patterns associated with tropical forcing. J. Clim. 15: 2184-2203.

Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation and relationships to regional temperature and precipitation. Science 269: 676-679.

Hurrell JW (1996) Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperatures. Geophys. Res. Lett. 23: 665-668.

Hurrell J, Van Loon H (1997) Decadal variations in climate associated with the North Atlantic Oscillation. Climatic Change 36: 301-326.

Hurrell JW, Kushnir Y, Ottersen G, Visbeck M (eds) (2003) *The North Atlantic Oscillation: Climatic Significance and Environmental Impact.* Washington DC: American Geophysical Union, Geophysical Monograph Series 134, 279pp.

IPCC (2000) *Special Report on Emissions Scenarios*. Cambridge: Cambridge University Press, 612pp.

IPCC (2001) *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge and New York: Cambridge University Press, 881pp.

Johns TC, Carnell RE, Crossley JF, Gregory JM, Mitchell JFB, Senior CA, Tett SFB, Wood RA (1997) The second Hadley Centre coupled ocean-atmosphere GCM: Model description, spinup and validation. Clim Dynam 13: 103-134.

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, Ropelweski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40 reanalysis project. B Am Meteorol Soc 77: 437-471.

Kang I-S, Jin K, Lau K-M, Shukla J, Krishnamurthy V, Schubert SD, Waliser DE, Stern WF, Satyan V, Kitoh A, Meehl GA, Kanamitsu M, Galin VYa, Sumi A, Wu G, Liu Y, Kim, J-K (2002) Intercomparison of atmospheric GCM simulated anomalies associated with the 1997-98 El Niño. J Climate 15, 2791-2805.

Kirchhofer W (1974) Classification of European 500 mb patterns. Schwiezerische Meteorologische Anstalt, Institut Suisse de Meteorologie, Zurich 43: 1-16.

Kistler R, Kalnay E, Collins W, Saha S, White G, Woollen J, Chelliah M, Ebisuzaki W, Kanamitsu M, Kousky V, van den Dool H, Jenne R, Fiorino M (2001) The NCEP-NCAR 50 year reanalysis: Monthly mean CD-ROM and documentation B Am Meteorol Soc 82: 247-267.

Kushnir Y (1994) Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. J Clim 7: 141-157.

Lapp S, Byrne J, Kienzle S, Townshend I (2002) Linking global circulation model synoptics and precipitation for western North America. Int J Climatol 22: 1807-1817.

Leathers DJ, Yarnal B, Palecki MA (1991) The Pacific/North American teleconnection pattern and United States climate. Part I: Regional temperature and precipitation associations. J Climate 4: 517-528.

Leathers DJ, Palecki MA (1992) The Pacific/North American teleconnection pattern and United States climate. Part II: Temporal characteristics and index specification. J Climate 5: 707-716.

Lund IA (1963) Map-pattern classification by statistical methods. J Appl Meteorol 2: 56-65.

McAveney, B. J., Covey, C., Joussaume, S., Kattsov, V., Kitoh, A., Ogana, W., Pitman, A. J., Weaver, A. J., Wood, R. A. and Zhao, Z.-C. (2001). Model Evaluation. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggset al. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 881pp.

McFarlane N, Boer G, Blanchet J, Lazare M (1992) The Canadian Centre second generation General Circulation Model and its equilibrium climate. J Clim 5: 1013-1044.

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. Atmos Ocean 33: 809-825.

Osborn TJ, Briffa KR, Tett SFB, Jones PD (1999) Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. Clim Dynam 15: 685-702.

Ott RL (1993) An Introduction to Statistical Methods and Data Analysis, 4th Edition. Belmont: Duxbury Press. 1051 pp.

Paeth H, Hense A, Glowienka-Hense R, Voss R, Cubasch, U (1999) The North Atlantic Oscillation as an indicator for greenhouse-gas induced regional climate change. Clim Dynam 15: 953-960.

Pan Z, Christensen J, Arritt RW, Gutowski W, Takle ES, Otieno F (2001) Evaluation of uncertainties in regional climate change simulations. J Geophys Res-Atmos 106: 17735-17751.

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. Clim Dynam 16: 123-146.

Pryor SC, Barthelmie RJ (2003) Long term trends in near surface flow over the Baltic. Int J Climatol 23: 271-289.

Renshaw AC, Rowell DP, Folland CK (1998) Wintertime low-frequency weather variability in the North Pacific-American sector 1949-93. J Climate 11: 1073-1093.

Rogers JC (1984) The association between the North Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere. Mon Wea Rev 112: 1999-2015.

Schubert S (1998) Downscaling local extreme temperature changes in south-eastern Australia from the CSIRO Mark2 GCM. Int J Climatol 18: 1419-1438.

Sheridan SC (2003) North American weather-type frequency and teleconnection indices. Int J Climatol 23: 27-45.

Stephenson DB, Pavan V (2003) The North Atlantic Oscillation in coupled climate models: a CMIP1 evaluation. Clim Dynam 20: 381-399.

Stratton RA (1999) A high resolution AMIP integration using the Hadley Centre model HadAM2b. Clim Dynam 15: 9-28.

Tourre Y, Rajagopalan B, Kushnir Y (1999) Dominant patterns of climate variability in the Atlantic Ocean during the last 136 years. J Clim 12: 2285-2299. Walker GT (1924) Correlation of seasonal variations in weather IX: A further study of world weather. Mem. Indian Meteor. Dep. 24: 275-332.

Walker GT, Bliss EM (1932) World Weather V. Mem. of the Roy. Meteorol. Soc. 4: 53-84.

Wallace JM, Gutzler DS (1981) Teleconnections in the geopotential height field during the Northern Hemisphere Winter. Mon Weather Rev 109: 784-812.

Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: a comparison of methods, Water Resour Res 34: 2995-3008.

Willmott CJ (1987) Synoptic weather map classification: correlation versus sums-of-squares. Prof Geogr 39: 205-207.

Yarnal B (1985) A 500 mb synoptic climatology of Pacific North-west Coast winters in relation to climatic variability, 1948-49 to 1977-78. J Climatol 5: 237-252.

Yarnal B, Leathers DJ (1988) Relationships between interdecadal and interannual climatic variations and their effect on Pennsylvania climate. Ann Assoc Amer Geogr 78: 624-641.

Yarnal (1993) Synoptic Climatology in Environmental Analysis. Belhaven Press, London. 195pp.

Yin Z-Y (1994) Moisture conditions in the south-eastern USA and teleconnection patterns. Int J Climatol 14: 947-967.

CHAPTER 8: EMPIRICAL DOWNSCALING OF GCM OUTPUT

Schoof JT, Pryor SC. 2004. A comparison of empirically derived regional climate change scenarios for the Midwestern USA. In preparation for submission to the *Journal of Climate*

1 Introduction

Coupled atmosphere-ocean general circulation models (GCMs) indicate that increases in atmospheric greenhouse gas concentrations will result in global average surface air temperature increases of at least a few degrees Celsius by the end of the 21st century (Cubasch et al. 2001) and concurrent increases in precipitation associated with an enhanced hydrological cycle. However, the demonstrated ability of GCMs to credibly simulate climate is limited to annual mean climate and the climatological seasonal cycle over broad continental scales (McAveney et al. 2001). This shortcoming is primarily due to coarse horizontal grid spacing within GCMs (typically 300 to 500 km), resulting in a lack of information regarding the effects of mesoscale circulation and other regional effects on climate. Hence, techniques have been developed to relate information at large scales (at which GCMs are reliable) to local scale climate information, allowing generation of regional climate change scenarios (see reviews by Wilby and Wigley (1997) and Giorgi et al. (2001)).

The focus of this study is the generation of regional temperature and precipitation scenarios for the Midwest region of the USA using empirical downscaling techniques. Specific objectives of this study are (1) evaluation of GCM simulations for use in

empirical downscaling and (2) documentation of differences in scenarios produced with different GCMs and empirical downscaling methodologies.

Empirical downscaling has its roots in the numerical-statistical daily forecast methods of the National Weather Service (e.g., perfect prognosis (PP), Klein et al. 1959; model output statistic (MOS), Glahn and Lowry 1972), which produce weather forecasts by establishing statistical relationships between free atmosphere variables and surface variables. While the PP approach uses observed free atmosphere predictors to condition the transfer functions, the MOS approach employs model output variables and can therefore account for the internal bias and inaccuracy of the large-scale fields (Klein 1982, Sailor and Li 1999). Most empirical downscaling studies have employed PP approaches, although many have failed to adequately demonstrate that the GCM accurately simulates the predictor variables. Regardless of which approach is adopted, there are several requirements for successful application of empirical downscaling:

1) The free atmosphere predictors must be relevant to the local climate variable of interest. Typically, surface air temperature or precipitation are modeled from atmospheric circulation variables (e.g., mean sea-level pressure or 500-hPa geopotential height), either at a nearby grid points (e.g., Sailor and Li 1999) or in weather classification schemes (e.g., Schoof and Pryor 2001).

2) The free atmosphere predictor variables must be adequately simulated by the GCM. While this requirement is less important in studies which employ MOS-based empirical downscaling techniques, the inability of a GCM to simulate

recent climate and climate variability decreases the confidence in any projections derived from the model.

3) The statistical links between the free atmosphere predictors and local climate variables must be valid under changed climate conditions. While the inability to test this requirement is the major drawback of the empirical downscaling method, availability of long data series for model development may allow some testing of the robustness of relationships under a range of climate conditions.

4) The climate change signal must be fully represented by the predictor variable(s). While many downscaling studies have used circulation metrics as the only predictor variables, changes in regional climate will likely have sources other than changes in circulation, and the need to include other parameters (e.g., large scale temperature as a surrogate for increased radiation, or atmospheric humidity to differentiate between moist and dry episodes within weather classes) is being recognized (Charles et al. 1999; Giorgi et al. 2001; Huth 2002).

The availability of atmospheric reanalysis products now makes it possible to consider a larger number of atmospheric downscaling predictor variables (over a large, regular grid), although the variables archived from GCM experiments place a limitation on which predictors can be used for projecting regional climate change. Additionally, after several downscaling review articles (Hewitson and Crane 1996; Kattenberg et al.

1996; Wilby and Wigley 1997), there is still little consensus among downscaling studies regarding the proper choice of atmospheric predictor variables (Wilby and Wigley 2000).

Although many studies use a single downscaling method, application of multiple methods can help to identify the relative merits of a particular approach and consistent results from different downscaling methods can increase the overall confidence in the generated regional scenarios. The most common empirical downscaling approaches are based on transfer functions, which use direct relationships between the free atmosphere predictors and the surface variables of interest. These transfer functions vary in complexity from simple regression models to artificial neural networks (e.g., Kim et al. 1984; Schoof and Pryor 2001). Weather typing approaches, which relate surface climate variables to the large-scale state (usually defined in terms of circulation), are also commonly used (Goodess and Palutikof 1998; Schnur and Lettenmaier 1998; Bellone et al. 2000; Bardossy et al. 2002). Another alternative is use of weather generators, a class of random number generators which produces realistic climate sequences conditioned on the large-scale state (e.g., Wilks 1992; Wilks 1999). Many downscaling methods are really hybrid in nature, possessing characteristics of more than one approach (Wilby and Wigley 1997).

In this study, regional climate change scenarios are constructed by: (1) identifying optimal predictors (of daily maximum and minimum air temperature and total daily precipitation) from a suite of free atmosphere predictors (including sea level pressure, geopotential height, upper air temperature and specific humidity, and derived variables such as layer thickness, and geostrophic flow and vorticity components), (2) assessing the accuracy of the simulation of the key predictor variables by two GCMs, and (3)

developing, applying and evaluating downscaling techniques. We introduce an innovative two-stage downscaling method, in which changes in the seasonal cycle are developed using a MOS-based regression approach and deviations from the seasonal cycle are constructed using the analog method (Zorita and von Storch 1999), and a stochastic weather generator (Wilks 1992) in a PP context.

2 Study Region

The geographic focus of this study is the Midwest region of the USA, which extends from approximately 35° to 50°N latitude and from 75° to 95°W longitude (Figure 1), although the techniques used have much broader applications. This region is highly sensitive to climate change for several reasons. First, the region is a major agricultural center, producing a large proportion of the nation's corn and soybeans. Secondly, with the exception of the polar ice caps, the Great Lakes are the world's largest source of fresh water and provide drinking water and hydroelectric power to the region. The lakes additionally serve as a major transportation system, linking the region with the Atlantic Ocean via the St. Lawrence Seaway. Key climate change issues for the Midwest region of the United States include reductions in lake and river levels, increases in heat related stress and mortality, shifts in agricultural productivity, and negative ecological impacts, such as runoff of excess nutrients used for fertilization (Sousounis and Albercook 2000). The specific parameter foci for this downscaling study are maximum and minimum daily surface air temperature (T_{max} and T_{min}) and precipitation (P) due to their influence on the impacts listed above.

During the winter, the synoptic scale climate of the Midwestern United States is characterized by high day-to-day variability associated with the position of the polar jet stream, which is often located over the region, and the associated polar front. Mid-latitude cyclones, which form in the presence of these features, are the main sources of precipitation in the region during the winter months. During summer, the polar front retreats into Canada and frontal passages are less common. During these months, most precipitation is convective in nature and moisture available is largely determined by the position and intensity of the subtropical (Bermuda) high (often located within or to the southeast of the study area). Variations in the position and intensity of the subtropical high and other semi-permanent pressure systems affecting the region (e.g., the Hudson Bay low), as well as the tracking and intensity of synoptic scale phenomena in the study region, have been linked to larger pressure oscillations (or teleconnection patterns), specifically the North Atlantic Oscillation (NAO) and Pacific North American (PNA) pattern.

According to Folland et al. (2001), the northern part of the study region (i.e., the upper Great Lakes) has warmed by approximately 2K over the past 100 years, while the southern and eastern parts of the region have warmed only slightly or even cooled. However, during the last quarter century, all parts of the region have warmed (trends of 0.4 to 0.8K per decade), with the largest warming during the winter. Precipitation also increased over most of the region during the last century, with increases as large as 30% in some regions, although in the southeast, precipitation has been relatively stable or decreased slightly.

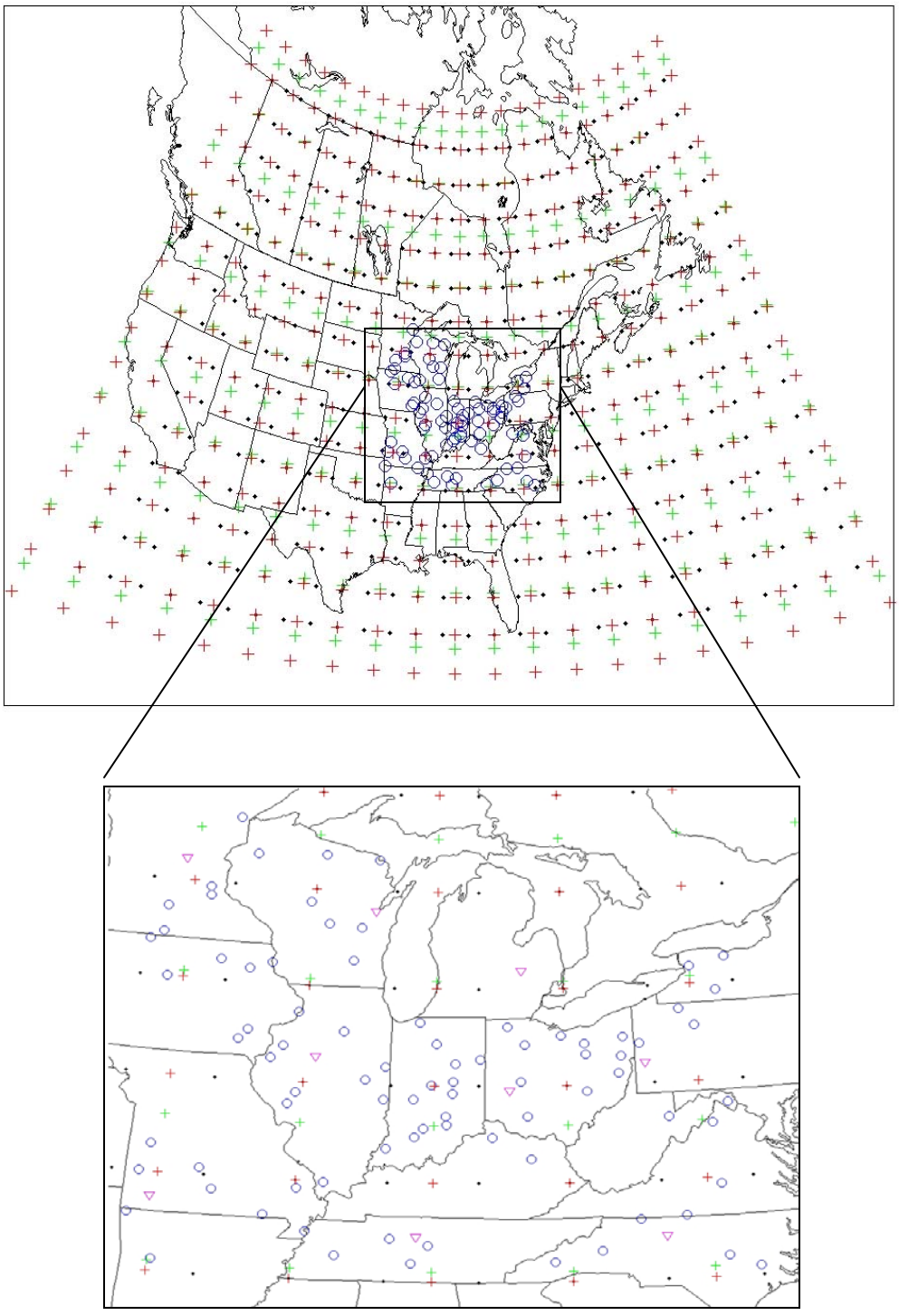

Figure 1. Map of the Midwest region of the United States, including the locations of the NCEP/NCAR and ECMWF reanalysis predictor grid points (•), HadCM3 grid points (+), CGCM2 grid points (+), and surface (USHCN/D) stations (o). See Section 3 for a complete description of each data set.

3 Data

3.1 Reanalysis data

In this research, data are used from both the NCEP/NCAR reanalysis project (Kalnay et al. 1996) and the ECMWF reanalysis project (Gibson et al. 1997). These data are gridded at a horizontal resolution of 2.5° × 2.5°, with either daily or twice daily observations on multiple atmospheric levels. The atmospheric model used for the NCEP/NCAR reanalysis is spectral (T62) with transformation to a Gaussian grid for calculation of nonlinear quantities and physics. The vertical domain is divided into 28 levels with enhanced resolution near the bottom and the top. The atmospheric model used for the ECMWF reanalysis has T159 spherical-harmonic representation for basic dynamic fields, with a reduced Gaussian grid of approximately uniform 125 km spacing for surface and other grid-point fields. The vertical domain is divided into 60 levels. Here we use reanalysis data from 1958-2001. Specific humidity is the only ECMWF variable used. For other large-scale variables, data are extracted from the NCEP/NCAR reanalysis.

3.2 Surface station data

While historical surface temperature and precipitation data are susceptible to a number of biases and inhomogeneities resulting from changes in the station environment or observing practices (e.g., urbanization, station moves, instrumentation and time of observation changes (Jones 1994; Jones et al. 1997; Peterson et al. 1998), construction of the downscaling models requires high quality surface station data from throughout the study region. Hence, we use data from the Daily United States Historical Climatology

Network (USHCN/D; Easterling et al. 1999) and include only stations that have a data record that is at least 95% complete over the period 1958-2001 and at least 90% complete within each individual year. To ensure consistency between the station data and large-scale free atmosphere predictors, only stations with a consistent observation time are included in this study. With these constraints, 84 stations are available within the study region (Figure 1, Table 1). For each station, maximum and minimum daily air temperature and total daily precipitation are available for at least the period 1958-2001.

Table 1 List of the 84 USHCN/D stations used in this study.

Station Name	Lat (°N)	Lon (°W)	Station Name	Lat (°N)	Lon (°W)
Gravette, AR	36.44	94.45	Lamar, MO	37.51	94.27
Subiaco, AR	35.30	93.66	Lebanon, MO	37.67	92.66
Anna, IL	37.47	89.24	Marble Hill, MO	37.30	89.97
Hoopston, IL	40.47	87.67	Mountain Grove, MO	37.16	92.27
Jacksonville, IL	39.74	90.20	Angelica, NY	42.30	78.02
La Harpe, IL	40.59	90.97	Buffalo, NY	42.92	78.74
Monmouth, IL	40.92	90.64	Rochester, NY	43.14	77.67
Ottawa, IL	41.34	88.92	Kinston, NC	35.22	77.54
Paris, IL	39.62	87.70	Morganton, NC	35.76	81.69
Urbana, IL	40.11	88.24	Mount Airy, NC	36.52	80.62
White Hall, IL	39.44	90.39	Smithfield, NC	35.52	78.35
Anderson, IN	40.11	85.72	Waynesville, NC	35.49	82.97
Berne, IN	40.67	84.95	Cadiz, OH	40.27	81.00
Columbus, IN	39.21	85.92	Chippewa Lake, OH	41.05	81.94
Greencastle, IN	39.64	86.85	Findlay, OH	41.05	83.67
Greenfield, IN	39.79	85.75	Hillsboro, OH	39.21	83.62
LaPorte, IN	41.61	86.72	Millport, OH	40.72	80.91
Marion, IN	40.57	85.67	Norwalk, OH	41.27	82.62
Oolitic, IN	38.89	86.55	Philo, OH	39.84	81.92
Princeton, IN	38.36	87.59	Urbana, OH	40.11	83.79
Rochester, IN	41.07	86.22	Warren, OH	41.21	80.82
Seymour, IN	38.99	85.91	Wauseon, OH	41.52	84.16
Shoals, IN	38.67	86.80	Wooster, OH	40.79	81.92
Whitestown, IN	40.01	86.35	New Castle, PA	41.02	80.37
Charles City, IA	43.05	92.67	Ridgway, PA	41.42	78.75
Clinton, IA	41.80	90.27	Warren, PA	41.86	79.16
Estherville, IA	43.42	94.84	Dickson, TN	36.07	87.39
Fairfield, IA	41.04	91.95	Jackson, TN	35.62	88.84
Fayette, IA	42.86	91.80	Lewisburg, TN	35.46	86.80
Fort Dodge, IA	42.51	94.2	Murfreesboro, TN	35.92	86.37
Washington, IA	41.29	91.69	Danville, VA	36.59	79.39
Farmers, KY	38.12	83.55	Farmville, VA	37.34	78.39
Williamstown, KY	38.66	84.62	Woodstock, VA	38.91	78.47
Iron Mountain, MI	45.79	88.09	Martinsburg, WV	39.41	77.99
Cloquet, MN	46.71	92.52	Parsons, WV	39.11	79.67
Fairmont, MN	43.64	94.47	Hancock, WI	44.12	89.54
Farmington, MN	44.67	93.19	Marshfield, WI	44.66	90.14
Minneapolis, MN	44.89	93.22	Minocqua Dam, WI	45.89	89.74
New Ulm, MN	44.30	94.45	Oshkosh, WI	44.04	88.55
Appleton City, MO	38.21	94.04	Prairie Du Chien, WI	43.04	91.16
Caruthersville, MO	36.21	89.67	Spooner, WI	45.82	91.89
Doniphan, MO	36.59	90.82	Watertown, WI	43.19	88.74

3.3 General circulation model (GCM) data

Because substantial differences exist between projections from climate models, application of a single model does not provide results that adequately reflect the uncertainty inherent in regional climate scenario generation (Cubasch et al. 2001). Therefore, in this study, the output from transient simulations from two GCMs is used: 1) the Hadley Centre 3rd generation coupled oceanic-atmospheric general circulation model (HadCM3; Gordon et al. 2000; Pope et al. 2000) and 2) the Canadian Centre for Climate Modelling and Analysis (CCCma) 2nd Generation coupled general circulation model (CGCM2; Flato et al. 2000; Flato and Boer 2001). HadCM3 and CGCM2 are chosen to provide a range of characteristics associated with the current generation of coupled models and therefore differ in several respects. First, HadCM3 is a Cartesian model (approximate horizontal resolution of 2.5° latitude × 3.75° longitude with 19 vertical levels, see Figure 1), whereas CGCM2 is a spectral model (T32, approximate horizontal resolution of 3.75° latitude × 3.75° longitude with 10 vertical levels, see Figure 1). Second, CGCM2 employs flux adjustments (Hansen et al. 1984), while HadCM3 has slightly higher oceanic resolution (1.25°×1.25° vs. 1.8°×1.8°) and does not employ flux adjustments.

Prior to use in this study, both GCM grids were interpolated to the 2.5° × 2.5° reanalysis grid (Figure 1) using an inverse distance based interpolation algorithm. Additionally, the output from HadCM3 was archived with a 360-day year. For comparison with observed data and output from CGCM2, this 360-day year was projected onto a 365-day year.

Consistent with studies that report lack of agreement between GCMs and observations near the surface and a large degree of variability in surface parameters within GCM grid cells (e.g., Sailor and Li 1999), Figure 2 indicates large differences in the mean daily T_{max}, T_{min} and precipitation during 1990-2001 from the two GCMs and the USHCN/D station data.

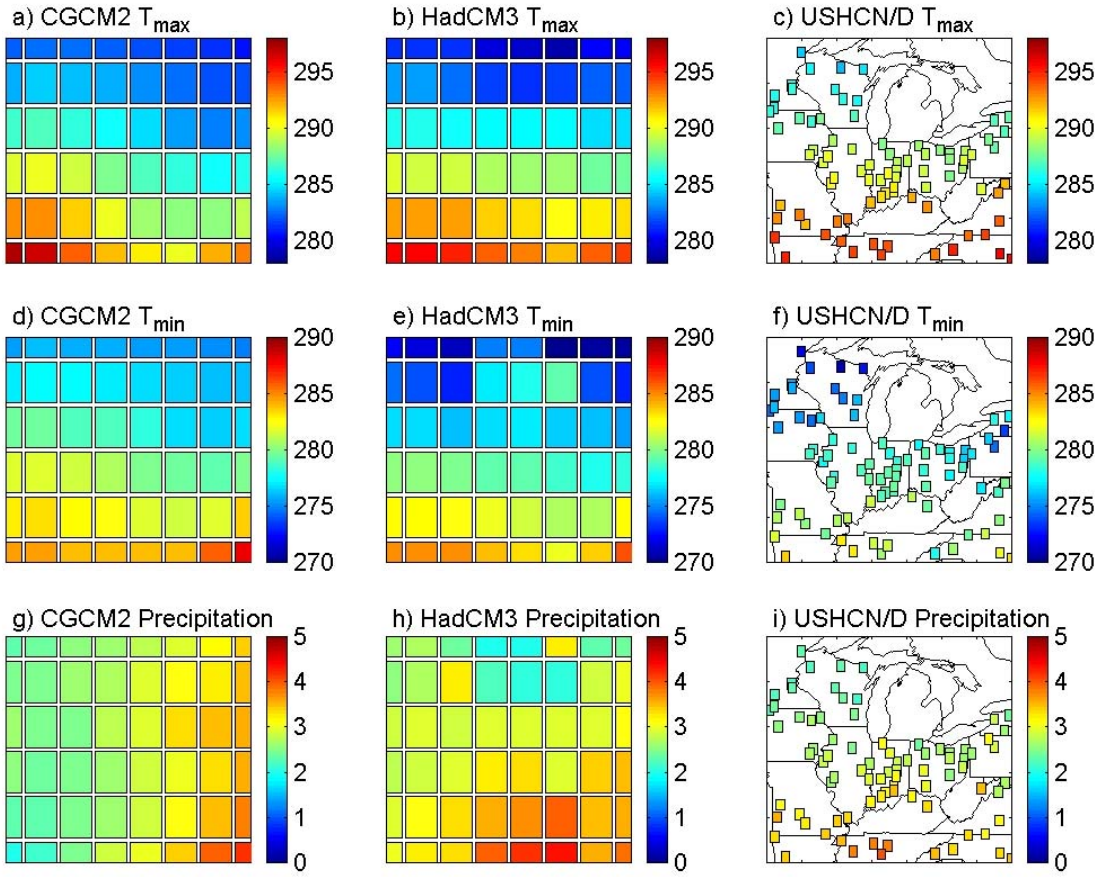

Figure 2. 1990-2001 mean values of daily T_{max} and T_{min} (K), and precipitation (mm/day) within the study area: a) CGCM2 T_{max}, b) HadCM3 T_{max}, c) USHCN/D T_{max}, d) CGCM2 T_{min}, e) HadCM3 T_{min}, f) USHCN/D T_{min}, g) CGCM2 precipitation, h) HadCM3 precipitation, and i) USHCN/D precipitation. The area shown in each frame is the inset on Figure 1.

The GCM experiments used here were conducted using the SRES A2 emissions scenario (IPCC 2000), which results in global carbon dioxide (CO_2) emissions from industry and energy in 2100 that are almost four times the 1900 value and emissions from land use change by 2100 are close to zero, leading to a global CO_2 emission in 2100 of almost 28 GtC yr^{-1}. This emissions scenario equates to a moderate to high greenhouse gas cumulative emission for 1990 to 2100 as a result of projected population growth and fairly slow introduction of alternative technologies, and is used in this study to provide an upper bound on likely climate change and hence a high signal to noise ratio when comparing current and future climates.

4 Methodology

4.1 Identification of downscaling predictor variables

The first step in the construction of climate change scenarios is the identification of large scale predictor variables which are strongly linked to the surface variables of interest. Table 2 shows the potential downscaling predictors drawn from previous studies and considered herein. The predictors are chosen to reflect (1) atmospheric circulation, (2) lower atmospheric air temperature, (3) atmospheric humidity, and (4) flow and vertical motion. The geostrophic flow and vorticity components were derived from the sea-level pressure field using previously published methods (see Dessouky and Jenkinson 1975; Jenkinson and Collison 1977; Jones et al. 1993).

Table 2. Upper-atmosphere and surface predictor variables and examples of their use in previous downscaling studies, where the predictands are shown in parentheses (T=surface air temperature, P=surface precipitation).

Predictor (units)	Abbrev.	Previous studies
Upper-atmosphere variables		
850 hPa geopotential height (m)	Z_{850}	Sailor and Li 1999 (T, P)
500 hPa geopotential height (m)	Z_{500}	Hewitson and Crane 1996 (P) Weichert and Burger 1998 (T, P, vapor pressure) Sailor and Li 1999 (T, P) Palutikof et al. 2002 (T, P)
850-500 hPa thickness (m)	THICK	Cavazos 1997 (P) Kidson and Thompson 1998 (T, P) Cavazos 1999 (P)
850 hPa specific humidity (kg/kg)	Q_{850}	Crane and Hewitson 1998 (P) (uses Q at 1000 and 500 hPa) Murphy 1999 (T, P)
850 hPa relative humidity (%)	RH_{850}	Sailor and Li 1999 (T, P)
Surface variables		
Mean sea-level pressure (hPa)	SLP	Zorita et al. 1995 (P) Hewitson and Crane 1996 (P) Heyen et al. 1996 (sea level) Schubert 1998 (T_{max}, T_{min}) Palutikof et al. 2002 (T, P)
Zonal component of geostrophic flow (hPa /10° latitude at grid point latitude)	GEOW	Buishand and Brandsma 1997 (T, P)
Meridional component of geostrophic flow (hPa /10° latitude at grid point latitude)	GEOS	Wilby and Wigley 1997 (P) Kidson and Thompson 1998 (T, P)
Strength of the resultant geostrophic flow (hPa /10° latitude at grid point latitude)	GEOWS	Wilby 1998 (T, P)
Westerly shear vorticity (hPa /10° latitude at grid point latitude, per 10° latitude)	GEOZW	
Southerly shear vorticity (hPa /10° latitude at grid point latitude, per 10° latitude)	GEOZS	
Total shear vorticity (hPa /10° latitude at grid point latitude, per 10° latitude)	GEOZT	

To assess these predictors, correlations were computed between the daily values of each of the predictors (at each of the 435 grid points shown in Figure 1) and each surface variable at each station for each climatological season (DJF, MAM, JJA, SON). In each season, the thickness of the 850-500 hPa layer exhibits the strongest correlation with daily T_{max} and T_{min} (>0.80 for most stations during spring and autumn and >0.60 for most stations during winter and summer), while 500-hPa geopotential height, 850-hPa specific humidity, and sea level pressure also exhibit strong correlations. For precipitation, correlation coefficients with the large-scale variables are lower and the total geostrophic shear vorticity was identified as the single best variable (with correlations generally in the 0.30-0.50 range), while 850-hPa specific humidity also exhibits strong correlations with precipitation at most stations. In accord with previous studies (e.g., Brinkmann 2002), for all three surface variables and most predictor variables, the maximum correlation was found at grid points located remotely from the station location.

4.2 Evaluation of grid point GCM simulations

A key prerequisite for use of grid point GCM data in empirical downscaling is that the GCM accurately reproduces the large-scale variables. To evaluate this assumption, the data were evaluated using the method proposed by Chervin (1981) and later used and refined by Portman et al. (1992). Following their approach, 100 expanded datasets were constructed from the reanalysis data and compared to the GCM output using a bootstrap resampling technique. First, 12 years were randomly chosen from the reanalysis time series (1958-2001). For each GCM and variable (Z_{500}, Z_{850}, THICK, Q_{850}, RH_{850}, SLP, GEOS, GEOW, GEOWS, GEOZW, GEOZS, and GEOZT, see Table

2), two test variates were computed for each grid box in the study area for each of the 100 random samples of 12-years from the reanalyses. The first variate, a scaled difference of means of the GCM and observed time series is given by:

$$r_1 = \frac{(\mu_a - \mu_b)}{\sqrt{\sigma_a^2 + \sigma_b^2}} \quad (1)$$

where μ_a and μ_b are the means of the GCM-simulated and observed time series, respectively, and σ_a^2 and σ_b^2 are the variances of the GCM-simulated and observed time series, respectively. The second variate is a simple ratio of standard deviations of the GCM simulated (σ_a) and reanalysis (σ_b) time series given by:

$$r_2 = \frac{\sigma_a}{\sigma_b} \quad (2)$$

The null hypotheses that the observations (bootstrap samples) and GCMs (first 12 years) have the same means and variances can be rejected if a sufficient number of r_1 and r_2 values lie outside of the specified confidence intervals. To examine a range of temporal scales, separate analyses are conducted at annual, seasonal, and monthly time steps. It should be noted that 12 years represents a relatively small amount of data for model evaluation and precludes analysis of grid point GCM data at higher (e.g., daily) time steps. This 12-year evaluation period is dictated by the short period of overlap between the transient GCM and reanalysis data (1990-2001).

Comparison of the 1990-2001 GCM simulations with the 1958-2001 reanalysis data is based on the assumption that the 1958-2001 reanalysis data are also representative of the 1990-2001 period. To test this assumption, 12 years were randomly chosen (100 times) from the 1958-2001 reanalysis data and compared to the 1990-2001 data using the

r_1 and r_2 statistics. For each variable, r_1 and r_2 values were within the confidence intervals for each random sample and annual, seasonal, and monthly timescales. Table 3 shows the number of grid points (out of 435) for which less than 95% of r_1 and r_2 values lie within the reanalysis-derived confidence interval for each variable and season and indicates a poor level of agreement between GCM and reanalysis grid point data for several variables.

All HadCM3 Z_{500} grid points exhibit values of r_1 and r_2 within the confidence bounds for at least 95% of the bootstrap samples (Table 3). In the extreme southeast corner of the study area, CGCM2 Z_{500} grid point means are statistically different from those in the reanalysis, with the greatest spatial extent of significant differences during summer. Additionally, due to underestimated variance, CGCM2 Z_{500} r_2 values are outside of the specified confidence bounds for large areas of the western Atlantic Ocean, especially during winter.

For both GCMs, only a small number of grid points have Z_{850} r_1 values that are outside of the confidence interval (Table 3). For CGCM2, these grid points are in the southeast corner of the domain, while for HadCM3, they are located in the southwest corner of the domain. With the exception of a small number of grid points during winter in CGCM2, the data and statistical tests used here do not imply a difference in the variances of observed and GCM simulated Z_{850} values on seasonal timescales for either model (Table 3). As expected, results for 850-500 hPa thickness are similar to those for Z_{500} and Z_{850}.

Both GCMs exhibit disagreement with observations with respect to mean SLP in the southwest part of the study region, particularly during summer, while both GCMs

simulate the variance of SLP well on seasonal timescales, with nearly all grid points producing values of r_2 within the confidence bounds for all 100 bootstrap samples in each season (exceptions are a few grid points in the southwest during summer).

Differences between observed and simulated grid point means and variances of the geostrophic flow variables (GEOS, GEOW, and GEOWS) are not statistically significant at most grid points, with the exception of mean values during the summer (Table 3). The mean values of the geostrophic vorticity variables are also well simulated by both GCMs, with significant differences at only a small percentage of grid points during each season. Similar results were found for seasonal variances of the vorticity variables as simulated by CGCM2. The seasonal variances of HadCM3-derived geostrophic vorticity variables lack agreement with the observed variances (Table 3). These results suggest that HadCM3 vorticity values are more variable (both spatially and temporally) than those derived from the reanalysis, but that the mean values generally agree on seasonal timescales.

For each season both GCMs have grid points at which humidity simulations (Q_{850} for CGCM2 and RH_{850} for HadCM3) are significantly different from the reanalysis in terms of both the mean and variance (Table 3). For both GCMs, the greatest number of poorly simulated grid points occurs during summer when r_1 values lie outside of the confidence interval for large areas in the western and southeastern regions of the domain. Regions in which the variances are found to be significantly different from the reanalysis generally coincide with those for which the means are different, although they are less extensive.

The results of this GCM evaluation suggest that for many of the predictor variables grid point means and variances are not in agreement with observations and hence, better downscaling models will result from MOS-based downscaling (rather than PP) or by using GCM information aggregated across larger scales. The evaluation of synoptic phenomena presented in Schoof and Pryor (2004) shows that both HadCM3 and CGCM2 are capable of reproducing the range of 500-hPa circulation types in the study area as manifest in the reanalysis data. Hence, the downscaling method employed in this study is a combination of MOS-based downscaling and PP downscaling using synoptic circulation methods, which incorporate GCM information on larger, and hence, more reliable, scales.

Table 3. Number of grid points (out of 435) for which less than 95% of r_1 and r_2 values lie within the specified confidence interval at the seasonal timescale.

CGCM2 r_1											
	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	67	4	87	12	22	2	14	0	10	15	15
MAM	75	3	97	21	52	6	4	3	14	10	10
JJA	153	36	151	35	109	52	57	78	41	50	44
SON	93	25	87	0	48	2	5	0	16	7	9
CGCM2 r_2											
	Z_{500}	Z_{850}	THICK	SLP	Q_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	105	9	65	0	5	0	0	0	3	0	2
MAM	53	0	5	0	22	0	0	0	0	1	0
JJA	0	0	0	7	47	3	13	6	22	54	42
SON	13	0	10	3	51	0	3	0	3	10	0
HadCM3 r_1											
	Z_{500}	Z_{850}	THICK	SLP	R_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	0	3	0	29	48	17	0	3	40	27	24
MAM	0	0	0	24	22	10	0	17	21	51	48
JJA	0	28	78	120	107	80	68	79	84	80	80
SON	0	3	0	47	43	24	5	4	26	13	15
HadCM3 r_2											
	Z_{500}	Z_{850}	THICK	SLP	R_{850}	GEOS	GEOW	GEOWS	GEOZS	GEOZW	GEOZT
DJF	0	0	0	0	64	0	6	0	399	317	349
MAM	0	0	0	1	5	0	0	0	424	430	427
JJA	0	0	0	0	21	0	0	9	435	435	435
SON	0	0	0	0	21	0	0	0	402	413	405

4.3 Downscaling methods

The findings presented in Sections 4.1 and 4.2 place limitations on the types of downscaling strategies that can be used for these particular models and variables in our study region. Therefore, we devised a two-stage downscaling methodology consisting of (1) MOS-based downscaling of the seasonal cycles of daily T_{max} and T_{min} using multiple linear regression with GCM-derived grid cell level predictors and (2) downscaling of perturbations of T_{max} and T_{min} from their respective seasonal cycles and daily precipitation amounts using an analog method and a stochastic weather generator, both of which incorporate synoptic-scale GCM information. The former generates the average seasonal cycles of daily T_{max} and T_{min} for the projected window, while the latter allows

examination of the daily evolution (i.e., variability of daily values around the seasonal cycle) of both temperature and precipitation.

4.3.1 MOS-based downscaling of the seasonal cycle

For each station, annual cycles of T_{max} and T_{min} were constructed by fitting harmonics (3) to the mean values on each day of the year. In the same way, annual cycles were constructed for each of the large-scale predictors from the GCMs on a decadal basis. Regression models were then used to predict the values of the surface parameters on any day of the year based on the values of the large scale GCM predictors on that day:

$$T_{max} = \beta X + \varepsilon \quad (3a)$$

$$T_{min} = \beta X + \varepsilon \quad (3b)$$

where T_{max} and T_{min} are the daily values of maximum and minimum temperature at a USHNC/D station, X are the daily values of GCM grid point predictors, β is a matrix of coefficients (defined separately for T_{max} and T_{min} at each station), and ε is an error term. The models were trained using data from the period of overlapping observations and GCM simulations (1990-2001) and applied to the transient GCM simulations.

The MOS-based downscaling of the seasonal cycles of T_{max} and T_{min} began with a pool of variables analyzed in a stepwise fashion. However, it was found that nearly the same variables appeared in the models for each station and that three predictor variables were generally needed to model the full variability in the seasonal cycles (each of the two predictor models tested failed to reproduce the range of the observed seasonal cycles): 850-500 hPa thickness, sea level pressure and 850-hPa specific humidity. Figure 3 shows

the locations of the 850-500 hPa thickness, sea level pressure, and 850-hPa humidity grid points which have the maximum absolute correlation with T_{max} and T_{min} at each of the 84 stations, and hence were used in the seasonal cycle downscaling. For both models, the most explanatory 850-500 hPa thickness grid points for both temperature variables (T_{max} and T_{min}) were located to the southwest of the primary study region, and remote from the station locations (see Figure 3). For sea level pressure, the predictor grid points for CGCM2 were consistently located at grid points southwest of the station location and exhibit negative correlation with station temperatures consistent with warm air advection into the study region when low pressure exists in those locations. HadCM3 sea level pressure predictor grid points were less consistently located, but were still generally indicative of the long-wave pattern which occurs over the study region. Specifically, the sea level pressure grid points in the northwest and southeast corners of the study region (Figure 3 b, d) are very near the locations used to define the PNA index (Wallace and Gutzler 1981) which describes this pattern. The CGCM2 850-hPa specific humidity predictors were located southwest of the station locations for both T_{max} and T_{min}, although some stations in the eastern part of the study area exhibit higher correlations with 850-hPa specific humidity grid points in oceanic grid cells. For T_{min}, there is greater variability in the location of the optimum grid point location, although for most stations, the grid point location is similar to that for T_{max}. The large degree of variability in the locations of the 850-hPa relative humidity predictor grid points results from the weaker relationship with the temperature variables relative to specific humidity.

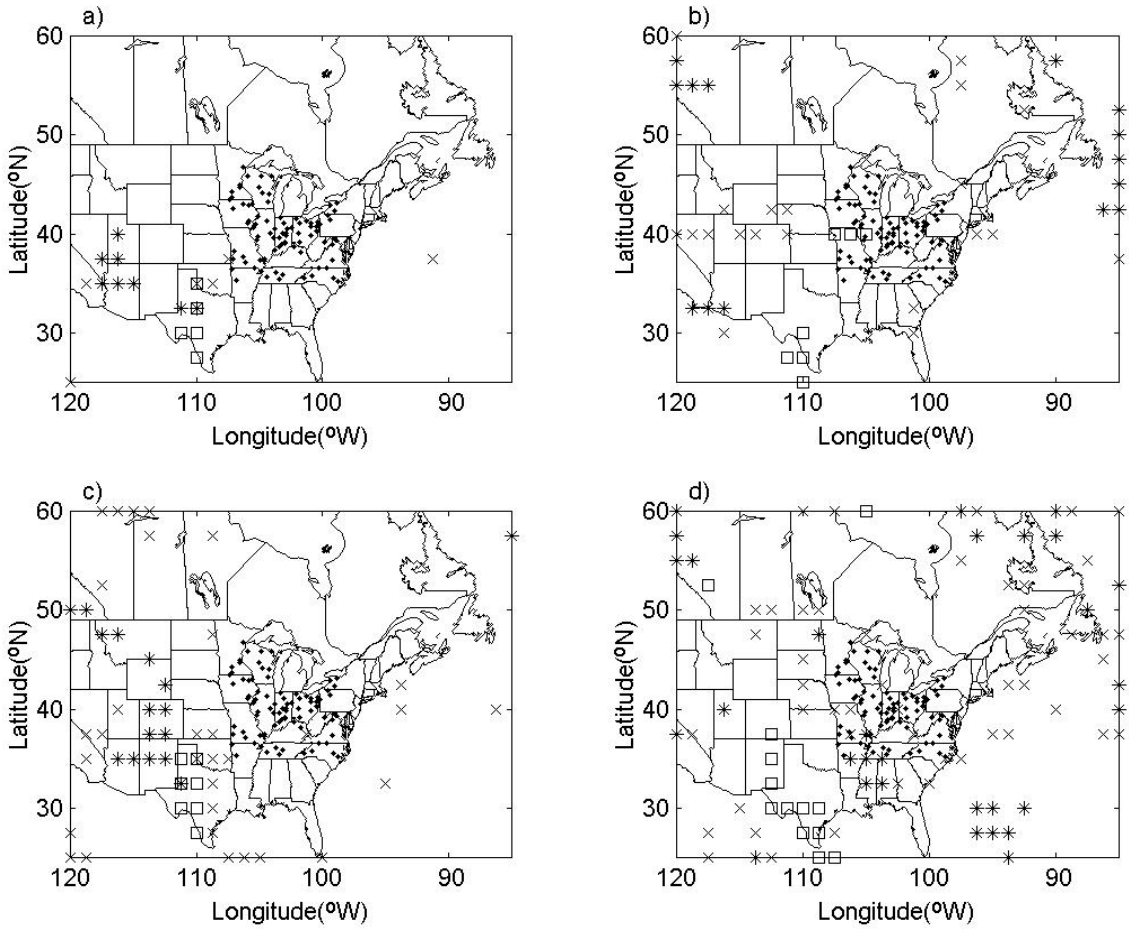

Figure 3. Optimum grid point locations for downscaling of the seasonal cycles of T_{max} and T_{min} at USHCN/D stations. The maps shown the USHCN/D station locations (•) and the 850-500 hPa thickness (□), sea-level pressure (*), and 850-hPa humidity (x) grid points for a) CGCM2 T_{max}, b) HadCM3 T_{max}, c) CGCM2 T_{min}, and d) HadCM3 T_{min}.

4.3.2 The analog method for downscaling daily variability

In the analog method, historical observations are searched to find the most similar 'state' (usually defined as the minimum value of the sum of the squared differences over all grid points) to that produced by the GCM. The values of the surface variables are then modeled as those which occurred with the historical analog. While only a few studies have applied the analog method to climate downscaling (Zorita et al. 1995; Cubasch et al. 1996; Zorita and von Storch 1999), it has long been used in weather forecasting and

short-term climate prediction (Lorenz 1969; Kruizinga and Murphy 1983; van den Dool 1994).

Based on the findings presented in Section 4.1, we used gridded values of 850-500 hPa thickness to find analogs for T_{max} and T_{min}, and total geostrophic shear vorticity to find analogs for precipitation. Because similar large scale states could occur in different seasons, the residual (from the seasonal cycle) derived from the analog day was added to the seasonal cycle (as predicted using the method described in Section 4.3.1) on the downscaled day.

4.3.3 Stochastic weather generator for downscaling daily variability

The stochastic weather generator used in this research is based on a modified version of the WGEN model (Richardson and Wright 1984). Using parameters estimated from observed data, the model traditionally generates daily values of precipitation occurrence (binary), precipitation amount, T_{max}, T_{min}, and solar radiation. Since the variables of interest in this study are T_{max}, T_{min}, and precipitation, solar radiation was not simulated, resulting in a modified version of WGEN. An additional modification, which allows adaptation of the weather generator for GCM downscaling, is the specification of changes in the seasonal cycles of the surface variables (described in Section 4.1) and separate parameter definitions for each large-scale weather type (defined below). Based on findings presented in a previous study (Schoof and Robeson 2003), all weather generator parameters were also computed separately for each station and, as with the analog approach, the resulting surface variables are perturbations from the seasonal cycle.

4.3.3.1 Weather classification

To derive weather types, the eigenvector/clustering technique was applied to the 1958-2001 reanalysis data. To accommodate differences in variables archived from the GCM experiments, two separate classifications were conducted. Specifically, the CGCM2 classification is based on 850-500 hPa layer thickness, sea-level pressure, and 850-hPa specific humidity, while the HadCM3 classification utilizes 850-500 hPa thickness, sea-level pressure, and 850-hPa relative humidity. Cluster analysis of significant principal components (PCs) from both classifications resulted in 11 weather types for the 1958-2001 data. Each GCM day was then classified by using the projection of the GCM output on the reanalysis-derived PCs and assigning the day to the reanalysis-derived class with the most similar PC scores. For each weather type, the weather generator then produced daily values of T_{max}, T_{min}, and precipitation using the appropriate parameters and inter-relationships, as described below.

4.3.3.2 Precipitation occurrence

The stochastic weather generator used in this study simulates precipitation occurrence using a Markov chain model, in which the occurrence of precipitation depends on two parameters: p_{01}, the probability of a wet day following a dry day, and p_{11}, the probability of a wet day following a wet day. Depending on the wet/dry status and weather type of the previous day, a uniform [0,1] random number was compared to the appropriate transition probability. If the random number was less than the transition probability, a wet day was simulated. Otherwise, a dry day was simulated. Wilks (1999) found this type of model correctly reproduced the precipitation occurrence characteristics

for stations in the central and eastern USA, but suggested that higher order models may be needed in other regions.

4.3.3.3 Precipitation amount

Once the precipitation occurrence algorithm indicated a wet day, a precipitation amount must be drawn from a chosen statistical distribution. The most common choice for precipitation amount simulation has been the gamma distribution (Wilks and Wilby 1999). However, several studies have indicated that the mixed-exponential distribution provides a better overall fit (and a particularly better fit for large precipitation amounts) than the gamma distribution (Foufoula-Georgiou and Lettenmaier 1987; Wilks 1999), and therefore, the mixed-exponential distribution was used here. The mixed-exponential distribution is a probability mixture of two single parameter exponential distributions, with probability density function:

$$f_r = \frac{\alpha}{\mu_1}\exp\left[\frac{-r}{\mu_1}\right] + \frac{1-\alpha}{\mu_2}\exp\left[\frac{-r}{\mu_2}\right] \quad (4)$$

where r is again the daily precipitation amount, with the mean and variance of wet day precipitation amount given by $\alpha\mu_1 + (1-\alpha)\mu_2$ and $\alpha\mu_1^2 + (1-\alpha)\mu_2^2 + \alpha(1-\alpha)(\mu_1-\mu_2)^2$, respectively. These parameters were fitted separately for each weather type and each station.

4.3.3.4 Simulation of other variables

SWGs produce variables other than precipitation with a first order multiple autoregression, first described by Matalas (1967) and given by:

$$X_i = AX_{i-1} + B\varepsilon_i \quad (5)$$

where X_i is a matrix containing the current day's standardized values of the variables and X_{i-1} is a matrix containing the previous day's standardized values of the variables, ε_i is a vector of independent values from a standard Normal distribution, and A and B are matrices given by

$$A = M_1 M_0^{-1} \quad (6)$$

$$BB^T = M_0 - M_1 M_0^{-1} M_1^T \quad (7)$$

where M_o is the matrix of lag-0 cross correlations and M_1 is the matrix of lag-1 cross correlations. While A can be directly computed, B is computed by defining a new matrix $Z=BB^T$ (see Greene 2000). Then $Z=CLC^T$, where C is the matrix of eigenvectors of BB^T and L has the eigenvalues of BB^T on the diagonal and zeros elsewhere. B can then be computed as $B=CL^{1/2}C^T$. After generation of the residual series with (5), daily harmonics were used to produce dimensional values of the input variables.

5 Results

5.1 MOS-based downscaling of the T_{max} and T_{min} seasonal cycles

The seasonal cycles of T_{max} and T_{min} constructed with using regression models with the optimum grid points for each USHCN/D station exhibit close agreement with the observed seasonal cycles (see example shown in Figure 4). For the entire annual cycle, and within each season, root mean squared differences between observed and modeled daily T_{max} and T_{min} values are less than 1K for all stations, and less than 0.5 K for most stations during 1990-2001.

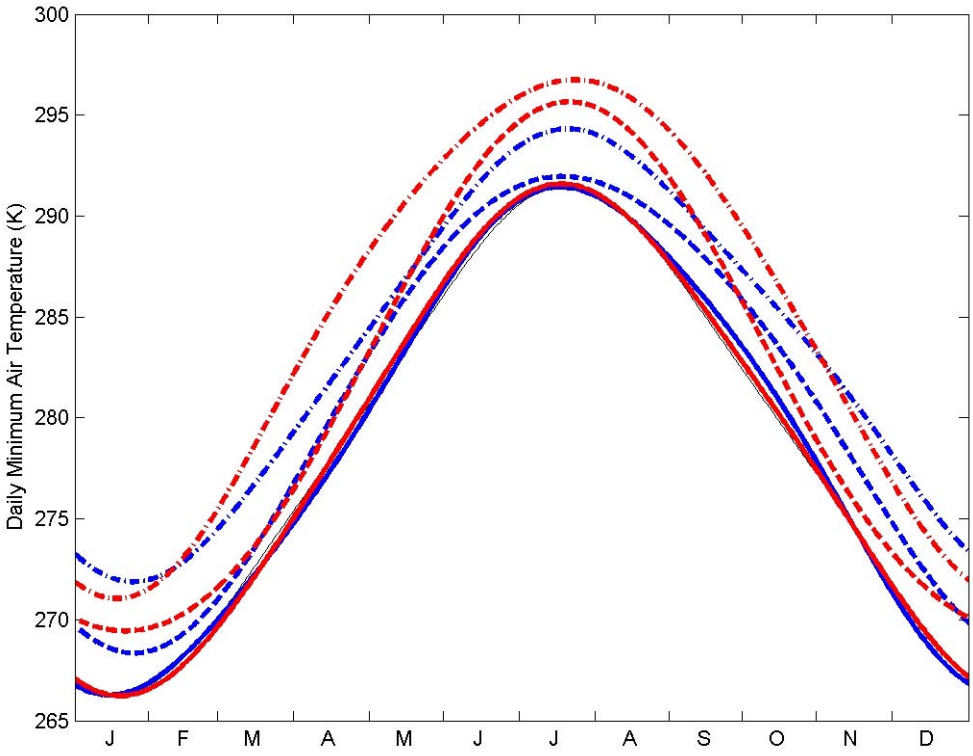

Figure 4. An example of the MOS-based seasonal cycle downscaling applied to minimum daily surface air temperature (T_{min}) at a central station (Berne, IN). The figure shows the observed (1990-2001) seasonal cycle (–), as well as the predicted seasonal cycles from CGCM2 (red lines) and HadCM3 (blue lines) for the periods 1990-2001 (–), 2040-2049 (––), and 2080-2089 (–·).

5.1.1 Projected changes in daily T_{max} and T_{min}: 2040-2049

The projected changes in daily T_{max} and T_{min} for 2040-2049 based on downscaling of the seasonal cycle differ substantially between the two GCMs (Table 4, Figure 5). These numbers indicate that, in accord with expectations: (1) changes in daily T_{min} are generally (excepting HadCM3 during autumn) greater than changes in daily T_{max} resulting in a decrease in diurnal temperature range, and (2) large differences exist in the projections of daily T_{max} and daily T_{min} by the two GCMs, primarily during spring (T_{max}) and summer (T_{max} and T_{min}), with much large regionally averaged changes derived from

HadCM3 output. Analysis of individual station results indicates that the differences in the temperature scenarios result largely from differences in GCM simulated sea-level pressure at the grid point locations used to downscale the seasonal cycles. Specifically, HadCM3 projects increases in SLP at these grid points throughout the year, while CGCM2 projects decreases in SLP during the summer.

Downscaled results from both GCMs indicate the largest increase in T_{max} during autumn and at stations in the northern part of the study region. Downscaled CGCM2 results indicate increases in T_{max} for all stations during each season, with the exception of a few stations in the southern part of the study region which show a slight cooling during summer. In HadCM3 decreases in T_{max} are projected for a large number of stations during spring, with decreases approaching 4 K at extreme southwest stations. These stations also exhibit slight decreases in T_{max} during winter and summer. At every station, CGCM2 projects much larger increases in T_{max} than HadCM3 during spring (differences as large as 4 K for stations in the extreme southwest), while HadCM3 projects larger increases in T_{max} than CGCM2 during summer.

Projections from both GCMs exhibit increases in T_{min} at all stations during each season, although during summer, minimum daily temperatures downscaled from HadCM3 exhibit changes in T_{min} which are an average (over the 84 stations) of 2.91 K larger than those projected by CGCM2. During the other seasons, differences between the projections are less than 1 K, with changes ranging from around 1-5 K, 1-3 K, and 1-5 K during winter, spring, and autumn, respectively.

Table 4. Projected seasonal changes in T_{max} and T_{min} (K) based on seasonal cycle downscaling averaged over all 84 stations within the study area.

	DJF		MAM		JJA		SON	
	HadCM3	CGCM2	HadCM3	CGCM2	HadCM3	CGCM2	HadCM3	CGCM2
T_{max} 2040-2049	2.19	1.86	0.42	1.66	2.17	0.36	2.67	2.07
2080-2089	5.50	3.93	6.43	2.95	5.90	1.60	7.97	3.19
T_{min} 2040-2049	2.80	2.19	2.12	1.98	4.02	1.10	2.50	2.63
2080-2089	5.12	5.36	6.95	4.12	5.76	2.83	6.22	4.65

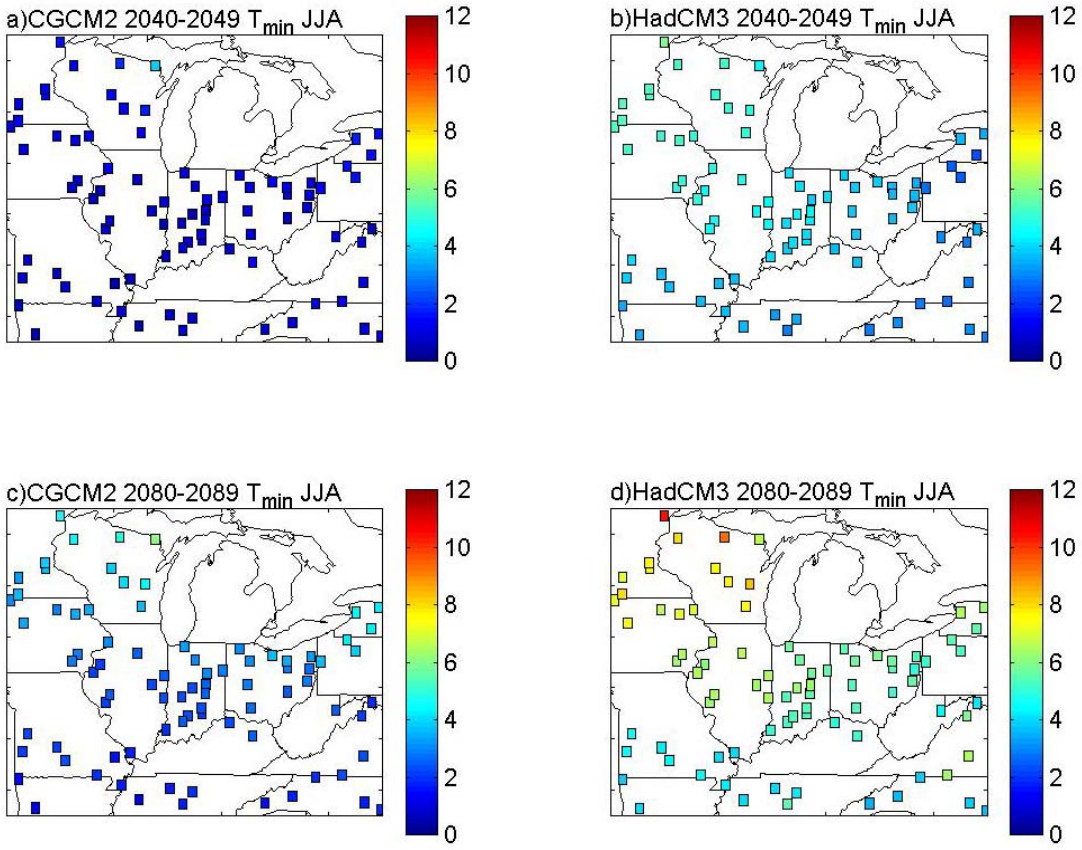

Figure 5. Projected changes in daily T_{min} (K) during summer relative to 1990-2001: a) CGCM2 2040-2049, b) HadCM3 2040-2049, c) CGCM2 2080-2089, d) HadCM3 2080-2089.

5.1.2 Projected changes in the daily T_{max} and T_{min}: 2080-2089

As expected, the projected changes in daily T_{max} and T_{min} for 2080-2089 are larger than those for 2040-2049 (Table 4, Figure 5). There are several important distinctions between temperature changes in the early and late 21st century. First, the differential warming in T_{max} and T_{min} projected by both GCMs for 2040-2049 is not mimicked by HadCM3 during the period 2080-2089. Second, relative to 1990-2001, the temperature change is positive and of greater magnitude in all seasons, as opposed to the earlier period in which HadCM3 projected decreases in spring T_{max} at some stations, and CGCM2 projected only slight increases in T_{max} on average during summer. This accelerated warming may be due to the increased rate of CO_2 accumulation in the latter half of the 21st century as specified by the SRES A2 scenario.

The seasonal characteristics of projected changes in T_{max} are similar to those for the earlier period for CGCM2, with the strongest warming occurring at northern stations during autumn, and the weakest warming occurring at southern stations during summer. As in the analysis for 2040-2049, increases in T_{max} downscaled from HadCM3 for 2080-2089 are strongest during the autumn. However, unlike the earlier period, large increases also occur during spring. With the exception of a few stations in the southwest part of the study area during winter and spring, station T_{max} values downscaled from HadCM3 are greater than those from CGCM2 in all seasons, and exceed 9 K for multiple stations in the northern part of the study area during spring and autumn. This result is particularly interesting given the global 21st century temperature projections from these models as reported by the IPCC TAR (IPCC 2001) and the magnitude of warming in the direct GCM output, which is generally larger in CGCM2 (see Figure 4).

Projections of 2080-2089 T_{min} from both models indicate warming at all stations during each season. However, with the exception of winter and a few stations in the other seasons, increases (relative to 1990-2001) in T_{min} values downscaled from HadCM3 are greater than those from CGCM2. While the differences between models are smaller on average than those for T_{max}, they approach 3 K during spring and summer. Both models project the largest warming at northern stations during the autumn and winter, with HadCM3 also indicating strong increases in T_{min} during the summer.

5.1.3 Summary

The downscaled seasonal cycles of T_{max} and T_{min} from both GCMs provide evidence of strong warming at most stations and continued decreases in diurnal temperature range, consistent with recent observations (Easterling et al. 1997) and the increases in greenhouse gases on which the model projections are based. However, the results from the models are substantially different on a seasonal basis. For example, during spring, downscaled HadCM3 T_{min} values for 2040-2049 at a large number of stations show a decrease or weak increase relative to 1990-2001, while those from CGCM2 indicate an increase of several degrees. Similarly, summer increases in T_{max} and T_{min} downscaled from CGCM2 are of smaller magnitude relative to those downscaled from HadCM3. For 2080-2089, results downscaled from both GCMs indicate an accelerated warming which is larger in HadCM3 during all seasons except winter. Additionally, while temperatures downscaled from CGCM2 continue to indicate differential warming between T_{max} and T_{min}, downscaled HadCM3 projections indicate similar increases in the two variables.

The spatial consistency of the results presented in the previous sections lends extra credence to the downscaled results, since the models used did not have a spatial component (i.e., the downscaling was performed independently for each station). However, as shown in Figure 5, there are differences in the downscaled projections for nearby stations which lie within the same GCM grid cell (often > 1K). Because these locations would be assigned a single value if raw GCM output were used, this finding offers further justification for downscaling to generate an appropriate level of spatial variability in the surface parameters at the sub-GCM grid box level.

5.2 Daily sequences of T_{max}, T_{min}, and precipitation

While the analysis of changes in the seasonal cycles of daily T_{max} and T_{min} presented in Section 5.1 provides information about temperature change, the construction of daily weather sequences is required for many climate change impact analysis. Additionally, regression-based approaches are not applicable to precipitation downscaling because precipitation only assumes non-negative values and exhibits a highly skewed distribution with a large number of days on which precipitation does not occur. Therefore, two additional downscaling techniques, the analog method and a stochastic weather generator, were used to derive daily sequences of T_{max}, T_{min} and precipitation. The T_{max} and T_{min} projections from these additional downscaling models are based on the previously discussed seasonal cycles, but also on the large scale weather state (see Figure 6). This approach allows examination of (1) differences in the projections due to changes in large scale parameters which largely determine day-to-day perturbations in the surface variables and (2) characteristics which may be important for

climate change impacts studies, such as the daily evolution of the surface variables (e.g., precipitation transition probabilities, wet and dry spell lengths, probabilities of extreme events, etc.).

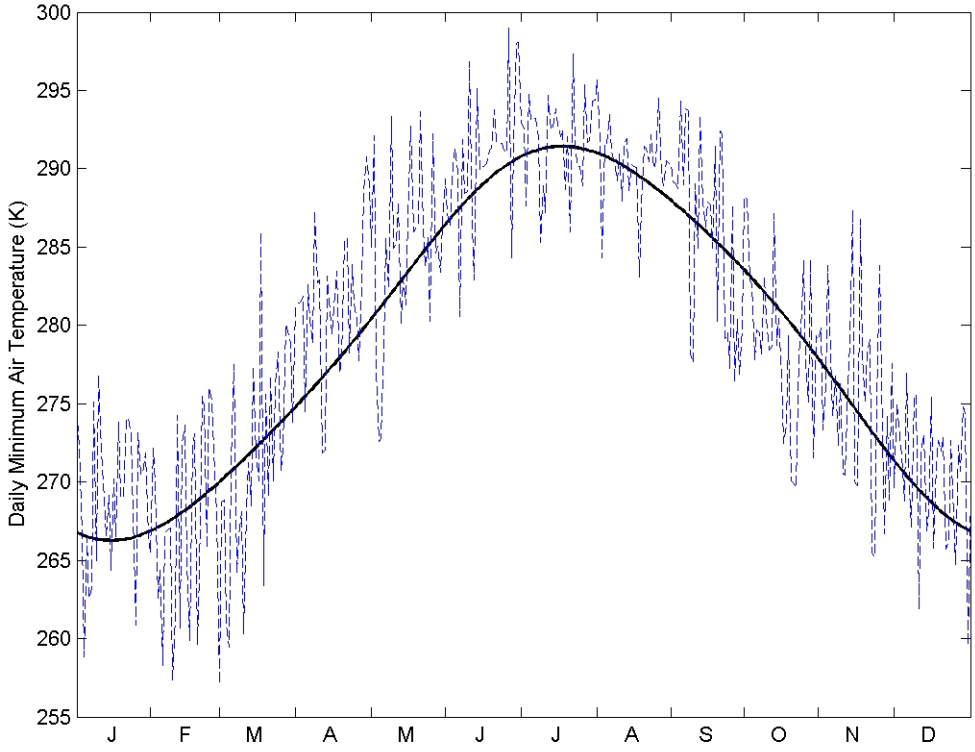

Figure 6. An example of the two-stage downscaling for T_{min} at a central station (Berne, IN) for 1990. The solid black line depicts the seasonal cycle downscaled from grid point CGCM2 predictors. The blue line depicts the residuals from the seasonal cycle resulting from application of the analog method to CGCM2 synoptic-scale simulations.

5.2.1 The analog method

5.2.1.1 Temperature scenarios

Like the MOS-based seasonal cycle downscaling presented in Section 5.1, the analog method was first applied to the period in which the observations and GCM simulations overlap (1990-2001) and then to the transient GCM simulations. For the former period, seasonal means of daily T_{max} and T_{min} downscaled with the analog method are generally within 1 K of the means derived from the downscaled seasonal cycles for

both GCMs. During spring (and winter for northern stations from CGCM2), mean T_{max} and T_{min} values at most stations are warmer (by 0.5 to 1.5 K) than the downscaled seasonal cycles for both GCMs. This pattern is also found in the transient GCM scenarios and suggests that the warming (relative to the seasonal cycle downscaling) in spring may be partially related to the simulation of synoptic phenomena in the GCMs. For example, HadCM3 has been shown to overestimate the frequency of a synoptic map pattern associated with anticyclonic features to the south of the study region, which could result in greater warm air advection into the study region (Schoof and Pryor 2004).

As mentioned previously, one of the advantages of the analog method relative to the MOS-based seasonal cycle downscaling is that it allows examination of the probability distributions of daily T_{max} and T_{min}. Examination of these distributions indicates that the increases in downscaled winter T_{max} and T_{min} from CGCM2 occur over the entire probability distribution, while spring and autumn T_{max} and T_{min} increases are much stronger in the lower tail of the probability distributions. The changes projected by downscaling T_{max} from HadCM3 (including the mid-21st century cooling at some stations during spring) also occur in the lower part of the distribution. However, for T_{min}, the increases occur over the entire distribution (winter and summer) or in the upper part of the distribution (spring and autumn). For both GCMs and both surface temperature variables, changes in variability are also indicated by the downscaling results. Specifically, T_{max} downscaled from CGCM2 is less variable (decreases in standard deviations during spring and autumn as large as 1 K by 2040-2049 and 2 K by 2080-2089). Winter and spring T_{min} standard deviations downscaled from CGCM2 are also projected to decrease by 1-2 K by 2080-2089. Projections from HadCM3 indicate less

variable summer T_{max} and T_{min} values, and increases in spring T_{max} and autumn T_{min} values.

These results have profound implications for projections of human heat-related mortality associated with heat waves. Both Schar et al. (2004) and Meehl and Tebaldi (2004) have recently published research that reports a potential increase in future heat waves due to increases in the variance of surface temperatures over Europe. While the results presented in this study do not indicate a widening of the temperature probability distributions, the analog-downscaled results do indicate an increase in the number of heat waves (defined as a three-day or longer period in which both T_{max} and T_{min} are greater than their respective 95th percentiles). As shown in Figure 7, temperatures downscaled from CGCM2 produce an increase of 1-4 heat waves per year by 2080-2089 (relative to 1990-2001). This is a very large change relative to the 1990-2001 period, in which most stations averaged around one heat wave per year. HadCM3 downscaled temperatures exhibit smaller increases in the number of annual heat waves.

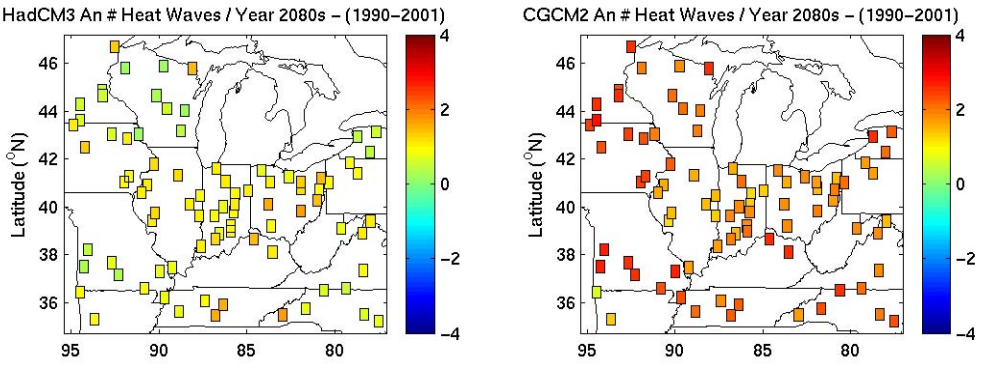

Figure 7. Increases in the number of annual heat waves as indicated by analog downscaling of HadCM3 (left) and CGCM2 (right).

5.2.1.2 Precipitation scenarios

Precipitation values downscaled using the analog method are in general agreement with observations over the period of overlapping data (1990-2001), although downscaled precipitation is consistently higher than observed for stations in the northwest part of the study region during winter for both GCMs, consistent with findings reported for a large suite of GCMs in the IPCC TAR (IPCC 2001). During the other seasons, downscaled mean wet-day precipitation amounts are within 25% of observed values. Downscaling from both GCMs successfully reproduces the wet-day probability within 10% of that observed during each season at each station. Further, results downscaled from CGCM2 successfully reproduce the observed length of wet and dry spells at each station during each season (all differences between downscaled and observed mean wet and dry spell lengths are less than 1 day), while downscaled HadCM3 results underestimate the length of dry spells by 1-2 days for most stations during autumn. For example, at Berne, IN, winter and spring dry spell lengths are well matched in downscaled results from both models. However, during both summer and autumn, mean dry spell lengths downscaled from HadCM3 are considerably shorter than those observed (3.1 and 3.4 days (observed) vs. 2.6 and 2.5 days (downscaled) for summer and autumn, respectively).

When applied to future periods from the GCMs, the results of downscaled precipitation are highly variable. The mean wet-day precipitation amounts exhibit no consistent spatial pattern, although downscaled seasonal precipitation estimates for 2040-2049 differ from 1990-2001 values by as much as 40%, with similar numbers of stations exhibiting increases and decreases (Figure 8). Averaged over all stations, 2040-2049

CGCM2-downscaled precipitation changes are largest during winter, increasing by about 6%, with smaller increases of 1-2% during summer and autumn. During spring, the average change in precipitation is a decrease of around 1%. In contrast, the results downscaled from HadCM3 for 2040-2049 show the largest increase in precipitation during spring, approximately 3.5% averaged over all stations. During summer, HadCM3-downscaled precipitation is about 2.5% lower, on average, than that from 1990-2001. Downscaled results from both GCMs are less spatially coherent that those for temperature, with both increases and decreases occurring at some stations in all seasons. The other parameters computed for evaluation of downscaled daily precipitation exhibit small and spatially variable changes.

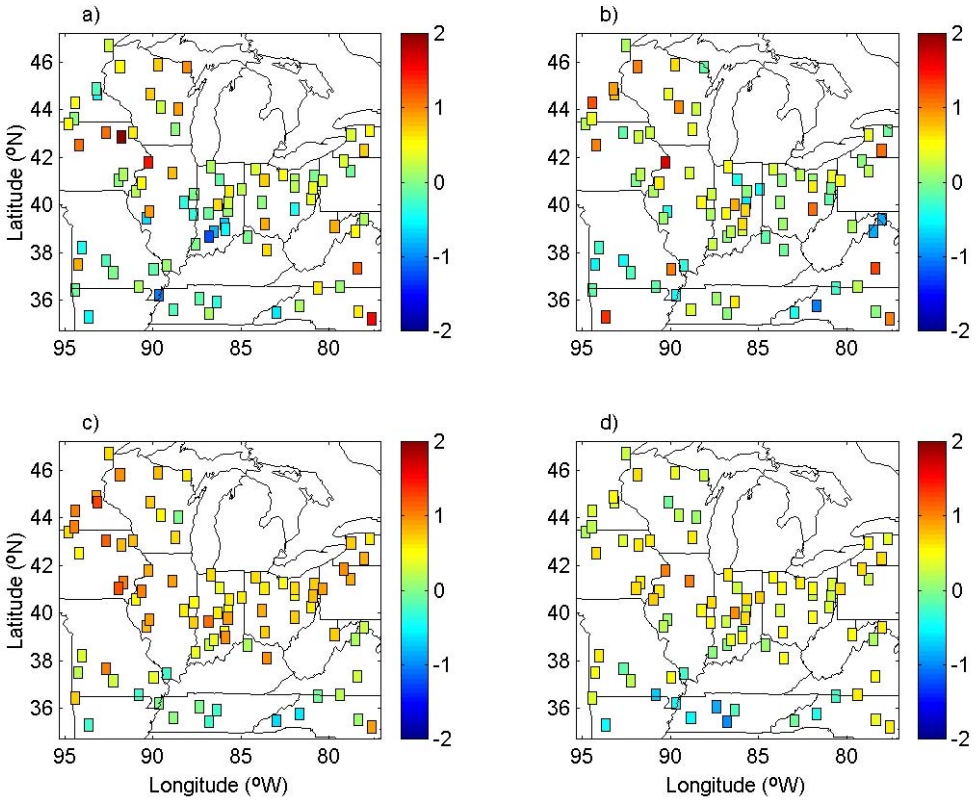

Figure 8. Changes in wet-day precipitation amount (mm/day) for 2049-2049 relative to 1990-2001: a) HadCM3 analog, b) CGCM2 analog, c) HadCM3 stochastic weather generator, and d) CGCM2 stochastic weather generator.

For the late 21st century (2080-2089), the analog-based downscaled changes in precipitation are more positive and more spatially coherent, although the downscaled results from the two GCMs differ, particularly during summer. The CGCM2 downscaled precipitation, averaged over all stations, is almost 9.5% higher than that for the reference period, 1990-2001, during winter and almost 3% lower during summer. During spring and autumn the CGCM2-downscaled precipitation estimates for 2080-2089, averaged over all stations, are within 1% of those for 1990-2001. For the same period (2080-2089), the area-averaged changed in precipitation downscaled from HadCM3 is positive in all seasons, with the largest increases during winter (5.5%) and summer (>6%). These results are consistent with historical analyses, indicating changes in the magnitude of wet-day events, rather than changes in the number of wet days (Folland et al. 2001).

5.2.2 Stochastic weather generator

5.2.2.1 Temperature scenarios

The stochastic weather generator was also first applied to the period in which the observations and GCM simulations overlap (1990-2001) and then to the transient GCM simulations. Seasonal means of daily T_{max} and T_{min} downscaled with the weather generator also show agreement with those derived from the downscaled seasonal cycles generally within 1 K, although there are larger differences in some regions during some seasons. Specifically, for both the reference periods and the transient simulations, CGCM2 temperatures downscaled using the weather generator are colder (>1 K) than those resulting from downscaling the seasonal cycle during and warmer (>1 K) during spring. A similar winter result is found for temperatures downscaled from HadCM3.

In accord with the results from analog downscaling, the increases in downscaled winter T_{max} and T_{min} from CGCM2 occur over the entire probability distribution while those during autumn occur in the lower tail of the distribution. However, spring T_{max} and T_{min} increases are much larger for the upper percentiles, in contrast to the results attained with the analog method. The changes projected by downscaling T_{max} from HadCM3 with the weather generator (excepting the mid-21st century cooling during spring) are also in accord with the analog downscaled results, and occur evenly over the T_{max} distribution. For T_{min}, the increases occur over the entire distribution (winter and summer) or in the upper part of the distribution (spring and autumn). Changes in the standard deviations of T_{max} and T_{min} downscaled from CGCM2 differ from the results obtained with the analog method, and indicate a decreases in standard deviations during autumn as large as 1-2 K by 2080-2089. When downscaled from HadCM3 using the weather generator, T_{max} is more variable during summer while T_{min} is more variable during autumn. The latter result in is in accord with the analog downscaled results.

Relative to the analog method, temperatures downscaled using the stochastic weather generator result in much smaller changes in the number of heat waves. While some stations exhibit decreases in the number of heat waves, most stations indicate an increase of 1-3 heat waves per decade, not per year, as indicated by the analog method.

5.2.2.2 Precipitation scenarios

As with the analog downscaled precipitation, precipitation amounts derived using the stochastic weather generator are consistently higher than observed for stations in the northwest part of the study region during winter for both GCMs during 1990-2001. The

magnitude and spatial extent of these differences is much smaller for the data downscaled with the stochastic weather generator than with the analog method. During the other seasons, downscaled mean wet-day precipitation amounts are within 20% of observed values. SWG downscaling from both GCMs successfully reproduces the wet-day probability within 10% during each season at each station. Mean wet and dry spell lengths are well simulated by both GCMs when the stochastic weather generator is used for downscaling. Precipitation results downscaled from the stochastic weather generator exhibit better agreement with observations with respect to extreme precipitation, especially for stations in the northwest part of the study area, for which analog downscaled 90th precipitation percentiles differ by more than 100% of the observed values and those downscaled by the stochastic weather generator are within 20% of observed values.

When applied to future periods from the GCMs, the precipitation changes downscaled from the weather generator are of similar magnitude to those projected with the analog method and more spatially consistent. Averaged over all stations, 2040-2049 CGCM2-downscaled precipitation changes are largest during winter, decreasing by about 6% and mostly confined to large precipitation events, with smaller decreases of 1-2% during summer and autumn. During spring, the average change in precipitation is an increase of around 1%. The results downscaled from HadCM3 for 2040-2049 exhibit only small changes of 1-2%.

For the late 21st century (2080-2089), the downscaled changes in precipitation are small (changing by only 1-2%) for both GCMs during spring, summer, and autumn. During winter, however, downscaled CGCM2 precipitation is an average (over the 84

stations) of 12% lower than that during the reference period (1990-2001). Results obtained by downscaling from HadCM3 with the weather generator indicate smaller decreases in precipitation during winter (around 3% on average).

6 Summary and Conclusions

The focus of this study was the generation of future temperature and precipitation scenarios for the Midwestern region of the USA. The scenarios were constructed by evaluating a range of free atmosphere predictors, assessing the validity of their simulation by two GCMs, and developing, applying and evaluating empirical downscaling methods.

In each season, the thickness of the 850-500 hPa layer exhibited the strongest correlation with daily T_{max} and T_{min}, while 500-hPa geopotential height, 850-hPa specific humidity, and sea level pressure also exhibited strong correlations. For precipitation, correlation coefficients were lower and the total geostrophic shear vorticity was identified as the single best variable. Specific humidity at 850-hPa was also found to exhibit strong correlation with precipitation at most stations. For all three surface variables, the maximum correlations were generally found at grid points located remotely from the station location. Evaluation of the predictor variables as simulated by the GCMs showed that for many of the predictor variables, grid point means and variances are not in agreement with observations and hence, downscaling models were constructed using a MOS-based approach coupled with two perfect prognosis (PP) approaches that can incorporate GCM information on larger scales: an analog method and a stochastic weather generator.

Downscaled temperature results from both GCMs and all three downscaling methods indicated a strong warming trend at most stations consistent with that observed over the period of observation and consistent with the increases in greenhouse gases. However, the results from the models and downscaling approaches were substantially different on a seasonal basis. For 2040-2049, winter and autumn changes in T_{max} and T_{min} were similar for downscaled projections from the two GCMs and indicate a differential warming in T_{max} and T_{min}, consistent with continued decreases in diurnal temperature range. During spring, however, downscaled HadCM3 T_{min} values for the mid-21st century showed a decrease at some stations relative to 1990-2001, while those from CGCM2 consistently indicated temperature increases. Similarly, summer increases in T_{max} and T_{min} were weak in CGCM2 relative to HadCM3. For 2080-2089, results from both GCMs and each downscaling method indicated an accelerated warming which is larger in HadCM3 during all seasons except winter, although in some seasons the downscaling methods attributed this warming to different parts of the probability distributions. Additionally, while CGCM2 indicated a continued differential warming between T_{max} and T_{min}, downscaled HadCM3 projections indicated similar increases in the two variables.

Downscaled precipitation results were much less consistent between models and downscaling methods. The large differences between downscaled precipitation series from the GCMs and downscaling methods reduces the confidence with which these prognoses can be viewed, but are consistent with the IPCC TAR which found inconsistent sign in precipitation change from several coupled models for central North America (Giorgi et al. 2001).

The results presented here have major implications for those interested in climate change impacts. The downscaling methods presented here were capable of reproducing the major features of the observed climate over the period in which stations observations and GCM simulations co-exist (1990-2001). However, even when downscaling was conducted using GCMs with identical forcing scenarios and similar predictor variables and methodologies, the resulting temperature and precipitation scenarios were drastically different. Future work will address these issues by focusing on identifying physical causes for these differences and expanding the work presented here to included systematic evaluation of other GCMs and downscaling methods.

References

Bardossy, A., Stehlik, J. and Caspary, H.-J. (2002). Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules. *Climate Research* **23**: 11-22.

Bellone, E., Hughes, J. P. and Guttorp, P. (2000). A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Climate Research* **15**: 1-12.

Brinkmann, W. A. R. (2002). Local versus remote grid points in climate downscaling. *Climate Research* **21**: 27-42.

Charles, S. P., Bates, B. C., Whetton, P. H. and Hughes, J. P. (1999). Validation of downscaling models for changed climate conditions: case study of southwestern Australia. *Climate Research* **12**: 1-14.

Cubasch, U., Meehl, G. A., Boer, G. J., Stouffer, R. J., Dix, M., Noda, A., Senior, C. A., Raper, S. and Yap, K. S. (2001). Projections of Climate Change. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggset al. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 881.

Cubasch, U., von Storch, H., Waszkewitz, J. and Zorita, E. (1996). Estimates of climate change in Southern Europe derived from dynamical climate model output. *Climate Research* **7**: 129-149.

Easterling, D. R., Horton B, Jones P.D., Peterson T.C., Karl T.R., Parker D.E., Salinger M.J., Razuvayev V., Plummer N., Jamason P., Folland C.K. (1997). Maximum and minimum temperature trends for the globe. Science **277**: 364-367.

Easterling, D. R., Karl, T. R., Lawrimore, J. H. and Del Greco, S. A. (1999). United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data for 1871-1997. Oak Ridge, Tennessee, Oak Ridge National Laboratory**:** 84.

Flato, G., Boer, G. J., Lee, W. G., McFarlane, N. A., Ramsden, D., Reader, M. C. and Weaver, A. J. (2000). The Canadian Centre for Climate Modelling and Analysis global coupled model and its climate. *Climate Dynamics* **16**: 451-467.

Flato, G. M. and Boer, G. J. (2001). Warming asymmetry in climate change simulations. *Geophysical Research Letters* **28**: 195-198.

Folland, C. K., Karl, T. R., Christy, J. R., Clarke, R. A., Gruza, G. V., Jouzel, J., Mann, M. E., Oerlemans, J., Salinger, J. and Wang, S.-W. (2001). Observed Climate Variability and Change. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report on the Intergovernmental Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggset al. Cambridge, United Kingdom, and New York, NY, USA, Cambridge University Press**:** 881pp.

Foufoula-Georgiou, E. and Lettenmaier, D. P. (1987). A Markov renewal model for rainfall occurrences. *Water Resources Research* **23**: 875-884.

Gibson, J. K., Kallberg, P., Uppala, S., Noumura, A., Hernandez, A. and Serrano, E. (1997). ERA Description. ECMWF Reanalysis Project Report Series. Reading, UK.

Giorgi, F., Hewitson, B. C., Christensen, J., Hulme, M., von Storch, H., Whetton, P., Jones, R., Mearns, L. and Fu, C. (2001). Regional Climate Information - Evaluation and Projections. Climate Change 2001: The Scientific Basis. Contribution of Working Group i to the Third Assessment Report of the Intergovernmentla Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer and P. J. van der Linden. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 881.

Glahn, H. R. and Lowry, D. A. (1972). The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology* **11**: 1203-1211.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B. and Wood, R. A. (2000). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **16**: 147-168.

Greene, W. H. (2000). Econometric Analysis. Upper Saddle River, N.J., Prentice Hall.

Hewitson, B. C. and Crane, R. G. (1996). Climate downscaling: techniques and application. *Climate Research* **7**: 85-95.

Huth, R. (2002). Statistical downscaling of daily temperature in Central Europe. *Journal of Climate* **15**: 1731-1742.

IPCC (2000). Special Report on Emissions Scenarios. Cambridge, Cambridge University Press**:** 612pp.

Jenkinson, A. F. and Collison, P. (1977). An initial climatology of gales over the North Sea. London, Meteorological Office**:** 18.

Jones, P. D. (1994). Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *Journal of Climate* **7**: 1794-1802.

Jones, P. D., Hulme, M. and Briffa, K. R. (1993). A comparison of Lamb circulation types with an objective classification scheme. *International Journal of Climatology* **13**: 655-663.

Jones, P. D., Osborn, T. J. and Briffa, K. R. (1997). Estimating sampling erros in large-scale temperature averages. *Journal of Climate* **10**: 2548-2568.

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, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437-470.

Karl, T. R., Williams Jr., C. N. and Quinlan, F. T. (1990). United States Historical Climatology Network (HCN) Serial Temperature and Precipitation Data. Oak Ridge, Tennessee, Oak Ridge National Laboratory.

Kattenberg, A., Giorgi, F., Grassl, H., Meehl, G. A., Mitchell, J. F. B., Stouffer, R. J., Tokioka, T., Weaver, A. J. and Wigley, T. M. L. (1996). Climate Models - Projections of Future Climate. Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the IPCC. J. T. Houghton, L. G. Meira Filho, B. A. Callanderet al. Cambridge, Cambridge University Press**:** 285-357.

Kim, J. W., Chang, J. T., Baker, N. L., Wilks, D. S. and Gates, W. L. (1984). The statistical problem of climate inversion: Determination of the relationship between local and large-scale climate. *Monthly Weather Review* **112**: 2069-2077.

Klein, W. H. (1982). Statistical weather forecasting on different time scales. *Bulletin on the American Meteorological Society* **63**.

Klein, W. H., Lewis, B. M. and Enger, I. (1959). Objective prediction of five-day mean temperatures during winter. *Journal of Meteorology* **16**: 672-682.

Kruizinga, S. and Murphy, A. H. (1983). Use of an analogue procedure to formulate objective probabilistic temperature forecasts in The Netherlands. *Monthly Weather Review* **111**: 2244-2254.

Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences* **26**: 636-646.

Matalas, N. C. (1967). Mathematical assessment of synthetic hydrology. *Water Resources Research* **3**: 937-945.

McAveney, B. J., Covey, C., Joussaume, S., Kattsov, V., Kitoh, A., Ogana, W., Pitman, A. J., Weaver, A. J., Wood, R. A. and Zhao, Z.-C. (2001). Model Evaluation. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggset al. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 881pp.

Meehl, G. A. and Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Nature* **305**: 994-997.

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P. and Parker, D. (1998). Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology* **18**: 1493-1517.

Pope, V. D., Gallani, M. L., Rowntree, P. R. and Stratton, R. A. (2000). The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics* **16**: 123-146.

Richardson, C. W. and Wright, D. A. (1984). WGEN: A model for generating daily weather variables, USDA-ARS**:** 83.

Sailor, D. J. and Li, X. (1999). A semiempirical downscaling approach for predicting regional temperature impacts associated with climate change. *Journal of Climate* **12**: 103-114.

Schar, C., Vidale, P. L., Luthi, D., Frei, C., Haberli, C., Liniger, M. A. and Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature* **427**: 332-336.

Schnur, R. and Lettenmaier, D. P. (1998). A case study of statistical downscaling in Australia using weather classification by recursive partitioning. *Journal of Hydrology* **212-213**: 362-379.

Schoof, J. T. and Pryor, S. C. (2001). Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. *International Journal of Climatology* **21**: 773-790.

Schoof, J. T. and Pryor, S. C. (2004). An evaluation of two GCMs: Simulations of North American teleconnections and synoptic phenomena. (in preparation for submission to *Climate Research*).

Schoof, J. T. and Robeson, S. M. (2003). Seasonal and spatial variations of cross-correlation matrices used by stochastic weather generators. *Climate Research* **24**: 95-102.

Semenov, M. A. and Barrow, E. M. (1997). Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change* **35**: 397-414.

Sousounis, P. J. and Albercook, G. M. (2000). The Great Lakes Region: Past, Present, and Future. Preparing For A Changing Climate. The Potential Consequences of Climate Variability and Change. P. J. Sousounis and J. M. Bisanz.

van den Dool, H. M. (1994). Searching for analogues, how long must we wait? *Tellus* **46A**: 314-324.

Wilby, R. L. and Wigley, T. M. L. (1997). Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography* **21**: 530-548.

Wilby, R. L. and Wigley, T. M. L. (2000). Precipitation predictors for downscaling: observed and general circulation model relationships. *International Journal of Climatology* **20**: 641-661.

Wilks, D. S. (1992). Adapting stochastic weather generation algorithms for climate change studies. *Climatic Change* **22**: 67-84.

Wilks, D. S. (1999). Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology* **93**: 153-169.

Wilks, D. S. (1999). Multisite downscaling of daily precipitation with a stochastic weather generator. *Climate Research* **11**: 125-136.

Wilks, D. S. and Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* **23**: 329-357.

Zorita, E., Hughes, J. P., Lettenmaier, D. P. and von Storch, H. (1995). Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *Journal of Climate* **8**: 1023-1042.

Zorita, E. and von Storch, H. (1999). The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate* **12**: 2474-2489.

CHAPTER 9: SUMMARY AND CONCLUSIONS

9.1 Summary

The focus of this dissertation has been the development and evaluation of a methodology for generating regional climate change scenarios for the Midwestern region of the USA using empirical methods. In the following sections, I restate my objectives and summarize my results as they pertain to those objectives.

9.1.1 To evaluate the homogenized, gridded data sets (i.e., reanalysis products) commonly using in empirical downscaling)

In Chapter 4, the ability of the NCEP/NCAR reanalysis to represent the synoptic climate of the Midwestern USA relative to radiosonde data was examined. A multi-method approach was used, which allowed both examination of independent classifications of the input fields and an examination of the improvement achieved by projecting the NCEP/NCAR data onto the classification derived using the radiosonde data. Using automated synoptic classifications based on rotated principal component analysis (PCA) and a two-step clustering algorithm, classifications were developed and 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. While the classifications were 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 analyze whether the differences in the classifications were 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 categorized the NCEP/NCAR reanalysis 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 and substantially improved the agreement between the classifications (in terms of both interpretability and mutual class occupancy), although the solutions still exhibit considerable discrepancies.

These results suggest that the NCEP/NCAR 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 consideration of the data set chosen to resolve synoptic scale phenomena and climate. Specific to the subsequent analyses presented in this dissertation, the analysis provides a context for comparisons of synoptic classifications derived from reanalysis data and output from GCMs.

9.1.2 To use the historical climate record to identify physically meaningful large-scale variables accounting for a large proportion of the variance of surface climate parameters.

In Chapter 5, a large suite of potential atmospheric predictors from the NCEP/NCAR and ECMWF reanalyses (500 and 850-hPa geopotential heights, 850-hPa relative and specific humidity, sea-level pressure) and several derived variables (850-500 hPa layer thickness and geostrophic flow and vorticity variables) were examined in terms

of their relationships with daily maximum and minimum surface air temperature (T_{max} and T_{min}) and total daily precipitation at 84 stations in Midwestern USA. The chosen predictor variables were found to be closely related to the surface variables, although the relationships vary substantially both geographically and seasonally. In most cases, the correlations could be explained in terms of the physical causes which are related to either thermal or moisture advection or enhanced vertical motion.

The variables exhibiting the strongest relationships with maximum and minimum daily surface air temperature (in all seasons) are 850-500 hPa thickness, 500-hPa geopotential height, 850-hPa specific humidity, sea-level pressure, the southerly and westerly components of the geostrophic flow, and 850-hPa geopotential height. Correlations were generally found to be higher and significant over larger areas during spring and autumn, consistent with higher spatial autocorrelation of the surface variables during those seasons. Correlations between the predictor variables and station precipitation were weaker than those for the temperature variables, but several predictor variables exhibit strong relationships with station precipitation. At most stations, the most highly correlated predictors were the geostrophic vorticity variables and 850-hPa specific humidity. Several other variables exhibited high correlations with precipitation at some stations during some seasons. These results did not offer support for use of lagged and differenced predictors in the downscaling models, but did provide support for the use of remote grid points in downscaling.

9.1.3 To evaluate multiple GCM simulations and identify well-simulated variables for use in generating regional climate prognoses

Chapters 6 and 7 focused on the evaluation of GCM simulations. Specifically, two GCMs are evaluated in terms of (1) mean fields during the period in which the observations and GCM simulations overlap (1990-2001), (2) individual grid point means and variances and (3) representation of synoptic phenomena, such as the Pacific/North American teleconnection (PNA) and North Atlantic Oscillation (NAO) and their links to atmospheric circulation in the study region.

The mean fields for the period of overlapping observations and GCM simulations (1990-2001) were found to exhibit close agreement for 850- and 500-hPa geopotential heights, 850-500 hPa layer thickness, and 850-hPa specific humidity, and, to a lesser extent, the geostrophic flow variables. For sea-level pressure and 850-hPa relative humidity, correlations between observed and GCM simulated mean fields were lower, and the spatial variability in the mean pattern was substantially overestimated during some seasons. While the mean fields of the CGCM2 geostrophic vorticity variables exhibited low correlations relative to the other variables, the variability in the mean fields was found to be well simulated. The mean fields of HadCM3 simulated geostrophic vorticity variables exhibited substantially greater spatial variability than observed.

For most variables (excepting HadCM3 geostrophic vorticity variables) and grid points, GCM simulated means and variances were not found to be different from observations on annual time scales. However, as the temporal resolution increases, more GCM grid point means and variances were found to be statistically different from those observed. In particular, CGCM2 500-hPa geopotential height means were found to be

significantly different for many grid points during the summer months, while variances were found to be significantly different during the winter months. These results also impacted 850-500 hPa thickness, for which many grid point means and variances were found to be significantly different from those observed during the same time periods. For HadCM3, means and variances of 500-hPa geopotential heights were not found to be statistically different nearly as often as those for geopotential heights closer to the surface (at 850-hPa). Multiple grid points exhibited statistically different means for humidity variables in both GCMs during the summer. For both GCMs, sea-level grid point means and variances were statistically different from observations less often than the other variables, with the exception of HadCM3 during the summer months.

The analysis of simulations at the synoptic scale, presented in Chapter 7, indicated that circulation regimes at 500 hPa as simulated by HadCM3 and CGCM2 for 1900-2001 show a high degree of correspondence with those manifest in the NCEP/NCAR reanalysis in terms of frequencies, persistence and progression. However, HadCM3 was found to overestimate the frequency of the most prevalent map type during all seasons. Additional discrepancies between mean frequencies of the reanalysis-derived and HadCM3 classifications resulted from different classification of days with weak anticyclonic characteristics. CGCM2 generally simulates the mean frequencies of the reanalysis-derived map types better than HadCM3, a surprising result given the poor performance of CGCM2 500-hPa geopotential heights on short time scales. Some, but not all, of the shortcomings of the GCM-based classifications can be explained by differences in the observed and simulated relationships between the map types and two teleconnection indices. Differences between the mean frequencies of observed and GCM

map types are larger than changes in the map type frequencies as manifest in the historical observations or prognostic changes in the mean map type frequencies as manifest in the 2030-2041 GCM simulations.

These results of the GCM evaluations support earlier findings that GCMs are more reliable over multiple grid points and longer time scales, than at single grid points and short time scales. These findings, taken in concert with the predictor variable identification results suggest that the best downscaling models will result from either the MOS-based downscaling approach or by using GCM information on larger scales.

9.1.4 To apply, compare, and evaluate multiple empirical downscaling methods.

Chapter 8 of this dissertation focused on the development, evaluation and comparison of empirical downscaling methods. The findings presented in Chapters 5-7 placed limitations on the types of downscaling strategies that could be used for the GCMs and predictor variables in our study region. Therefore, a new method was devised, in which the magnitude of changes in the temperature variables was first determined by downscaling the seasonal cycles of maximum and minimum daily surface temperature using a MOS-type approach. Two additional techniques, an analog method and a stochastic weather generator, were then used to describe precipitation projections and deviations of the temperature variables from their respective seasonal cycles. These methods provide additional detail associated with projections at the daily time scale.

Downscaled temperature results from both GCMs and all three downscaling methods indicate strong warming at most stations consistent with that observed over the period of observation and consistent with the increases in greenhouse gases. However,

the results downscaled from the models are substantially different on a seasonal basis. For 2040-2049, winter and autumn changes in T_{max} and T_{min} are similar for downscaled projections from the two GCMs and indicate a differential warming in T_{max} and T_{min}, consistent with continued decreases in diurnal temperature range. During spring, however, downscaled HadCM3 T_{min} values for the mid-21st century show an overall decrease relative to 1990-2001, while those from CGCM2 indicate an overall increase. Similarly, summer increases in T_{max} and T_{min} are weak in CGCM2 relative to HadCM3. For 2080-2089, results from both GCMs and each downscaling method indicate an accelerated warming which is larger in HadCM3 during all seasons except winter, although in some seasons the downscaling methods attribute this warming to different parts of the probability distributions. Additionally, while CGCM2 continues to indicate differential warming between T_{max} and T_{min}, downscaled HadCM3 projections indicate similar increases in the two variables. As a consequence of the differences in temperature projections from the different GCMs and downscaling methods, impacts-related parameters, such as heat wave frequency, remain largely uncertain, with analog-downscaled temperature series indicating much more frequent future heat waves relative to the stochastic weather generator.

Downscaled precipitation results are much less consistent between models and downscaling methods. Precipitation scenarios produced with the stochastic weather generator generally indicate decreases in precipitation within the study area during the 21st century, with the greatest decrease during winter, although the magnitude of the decrease varies between GCMs, while analog methods generally indicate increases in precipitation during the 21st century with the largest increases during winter. These

inconsistencies in the sign and magnitude of precipitation change in the study area agree with the findings of the IPCC TAR (Giorgi et al. 2001).

9.2 Discussion

There are several caveats to the findings reported in this dissertation. First, there is no way to satisfactorily test the assumption that the relationships between predictors and predictands do not change over time or to determine whether or not the climate change signal is fully characterized by the chosen predictors. Additionally, the 12-year period of overlap between observed data and the transient GCM simulations does not represent a full climatology, but was used for evaluation of several GCM characteristics, and for formulation of the MOS-based regression equations.

Nevertheless, the findings presented in this dissertation have broad implications for the use of GCMs to generate regional climate change scenarios. The two GCMs used in this study employed the same greenhouse gas emission scenario and represent the state-of-the-art in current coupled models. Even when driven with nearly identical predictor ensembles, the resulting downscaled scenarios of the surface variables are substantially different. This finding suggests that empirical downscaling tools will improve considerably when GCMs exhibit better agreement with one another.

9.3 Future Work

The findings presented in this research provide several avenues for future research. Examples include the comparison of one- and two-stage downscaling methods and the application of the methods presented here to other models, variables, regions, and

impacts. Additionally, further analysis is needed in terms of predictor variable comparison and evaluation, such as examination of the geostrophic flow and vorticity variables at vertical levels away from the surface.

As advancements in technology result in improvements in multiple aspects of climate modeling, GCM simulations should improve with respect to both model resolution and model realism. A larger number of variables archived at modeling centers will result in a greater diversity of possible analyses, providing greater insight into potential regional climate change impacts. Until that time, the two-stage downscaling method presented here provides a means of examining potential climate change impacts without reliance on GCM simulations that do not exhibit agreement with observations.

References cited

Arnell, N. W., Hudson, D. A. and Jones, R. G. (2003). Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. *Journal of Geophysical Research* **108**: 4519.

Bardossy, A. and Plate, E. G. (1992). Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resources Research* **28**: 1247-1259.

Bardossy, A., Stehlik, J. and Caspary, H.-J. (2002). Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules. *Climate Research* **23**: 11-22.

Bellone, E., Hughes, J. P. and Guttorp, P. (2000). A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Climate Research* **15**: 1-12.

Blackmon, M. L., Lee, Y.-H., Wallace, J. M. and Hsu, H.-H. (1984). Time variation of 500-mb height fluctuations with long, intermediate, and short time scales as deduced from lag-correlation statistics. *Journal of Atmospheric Science* **41**: 961-979.

Blair, D. (1998). The Kirchhofer technique of synoptic typing revisited. *International Journal of Climatology* **18**: 1625-1635.

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

Brinkmann, W. A. R. (1999). Application of non-hierarchically clustered circulation components to surface weather conditions: Lake Superior Basin winter temperatures. *Theoretical and Applied Climatology* **63**.

Brinkmann, W. A. R. (1999). Within-type variability of 700 hPa winter circulation patterns over the Lake Superior basin. *International Journal of Climatology* **19**.

Brinkmann, W. A. R. (2000). Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation. *International Journal of Climatology* **20**.

Brinkmann, W. A. R. (2002). Local versus remote grid points in climate downscaling. *Climate Research* **21**: 27-42.

Buishand, T. A. and Brandsma, T. (1997). Comparison of circulation classification schemes for predicting temperature and precipitation in The Netherlands. *International Journal of Climatology* **17**: 875-889.

Calinski, R. B. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics* **3**: 1-27.

Catell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research* **1**: 245.

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

Cavazos, T. (1999). Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. *Journal of Climate* **12**: 1506-1523.

Cayan, D. R. and Roads, J. O. (1984). Local relationships between US west coast precipitation and monthly mean circulation patterns. *Monthly Weather Review* **112**: 1276-1282.

Charles, S. P., Bates, B. C., Whetton, P. H. and Hughes, J. P. (1999). Validation of downscaling models for changed climate conditions: case study of southwestern Australia. *Climate Research* **12**: 1-14.

Chaston, P. R. (1997). Weather Maps. Kearney, MO, Chaston Scientific.

Chervin, R. M. (1981). On the comparison of observed and GCM simulated climate ensembles. *Journal of the Atmospheric Sciences* **38**: 885-901.

Coleman, J. S. M. and Rogers, J. C. (2003). Ohio River Valley winter moisture conditions associated with the Pacific-North American teleconnection pattern. *Journal of Climate* **16**: 969-981.

Conway, D., Wilby, R. L. and Jones, P. D. (1996). Precipitation and air flow indices over the British Isles. *Climate Research* **7**.

Corte-Real, J., Qian, B. D. and Xu, H. (1998). Regional climate change in Portugal: Precipitation variability associated with large-scale atmospheric circulation. *International Journal of Climatology* **18**: 619-635.

Covey, C., AchutaRao, K. M., Cubasch, U., Jones, P., Lambert, S. J., Mann, M. E., Phillips, T. J. and Taylor, K. E. (2003). An overview of results from the Coupled Model Intercomparison Project. *Global and Planetary Change* **37**: 103-133.

Craddock, J. M. and Flood, C. R. (1969). Eigenvectors for representing the 500-mb geopotential surface over the Northern Hemisphere. *Quarterly Journal of the Royal Meteorological Society* **95**.

Crane, R. G. and Hewitson, B. C. (1998). Doubled CO_2 precipitation changes for the Susquehanna Basin: down-scaling from the GENESIS general circulation model. *International Journal of Climatology* **18**: 65-76.

Cubasch, U., Meehl, G. A., Boer, G. J., Stouffer, R. J., Dix, M., Noda, A., Senior, C. A., Raper, S. and Yap, K. S. (2001). Projections of Climate Change. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggset al. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 881.

Cubasch, U., von Storch, H., Waszkewitz, J. and Zorita, E. (1996). Estimates of climate change in Southern Europe derived from dynamical climate model output. *Climate Research* **7**: 129-149.

Cubasch, U., Waszkewitz, J., Hegerl, G. and Perlwitz, J. (1995). Regional climate change as simulated in time-slice experiments. *Climatic Change* **31**: 273-304.

D'Agostino, R. B. and Stephens, M. A. (1986). Goodness-of-Fit Techniques, Marcel Dekker, Inc.

Davis, R. E. and Kalkstein, L. S. (1990). Using a spatial synoptic climatological classification to assess changes in atmospheric pollution concentrations. *Physical Geography* **11**: 320-342.

Dawdy, D. R. and Matalas, N. C. (1964). Statistical and probability analysis of hydrologic data, part III: Analysis of variance, covariance, and time series. Handbook of applied hydrology, a compendium of water-resources technology. V. T. Chow. New York, McGraw-Hill Book Company**:** 8.68-8.90.

Deque, M., Marquet, P. and Jones, R. (1998). Simulation of climate over Europe using a global variable resolution general circulation model. *Climate Dynamics* **14**: 173-189.

Dessouky, T. M. and Jenkinson, A. F. (1975). An objective daily catalogue of surface pressure, flow, and vorticity indices for Egypt and it's use in monthly rainfall forecasting. *Meteorological Research Bulletin, Egypt* **11**: 1-25.

Dickinson, R. E., Errico, E., Giorgi, F. and Bates, G. T. (1989). A regional climate model for the western United States. *Climatic Change* **12**: 383-422.

Dickson, R. R. and Namias, J. (1976). North American influences on the circulation and climate of the North Atlantic sector. *Monthly Weather Review* **104**: 1255-1265.

Easterling, D. R., Horton, B., Jones, P. D., Peterson, T. C., Karl, T. R., Parker, D. E., Salinger, M. J., Razuvayev, V., Plummer, N., Jamason, P. and Folland, C. K. (1997). Maximum and minimum temperature trends for the globe. *Science* **277**: 364-367.

Easterling, D. R., Karl, T. R., Lawrimore, J. H. and Del Greco, S. A. (1999). United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data for 1871-1997. Oak Ridge, Tennessee, Oak Ridge National Laboratory**:** 84.

El-Kadi, A. K. A. and Smithson, P. A. (1992). Atmospheric classifications and synoptic climatology. *Progress in Physical Geography* **16**: 432-455.

Flato, G., Boer, G. J., Lee, W. G., McFarlane, N. A., Ramsden, D., Reader, M. C. and Weaver, A. J. (2000). The Canadian Centre for Climate Modelling and Analysis global coupled model and its climate. *Climate Dynamics* **16**: 451-467.

Flato, G. M. and Boer, G. J. (2001). Warming asymmetry in climate change simulations. *Geophysical Research Letters* **28**: 195-198.

Folland, C. K., Karl, T. R., Christy, J. R., Clarke, R. A., Gruza, G. V., Jouzel, J., Mann, M. E., Oerlemans, J., Salinger, J. and Wang, S.-W. (2001). Observed Climate Variability and Change. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report on the Intergovernmental Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggset al. Cambridge, United Kingdom, and New York, NY, USA, Cambridge University Press**:** 881pp.

Foufoula-Georgiou, E. and Lettenmaier, D. P. (1987). A Markov renewal model for rainfall occurrences. *Water Resources Research* **23**: 875-884.

Fovell, R. G. and Fovell, M. C. (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. Available from http://www.ncdc.noaa.gov.

Fuentes, U. and Heimann, D. (2000). An improved statistical-dynamical downscaling scheme and its application to the Alpine precipitation climatology. *Theoretical and Applied Climatology* **65**: 119-135.

Gaffen, D. J. (1994). Temporal inhomogeneities in radiosonde temperature records. *Journal of Geophysical Research* **99**: 3667-3676.

Gaffen, D. J., Sargent, M. A., Habermann, R. E. and Lanzante, J. R. (2000). Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality. *Journal of Climate* **13**: 1776-1796.

Gibson, J. K., Kallberg, P., Uppala, S., Noumura, A., Hernandez, A. and Serrano, E. (1997). ERA Description. ECMWF Reanalysis Project Report Series. Reading, UK.

Giorgi, F., Hewitson, B. C., Christensen, J., Hulme, M., von Storch, H., Whetton, P., Jones, R., Mearns, L. and Fu, C. (2001). Regional Climate Information - Evaluation and Projections. Climate Change 2001: The Scientific Basis. Contribution of Working Group i to the Third Assessment Report of the Intergovernmentla Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer and P. J. van der Linden. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 881.

Giorgi, F., Shields, C. and Bates, G. T. (1994). Regional climate change scenario over the United States produced with a nested regional climate model. *Journal of Climate* **7**: 375-399.

Glahn, H. R. and Lowry, D. A. (1972). The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology* **11**: 1203-1211.

Glowienka-Hense, R. (1990). The North Atlantic Oscillation in the Atlantic-European SLP. *Tellus* **42A**: 497-507.

Gong, X. and Richman, M. B. (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.

Goodess, C. M. and Palutikof, J. P. (1998). Development of daily rainfall scenarios for Southeast Spain using a circulation-type approach to downscaling. *International Journal of Climatology* **10**: 1051-1083.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B. and Wood, R. A. (2000). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **16**: 147-168.

Greene, W. H. (2000). Econometric Analysis. Upper Saddle River, N.J., Prentice Hall.

Grotch, S. L. and MacCracken, M. C. (1991). The use of general circulation models to predict regional climate change. *Journal of Climate* **4**: 286-303.

Hanssen-Bauer, I., Forland, E. J., Haugen, J. E. and Tveito, O. E. (2003). Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling. *Climate Research* **25**: 15-27.

Hewitson, B. C. and Crane, R. G. (1996). Climate downscaling: techniques and application. *Climate Research* **7**: 85-95.

Heyen, H., Zorita, E. and 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.

Hirakuchi, H. and Giorgi, F. (1995). Multi-year present day and 2 x CO_2 simulations of monsoon climate over Asia and Japan with a regional climate model nested in a general circulation model. *Journal of Geophysical Research: Atmospheres* **100**: 21105-21126.

Hurrell, J. W. (1995). Decadal trends in the North Atlantic Oscillation and relationships to regional temperature and precipitation. *Science* **269**: 676-679.

Hurrell, J. W. (1996). Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperatures. *Geophysical Research Letters* **23**: 665-668.

Hurrell, J. W., Kushnir, Y., Ottersen, G. and Visbeck, M. (2003). The North Atlantic Oscillation: Climatic Significance and Environmental Impact.

Huth, R. (1999). Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Climate Research* **13**: 91-101.

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

Huth, R. (2002). Statistical downscaling of daily temperature in Central Europe. *Journal of Climate* **15**: 1731-1742.

IPCC (2000). Special Report on Emissions Scenarios. Cambridge, Cambridge University Press**:** 612pp.

IPCC (2001). Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom, and New York, NY, USA, Cambridge University Press.

Jenkinson, A. F. and Collison, P. (1977). An initial climatology of gales over the North Sea. London, Meteorological Office**:** 18.

Jones, P. D. (1994). Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *Journal of Climate* **7**: 1794-1802.

Jones, P. D., Hulme, M. and Briffa, K. R. (1993). A comparison of Lamb circulation types with an objective classification scheme. *International Journal of Climatology* **13**: 655-663.

Jones, P. D., Osborn, T. J. and Briffa, K. R. (1997). Estimating sampling erros in large-scale temperature averages. *Journal of Climate* **10**: 2548-2568.

Jones, R. G., Murphy, J. and Noguer, M. (1995). Simulation of climate change over Europe using a nested regional-climate model. I: Assessment of control climate,

including sensitivity to location of boundaries. *Quarterly Journal of the Royal Meteorological Society* **121**: 1413-1450.

Kalkstein, L. S., Tan, G. and Skindlov, J. A. (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, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437-470.

Kang, L.-S., Jin, K., Lau, K.-M., Shukla, J., Krishnamurthy, V., Schubert, S. D., Waliser, D. E., Stern, W. F., Satyan, V., Kitoh, A., Meehl, G. A., Kanamitsu, M., Galin, V. Y., Sumi, A., Wu, G., Liu, Y. and Kim, J.-K. (2002). Intercomparison of atmospheric GCM simulated anomalies associated with the 1997/98 El Nino. *Journal of Climate* **15**: 2791-2805.

Karl, T. R., Williams Jr., C. N. and Quinlan, F. T. (1990). United States Historical Climatology Network (HCN) Serial Temperature and Precipitation Data. Oak Ridge, Tennessee, Oak Ridge National Laboratory.

Kattenberg, A., Giorgi, F., Grassl, H., Meehl, G. A., Mitchell, J. F. B., Stouffer, R. J., Tokioka, T., Weaver, A. J. and Wigley, T. M. L. (1996). Climate Models - Projections of Future Climate. Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the IPCC. J. T. Houghton, L. G. Meira Filho, B. A. Callanderet al. Cambridge, Cambridge University Press**:** 285-357.

Kidson, J. W. and Thompson, R. D. (1998). A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate* **11**: 735-753.

Kilsby, C. G., Copertwait, P. S. P., O'Connell, P. E. and Jones, P. D. (1998). Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology* **18**: 523-539.

Kim, J. W., Chang, J. T., Baker, N. L., Wilks, D. S. and Gates, W. L. (1984). The statistical problem of climate inversion: Determination of the relationship between local and large-scale climate. *Monthly Weather Review* **112**: 2069-2077.

Kirchhofer, W. (1974). Classification of European 500 mb patterns. *Schwiezerische Meteorologische Anstalt, Institut Suisse de Meteorologie, Zurich* **43**: 1-16.

Klein, W. H. (1963). Specification of precipitation from the 700-millibar circulation. *Monthly Weather Review* **91**: 527-536.

Klein, W. H. (1982). Statistical weather forecasting on different time scales. *Bulletin on the American Meteorological Society* **63**.

Klein, W. H., Lewis, B. M. and Enger, I. (1959). Objective prediction of five-day mean temperatures during winter. *Journal of Meteorology* **16**: 672-682.

Kruizinga, S. and Murphy, A. H. (1983). Use of an analogue procedure to formulate objective probabilistic temperature forecasts in The Netherlands. *Monthly Weather Review* **111**: 2244-2254.

Leathers, D. J., Yarnal, B. and Palecki, M. A. (1991). The Pacific North-American Teleconnection Pattern and United-States Climate .1. Regional Temperature and Precipitation Associations. *Journal of Climate* **4**: 517-528.

Leung, L. R. and Ghan, S. J. (1999). Pacific Northwest climate sensitivity simulated by a regional climate model driven by a GCM. Part I: Control simulations. *Journal of Climate* **12**: 2031-2053.

Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences* **26**: 636-646.

Lund, I. A. (1963). Map-pattern classification by statistical methods. *Journal of Applied Meteorology* **2**: 56-65.

Marshall, G. J. (2002). Trends in Antarctic geopotential height and temperature: A comparison between radiosonde and NCEP-NCAR reanalysis data. *Journal of Climate* **15**: 659-.

Matalas, N. C. (1967). Mathematical assessment of synthetic hydrology. *Water Resources Research* **3**: 937-945.

McAveney, B. J., Covey, C., Joussaume, S., Kattsov, V., Kitoh, A., Ogana, W., Pitman, A. J., Weaver, A. J., Wood, R. A. and Zhao, Z.-C. (2001). Model Evaluation. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. J. T. Houghton, Y. Ding, D. J. Griggset al. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 881pp.

McGregor, J. L. and Walsh, K. (1994). Climate change simulations of Tasmanian precipitation using multiple nesting. *Journal of Geophysical Research: Atmospheres* **99**: 20889-20905.

McKendry, I. G., Steyn, D. G. and 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.

Meehl, G. A., Boer, G. J., Covey, C., Latif, M. and Stouffer, R. J. (2000). The Coupled Model Intercomparison Project (CMIP). *Bulletin of the American Meteorological Society* **81**: 313-318.

Meehl, G. A. and Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Nature* **305**: 994-997.

Mitchell, T. D. and Hulme, M. (1999). Predicting regional climate change: living with uncertainty. *Progress in Physical Geography* **23**: 57-78.

Murphy, J. (1999). An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate* **12**: 2256-2284.

Osborn, T. J., Briffa, K. R., Tett, S. F. B. and Jones, P. D. (1999). Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. *Climate Dynamics* **15**: 685-702.

Ott, R. L. (1993). An Introduction to Statistical Methods and Data Analysis. Belmont, Duxbury Press.

Overland, J. E. and Preisendorfer, R. W. (1982). A significance test for principal components applied to a cyclone climatology. *Monthly Weather Review* **110**: 1-4.

Palutikof, J. P., Goodess, C. M., Watkins, S. J. and Holt, T. (2002). Generating rainfall and temperature series at multiple sites: Examples from the Mediterranean. *Journal of Climate* **15**: 3529-3548.

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P. and Parker, D. (1998). Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology* **18**: 1493-1517.

Pope, V. D., Gallani, M. L., Rowntree, P. R. and Stratton, R. A. (2000). The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics* **16**: 123-146.

Portman, D. A., Wang, W.-C. and Karl, T. R. (1992). Comparison of general circulation model and observed regional climates: daily and seasonal variability. *Journal of Climate* **5**: 343-353.

Richardson, C. W. and Wright, D. A. (1984). WGEN: A model for generating daily weather variables, USDA-ARS**:** 83.

Richman, M. B. (1986). Rotation of principal components. *Journal of Climatology* **6**: 293-335.

Robeson, S. M. (2002). Increasing growing-season length in Illinois during the 20th century. *Climatic Change* **52**: 219-238.

Robeson, S. M. (2004). Trends in time-varying percentiles of daily minimum and maximum temperature over North America. *Geophysical Research Letters* **31**: L04203.

Sailor, D. J. and Li, X. (1999). A semiempirical downscaling approach for predicting regional temperature impacts associated with climate change. *Journal of Climate* **12**: 103-114.

Schar, C., Vidale, P. L., Luthi, D., Frei, C., Haberli, C., Liniger, M. A. and Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature* **427**: 332-336.

Schnur, R. and Lettenmaier, D. P. (1998). A case study of statistical downscaling in Australia using weather classification by recursive partitioning. *Journal of Hydrology* **212-213**: 362-379.

Schoof, J. T. and Pryor, S. C. (2001). Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. *International Journal of Climatology* **21**: 773-790.

Schoof, J. T. and Robeson, S. M. (2003). Seasonal and spatial variations of cross-correlation matrices used by stochastic weather generators. *Climate Research* **24**: 95-102.

Schoof, J. T. and Pryor, S. C. (2004). An evaluation of two GCMs: Simulations of North American teleconnections and synoptic phenomena. In preparation for submission to *Climate Research*.

Schubert, S. (1998). Downscaling local extreme temperature changes in south-eastern Australia from the CSIRO Mark2 GCM. *International Journal of Climatology* **18**: 1419-1438.

Semazzi, F. H. M., Lin, N., Lin, Y. and Giorgi, F. (1994). A nested model study of the Sahelian climate response to sea surface temperature anomalies. *Geophysical Research Letters* **20**: 2897-2900.

Semenov, M. A. and Barrow, E. M. (1997). Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change* **35**: 397-414.

Sheridan, S. C. (2003). North American weather-type frequency and teleconnection indices. *International Journal of Climatology* **23**: 27-45.

Sokal, R. R. and Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* **38**: 1409-1438.

Solman, S. A. and Nunez, M. N. (1999). Local estimates of global climate change: A statistical downscaling approach. *International Journal of Climatology* **19**: 835-861.

Sousounis, P. J. and Albercook, G. M. (2000). The Great Lakes Region: Past, Present, and Future. Preparing For A Changing Climate. The Potential Consequences of Climate Variability and Change. P. J. Sousounis and J. M. Bisanz.

Stephenson, D. B. and Pavan, V. (2003). The North Atlantic Oscillation in coupled climate models: a CMIP1 evaluation. *Climate Dynamics* **20**: 381-399.

Stidd, C. K. (1954). The use of correlation fields in relating precipitation to circulation. *Journal of Meteorology* **11**: 202-213.

Stone, R. C. (1989). Weather types at Brisbane, Queensland: An example of the use of principal components and cluster analysis. *International Journal of Climatology* **9**: 3-32.

Swail, V. R. and Cox, A. T. (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.

Taylor, K. E. (2001). Summarizing in a single diagram multiple aspects of model performance. *Journal of Geophysical Research* **106**.

Trenberth, K. E. and Stepaniak, D. P. (2002). A pathological problem with NCEP reanalyses in the stratosphere. *Journal of Climate* **15**: 690-695.

van den Dool, H. M. (1994). Searching for analogues, how long must we wait? *Tellus* **46A**: 314-324.

Walker, G. T. (1924). Correlation of seasonal variations in weather IX: A further study of world weather. *Memorandum of the Indian Meteorological Department* **24**: 275-332.

Walker, G. T. and Bliss, E. M. (1932). World Weather V. *Memorandum of the Royal Meteorological Society* **4**: 53-84.

Wallace, J. M. and Gutzler, D. S. (1981). Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly Weather Review* **109**: 784-812.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**: 236-244.

Weichert, A. and Burger, G. (1998). Linear versus nonlinear techniques in downscaling. *Climate Research* **10**: 83-93.

Werner, P. C. and von Storch, H. (1993). Interannual variability of Central European mean temperature in January/February and its relation to the large-scale circulation. *Climate Research* **3**: 195-207.

Wilby, R. L. (1998). Modelling low-frequency rainfall events using airflow indices, weather patterns and frontal frequencies. *Journal of Hydrology* **212-213**: 380-392.

Wilby, R. L. (1998). Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices. *Climate Research* **10**: 163-178.

Wilby, R. L., Hassan, H. and Hanaki, K. (1998). Statistical downscaling of hydrometeorological variables using general circulation model output. *Journal of Hydrology* **205**: 1-19.

Wilby, R. L. and Wigley, T. M. L. (1997). Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography* **21**: 530-548.

Wilby, R. L. and Wigley, T. M. L. (2000). Precipitation predictors for downscaling: observed and general circulation model relationships. *International Journal of Climatology* **20**: 641-661.

Wilks, D. S. (1992). Adapting stochastic weather generation algorithms for climate change studies. *Climatic Change* **22**: 67-84.

Wilks, D. S. (1999). Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology* **93**: 153-169.

Wilks, D. S. (1999). Multisite downscaling of daily precipitation with a stochastic weather generator. *Climate Research* **11**: 125-136.

Wilks, D. S. and Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* **23**: 329-357.

Willmott, C. J. (1987). Synoptic weather map classification: correlation versus sums-of-squares. *Professional Geographer* **39**: 205-207.

Winkler, J. A., Palutikof, J. P., Andresen, J. A. and Goodess, C. M. (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**: 2514-2535.

Xu, C. (1999). From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography* **23**: 229-249.

Yarnal, B. (1985). A 500 mb synoptic climatology of Pacific Northwest coast winters in relation to climatic variability, 1948-1949 to 1977-1978. *Journal of Climatology* **5**: 237-252.

Yarnal, B. (1993). Synoptic Climatology in Environmental Analysis. London, Belhaven Press.

Yarnal, B. and Leathers, D. J. (1988). Relationships between interdecadal and interannual climatic variations and their effect on Pennsylvania climate. *Annals of the Association of American Geographers* **78**: 624-641.

Yin, Z.-Y. (1994). Moisture conditions in the south-eastern USA and teleconnection patterns. *International Journal of Climatology* **14**: 947-967.

Zelenka, M. P. (1997). An analysis of the meteorological parameters affecting ambient concentrations of acid aerosols in Uniontown, Pennsylvania. *Atmospheric Environment* **31**: 869-878.

Zorita, E., Hughes, J. P., Lettenmaier, D. P. and von Storch, H. (1995). Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *Journal of Climate* **8**: 1023-1042.

Zorita, E. and von Storch, H. (1999). The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate* **12**: 2474-2489.

APPENDIX A

Schoof JT, Robeson SM. 2003. Seasonal and spatial variations of cross-correlation matrices used by stochastic weather generators. *Climate Research* **24**, 95-102.

Abstract

We examine seasonal and spatial variations of stochastic-weather-generator parameters and their impact on simulated weather sequences. Using daily weather observations from 29 stations across the contiguous United States, we estimate monthly station-specific parameters that are compared with the constant parameters that frequently are used in weather generator applications. A WGEN-type stochastic weather generator (SWG) is then used to generate a 100-year record of daily maximum and minimum air temperature, and daily total solar radiation at each station. These sequences are compared to sequences generated with constant parameters. While the means and standard deviations of the generated sequences are in agreement, the weather generator with station-specific parameters preserves relationships between variables. This is evident in both the lag-0 and lag-1 cross-correlations between generated variables and derived variables, such as diurnal temperature range. These results suggest that literature-based stochastic-weather-generator parameters may be appropriate for applications where monthly values of the means and standard deviations of generated variables are of interest. For applications that require proper simulation of relationships between variables, station-specific parameterizations are recommended.

Keywords: stochastic weather generator, climate simulation, climate variability, autoregressive parameters

1 Introduction

For many applications, the historical climate record is inadequate due to short or incomplete data records, or lack of appropriate spatial coverage. As a result, models of observed daily weather sequences, or stochastic weather generators (SWGs), are often used to supplement the historical record or to provide data for locations where weather data are not routinely collected (Johnson et al. 1996, Wilks & Wilby 1999). As time-series models with several interconnected components, SWGs simulate sequences for a number of variables, which typically include daily maximum and minimum air temperature (T_{max} and T_{min}) and total daily solar radiation (R), using a multivariate autoregressive process.

The generated sequences are designed to have the desired cross-correlations between T_{max}, T_{min}, and R by two matrices, **A** and **B**, which are estimated using lag-0 and lag-1 cross-correlations (see Sec. 3.2). In many weather-generator implementations, **A** and **B** are treated as constant with respect to location, time of year, and wet/dry status. Hayhoe (2000) examined bi-monthly variations in weather generator parameters for 3 stations in Canada and found spatial and seasonal variability in observed cross-correlations. Several authors (e.g., Wilks and Wilby 1999) have suggested using location- and time-specific parameters in an effort to account for spatial and seasonal variability.

In this study, the magnitude of the spatial and seasonal variability of these stochastic model parameterizations is investigated over a larger number of stations and wider range of climates than have been studied in the past. Using daily data from the conterminous U.S.A, we examine the spatial and seasonal differences in the values of the lag-0 and lag-1 cross-correlations, and hence **A** and **B**. We also examine differences between

simulated weather series when **A** and **B** are held constant and when they are allowed to vary by location and time of year. Rather than focusing on traditional statistics used to evaluate SWGs (monthly means and variances, monthly maximum and minimum values, freeze-free periods, etc.), we primarily examine (1) the correlation structure of the generated sequences and (2) variables that are closely related to the correlation structure, such as diurnal temperature range.

2 Data

To estimate the impacts of seasonally and spatially varying autoregressive parameters, daily data were extracted for 29 climatically diverse locations across the contiguous U.S.A. (Fig. 1), essentially the same stations used in the landmark study of Richardson & Wright (1984). Hourly air temperature, solar radiation, and precipitation values for 1961 to 1990 were available at these 29 locations through the Solar and Meteorological Surface Observation Network dataset (SAMSON, NCDC/NREL 1993) and updates, available from the National Climatic Data Center, Asheville, NC. To allow comparison and combination with cooperative climatic data, daily T_{max} and T_{min} (°C) were calculated using the now-common 7am observation time (Janis 2002). Daily total solar radiation values (MJ m^{-2} day^{-1}) were produced by numerical integration of hourly solar radiation observations.

Fig. 1. Map of the contiguous United States showing the 29 stations used in this study.

3 Description of Weather Generator

The SWG used in this research is based on the well-known and commonly used WGEN model (Richardson and Wright 1984). Using a number of parameters estimated from observed data, the model traditionally generates daily values of precipitation occurrence, precipitation amount, T_{max}, T_{min}, and R. Our primary objective was to evaluate the impact of varying the parameterizations of **A** and **B**; therefore, precipitation amount was not simulated. The individual components of the model are described below.

3.1 Precipitation Occurrence Component

Precipitation occurrence is simulated by a two-state, first-order Markov chain. The occurrence of precipitation depends on two parameters: p_{01}, the probability of a wet day following a dry day, and p_{11}, the probability of a wet day following a wet day. Depending on the precipitation occurrence simulation for the previous day, a uniform [0,1] random number is compared to the appropriate transition probability. If the random number is less than the transition probability, a wet day is simulated. Otherwise, a dry day is simulated. Previous work has shown that changes in precipitation parameters can effect the moments of conditioned variables generated by the model. Therefore, if a SWG is to be used in a climate change context, additional adjustments to model parameters may be needed (Katz 1996).

3.2 Temperature and Radiation Component

Daily values of T_{max}, T_{min}, and R are simulated by a first-order multivariate stochastic process, as described by Matalas 1967). To produce stationary time series, harmonic analysis first is used to construct annual cycles of daily means and standard deviations of the input variables. Using a Fourier transform, annual-cycles are fit to daily means and standard deviations using the first three harmonics. These annual cycle harmonics are fit separately for wet and dry days. In some cases, a given day of the year may have few wet or dry occurrences; therefore, a 15-day moving window was used to construct the daily means and standard deviations. In particularly dry areas, SWG users may consider using longer moving windows. The time series are then reduced to

standardized residual elements by subtracting the daily means and dividing by the standard deviations, as defined by the harmonics.

The weather generator simulates daily residuals of T_{max}, T_{min}, and R for day *i* with the equation

$$\mathbf{X_i} = \mathbf{A}\,\mathbf{X_{i-1}} + \mathbf{B}\boldsymbol{\varepsilon_i} \qquad (1)$$

where $\mathbf{X_i}$ is a (3 × 1) matrix containing the current day's standardized values of T_{max}, T_{min}, and R, $\mathbf{X_{i-1}}$ is a (3 × 1) matrix containing the previous day's standardized values of T_{max}, T_{min}, and R, $\boldsymbol{\varepsilon_i}$ is a (3 × 1) vector of independent values from a standard Normal distribution, and **A** and **B** are (3 × 3) matrices given by

$$\mathbf{A} = \mathbf{M_1}\mathbf{M_0^{-1}} \qquad (2)$$

$$\mathbf{B}\mathbf{B}^{\mathbf{T}} = \mathbf{M_0} - \mathbf{M_1}\mathbf{M_0^{-1}}\mathbf{M_1^{T}} \qquad (3)$$

where $\mathbf{M_o}$ is the (3 × 3) matrix of lag-0 cross correlations and $\mathbf{M_1}$ is the (3 × 3) matrix of lag-1 cross correlations. For example, $\mathbf{M_o}$ (1,2) is the correlation between T_{max} and T_{min} and $\mathbf{M_1}$ (1,2) is the correlation between T_{max} and T_{min} lagged by one day. While **A** can be estimated directly, **B** is estimated by defining a new matrix $\mathbf{Z}=\mathbf{B}\mathbf{B}^{\mathbf{T}}$ (see Greene 2000). Then by spectral decomposition, $\mathbf{Z}=\mathbf{C}\mathbf{L}\mathbf{C}^{\mathbf{T}}$, where **C** is the matrix of eigenvectors of $\mathbf{B}\mathbf{B}^{\mathbf{T}}$ and **L** has the eigenvalues of $\mathbf{B}\mathbf{B}^{\mathbf{T}}$ on the diagonal and zeros elsewhere. Since $\mathbf{B}\mathbf{B}^{\mathbf{T}}=\mathbf{Z}^{1/2}\mathbf{Z}^{1/2\mathbf{T}}=\mathbf{Z}$, $\mathbf{B}=\mathbf{Z}^{1/2}$. Then by Greene's Theorem 2.10, estimates of **B** can then be computed as $\mathbf{B}=\mathbf{C}\mathbf{L}^{\mathbf{1/2}}\mathbf{C}^{\mathbf{T}}$. After generation of the residual series with Eq. 1, the daily harmonics described above are used to produce dimensional values of T_{max}, T_{min}, and R, based on wet/dry status.

The use of a standard normal distribution for all three elements of $\boldsymbol{\varepsilon_i}$ may not be appropriate in many situations, particularly for solar radiation. For this reason, some weather generators (e.g., LARS-WG (Semenov & Barrow 1997)) have used more complex distributions for R. Other SWGs (e.g., CLIGEN (Nicks and Gander, 1993, 1994)) have addressed this issue by constraining the generated R data between a maximum value based on station location and sun angle and a minimum value of 5% of the maximum value. In an examination of 15 US climate stations, Harmel et al. (2002) found that even T_{max} and T_{min} were not generally normally distributed in each month, results that have wide implications for further weather generator research. Nonetheless, the focus of this research is not on $\boldsymbol{\varepsilon_i}$ and the vast majority of stochastic weather generators are still based on assumptions of normality. Weather generator users must decide how these assumptions impact their particular application.

4 Observed Relationships

In many implementations, WGEN-type models use fixed values of $\mathbf{M_0}$ and $\mathbf{M_1}$ — and therefore **A** and **B** — irrespective of location and time of year. Richardson (1982) provides the following values:

$$\mathbf{M_0} = \begin{pmatrix} 1.000 & 0.633 & 0.186 \\ 0.633 & 1.000 & -0.193 \\ 0.186 & -0.193 & 1.000 \end{pmatrix} \qquad \mathbf{M_1} = \begin{pmatrix} 0.621 & 0.445 & 0.087 \\ 0.563 & 0.674 & -0.100 \\ 0.015 & -0.091 & 0.250 \end{pmatrix}$$

$$\mathbf{A} = \begin{pmatrix} 0.567 & 0.086 & -0.002 \\ 0.253 & 0.504 & -0.050 \\ -0.006 & -0.039 & 0.244 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 0.781 & 0.000 & 0.000 \\ 0.328 & 0.637 & 0.000 \\ 0.238 & -0.341 & 0.873 \end{pmatrix} \tag{4}$$

While the literature-based correlation matrices may be appropriate during some seasons at some locations, the observed and literature-based (constant) correlation matrices can be very different when location and the entire calendar year are considered. The differences in the correlation matrices ultimately dictate the variability in the estimated elements of the **A** and **B** matrices. In the following sections, observed values of $\mathbf{M_0}$ and $\mathbf{M_1}$, and hence the estimates of **A** and **B**, are examined.

4.1 Seasonal/Spatial Variability of Lag-0 Correlation Coefficients (M_0)

The literature-based values of $\mathbf{M_0}$ agree with observations at some locations during some months; however, examination of station-specific monthly correlations suggests that the literature values may not be appropriate for all locations year-round (Fig. 2, Fig. 3).

During the late summer months, the literature-based value of $\mathbf{M_0}(1,2)$ — 0.633— is similar to values observed at many stations. However, during the winter months, the literature-based correlation is lower than the observed correlation at most stations (Fig. 2a, Fig. 3a). The correlations between temperature and radiation ($\mathbf{M_0}(1,3)$ and $\mathbf{M_0}(2,3)$) are more seasonally and spatially variable than $\mathbf{M_0}(1,2)$ (Fig. 2b,c, Fig. 3b,c) . For these elements, the literature-based correlations are appropriate at some locations during the transition seasons, but are generally too strong during the winter months and too weak during the summer months.

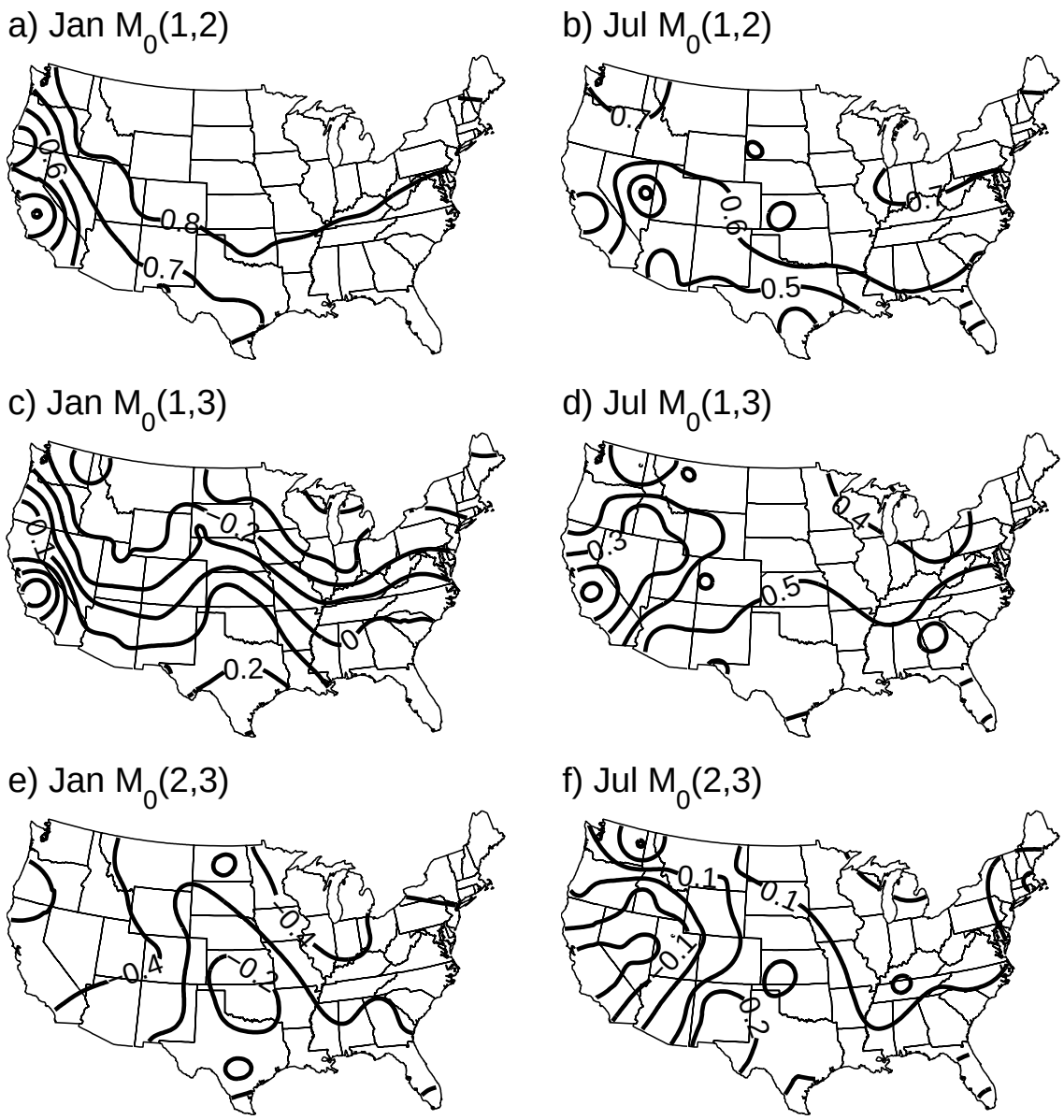

Fig. 2. Contour plots of lag-0 correlations (**M_0**): a) January correlation between daily T_{max} and T_{min} (**M_0**(1,2)), b) July correlation between daily T_{max} and T_{min} (**M_0**(1,2)), c) January correlation between T_{max} and R (**M_0**(1,3)), d) July correlation between T_{max} and R (**M_0**(1,3)), e) January correlation between daily T_{min} and R (**M_0**(2,3)) and f) July correlation between daily T_{min} and R (**M_0**(2,3)). Literature-based values of **M_o** are given in Eq. 4.

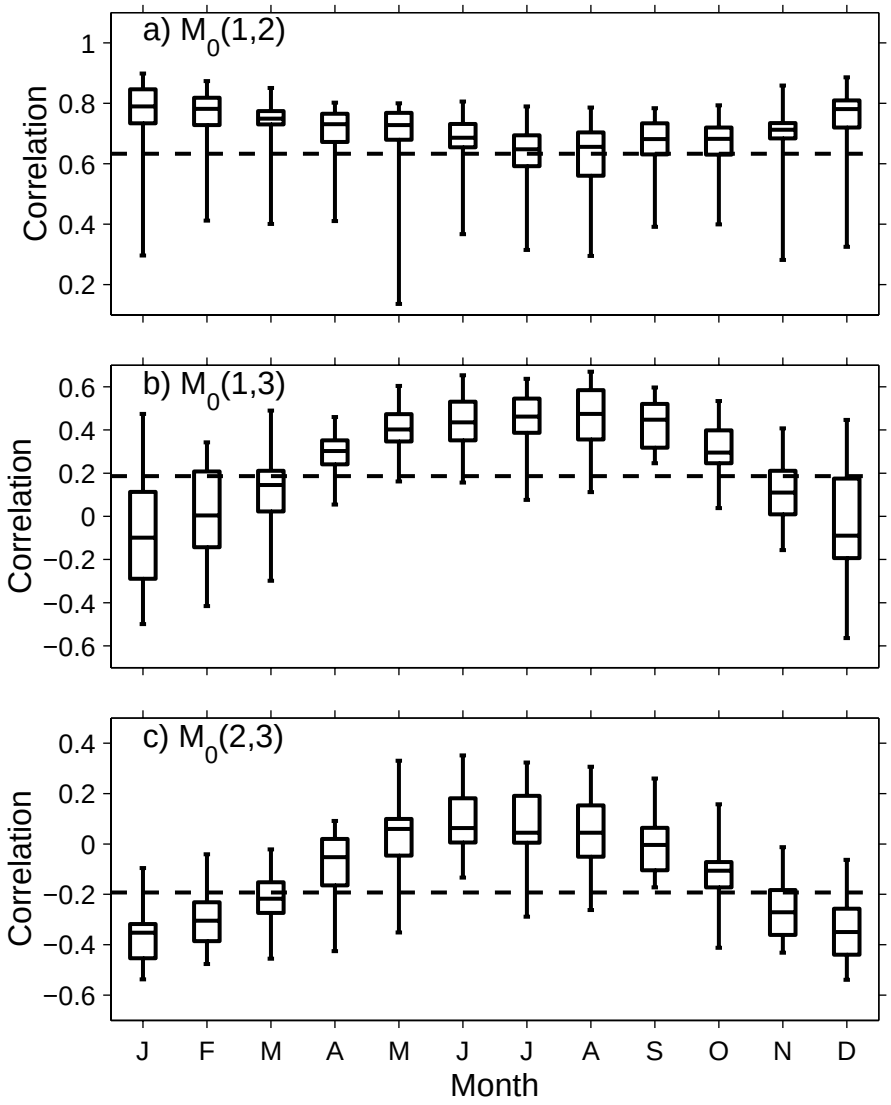

Fig. 3. Distributions (boxplots) of lag-0 correlations ($\mathbf{M_0}$): a) correlation between daily T_{max} and T_{min} ($\mathbf{M_0}(1,2)$), b) correlation between T_{max} and R ($\mathbf{M_0}(1,3)$), and c) correlations between daily T_{min} and R ($\mathbf{M_0}(2,3)$). Each box shows the distribution of correlations across the 29-station network, and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value.

4.2 Seasonal/Spatial Variability of Lag-1 Correlation Coefficients (M_1)

The data used in this study suggest that the elements of $\mathbf{M_1}$, the lag-1 correlation matrix, are also seasonally and spatially variable within the contiguous USA. As with the elements of $\mathbf{M_0}$, the literature-based values are appropriate at some locations and some

times of year (e.g., $\mathbf{M_1}(2,2)$ in November , Fig. 4e), but fail to accommodate the range of values observed over the study area in most months (Fig. 4). For some elements of $\mathbf{M_1}$ the literature-based value is entirely outside the range of observations during particular months (e.g., $\mathbf{M_1}(2,3)$ in November, December, January, and February, Fig. 4h; note that in all of the boxplots shown, spatial variability can be inferred from the amount of variation in any given box-and-whiskers, although some of the variation also is due to sampling variability).

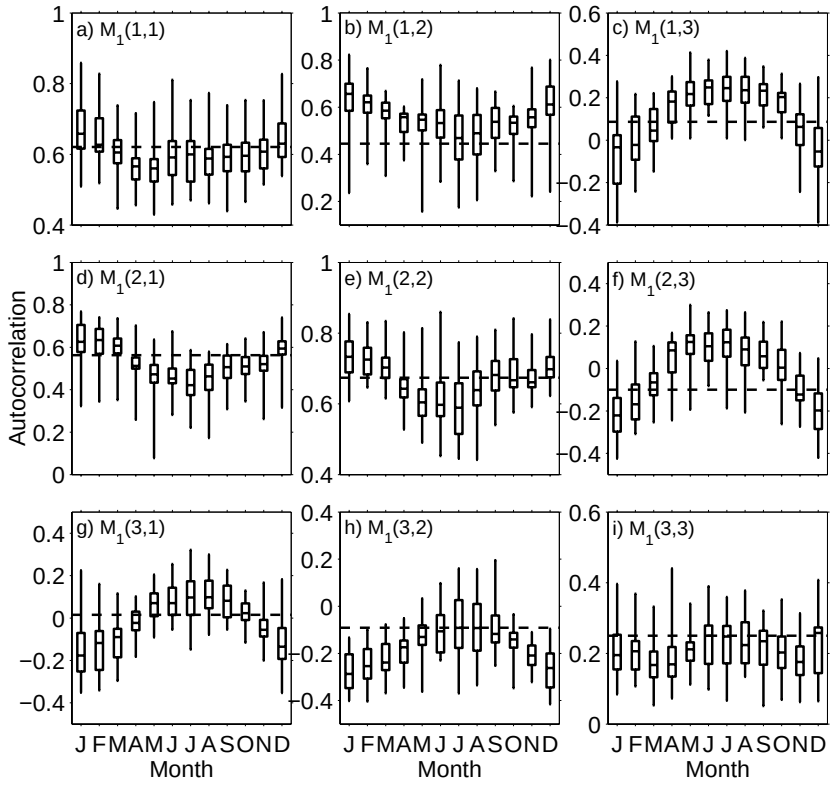

Fig. 4. Boxplots of lag-1 correlations ($\mathbf{M_1}$): a) lag-1 correlation between T_{max} and T_{max}, b) lag-1 correlation between T_{max} and T_{min}, c) lag-1 correlation between T_{max} and R, d) lag-1 correlation between T_{min} and T_{max}, e) lag-1 correlation between T_{min} and T_{min}, f) lag-1 correlation between T_{min} and R, g) lag-1 correlation between R and T_{max}, h) lag-1 correlation between R and T_{min}, and i) lag-1 correlation between R and R. Each box shows the distribution of correlations across the 29-station network, and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value.

4.3 Seasonal/Spatial Variability of A and B

Although seasonal variations in the elements of **A** are generally small, several elements of **A** exhibit substantial spatial variability in individual months (Fig. 5). The observed values of each element are generally in poor agreement with literature-based values, with most stations consistently having values that differ from the literature-based values. In general, the individual elements of **B** exhibit more seasonal variability and less spatial variability than the elements of **A** (Fig. 6). Observed values of **B**(1,2) and **B**(3,2) are different from the literature-based values at all stations during all months. For other elements of **B**, especially those exhibiting seasonal variability, the literature-based values are appropriate only at particular times and locations. In general, the literature-based values of **B** do not agree with those computed with spatially and seasonally variable values of $\mathbf{M_0}$ and $\mathbf{M_1}$ (Fig. 6). Although the estimated **A** and **B** elements vary over relatively large spatial scales, they do not have obvious relationships with physiographic characteristics, such as latitude, longitude, and elevation. This result has implications for interpolation of weather generator parameters (e.g., Semenov & Brooks 1999).

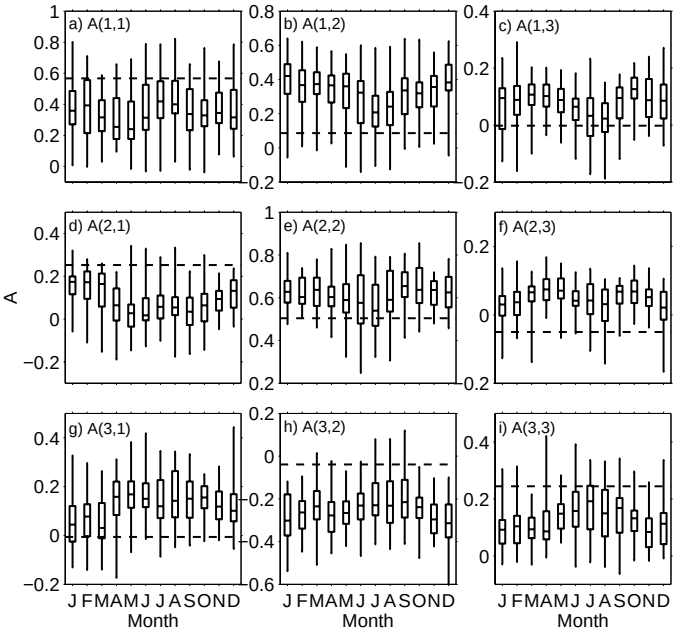

Fig. 5. Boxplots of elements of **A**: a) **A**(1,1), b) **A**(1,2), c) **A**(1,3), d) **A**(2,1), e) **A**(2,2), f) **A**(2,3), g) **A**(3,1), h) **A**(3,2), and i) **A**(3,3). Each box shows the distribution of coefficients across the 29-station network, and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value.

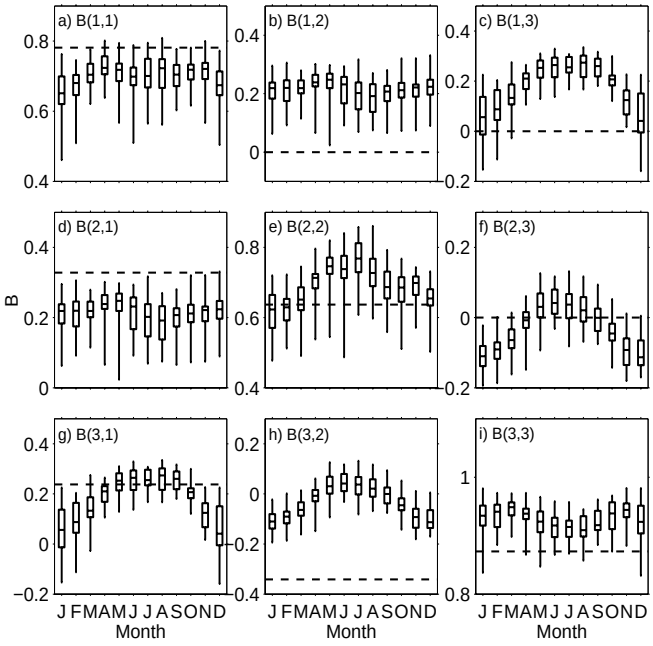

Fig. 6. Boxplots of elements of **B**: a) **B**(1,1), b) **B**(1,2), c) **B**(1,3), d) **B**(2,1), e) **B**(2,2), f) **B**(2,3), g) **B**(3,1), h) **B**(3,2), and i) **B**(3,3). Each box shows the distribution of coefficients across the 29-station network, and depicts the maximum and minimum values, as well as the inter-quartile range and median. The dashed line represents the literature-based value.

5 Weather Generator Implementation

Variability in the elements of the **A** and **B** matrices suggests that using station-specific parameters may have important impacts on data generated with a SWG. To investigate the effects of these spatially and seasonally varying parameterizations, the weather generator described in Sec. 3 was used to produce100-year sequences of daily T_{max}, T_{min}, and R for each station in our analysis (Fig. 1, as described in Sec. 2). The weather generator was run in two modes. First, the elements of **A** and **B** were held constant according to the literature-based values (i.e., values given by Richardson 1982), producing generated data that is hereafter referred to as LGEN. In the second mode, the monthly values of **A** and **B** estimated from historical data for each individual station were used, producing data that is hereafter referred to as ABGEN.

6 Evaluation of Generated Data

6.1 Means and Standard Deviations

Means and standard deviations of the generated variables (not shown) are in general agreement with observations for both versions of the generator (LGEN and ABGEN). These means and standard deviations are largely dependent on the harmonics used to depict the means and standard deviations – and not on the correlation structure of the variables. Since these harmonics do not differ between the ABGEN and LGEN simulations, the small differences in these values are not unexpected. For the eastern half of the contiguous USA, differences between observed and generated T_{max} are less than 1ºC in all months. For stations in the western USA, differences are less than 1ºC during the summer months and 1-2ºC during the winter months. Differences between

observed and generated T_{min} are less than 1ºC at all locations during the spring and summer months, with differences of 1-2ºC in the east during the winter. During the late autumn and winter months (Oct-Feb), differences between observed and generated R are less than 1 $MJm^{-2}day^{-1}$ in the western half of the US. During the spring, these differences are as large as 3 $MJm^{-2}day^{-1}$. The errors in R are not unexpected and result from poor agreement in the lower tail of the distribution. While observed values are physically bounded at zero, modeled values of R can become negative. In the 100-year simulations conducted here, negative R values were not generated at most stations. When this fundamental simulation error did occur, R was set to zero. Some alternatives, such as constraining the generated R values in physically plausible ways are available (see Section 3.2).

ABGEN and LGEN produce means and standard deviations with nearly identical spatial and temporal variability (and therefore are not shown). In terms of these means and standard deviations, the station-specific ABGEN does not provide any improvement over the constant-parameter LGEN.

6.2 Correlations Between Generated Variables

While correlations between simulated variables are not routinely used to evaluate SWGs, preservation of the correlation structure between the variables is critical for impacts modeling in agriculture and hydrology where multiple input series of daily weather variables are routinely employed. Since the correlation structure between the generated variables is fundamentally dependent on the values of **A** and **B**, the station-specific generator should better replicate the observed correlations between variables.

Lag-0 and lag-1 cross-correlations confirm that station-specific, monthly parameterization of **A** and **B** produces a better match between simulated and observed correlations (Fig. 2, Fig. 7). Data simulated with the constant, literature-based values for **A** and **B** resulted in larger differences between observed and generated correlations (Fig. 2, Fig. 7; note that the correlations depicted in Fig. 7 are computed using the generated data and include the annual cycle harmonics; therefore, some of the variability in these maps results from differences in the harmonics).

For each element of $\mathbf{M_0}$, absolute differences between observed and generated values are larger for LGEN than ABGEN. The maximum absolute differences between observed and generated $\mathbf{M_0}(1,2)$, the correlation between T_{max} and T_{min}, are 0.49 for LGEN and 0.27 for ABGEN. For $\mathbf{M_0}(1,3)$, the correlation between T_{max} and R, the maximum absolute differences are 0.68 for LGEN, compared to 0.23 for ABGEN. Maximum absolute differences between observed and generated $\mathbf{M_0}(2,3)$, the correlation between T_{min} and R, are 0.46 for LGEN and 0.22 for ABGEN.

The lag-1 cross-correlations computed with the generated data are also different from the observed lag-1 cross-correlations ($\mathbf{M_1}$). For ABGEN, the absolute differences between generated and observed elements of $\mathbf{M_1}$ are typically small (<0.1 for all temperature-temperature correlations and <0.2 for all temperature-radiation correlations). For LGEN, the differences can be quite large. For example, the maximum absolute difference between generated and observed $\mathbf{M_1}(2,1)$ for the month of May is 0.48 for LGEN, compared to only 0.10 for ABGEN.

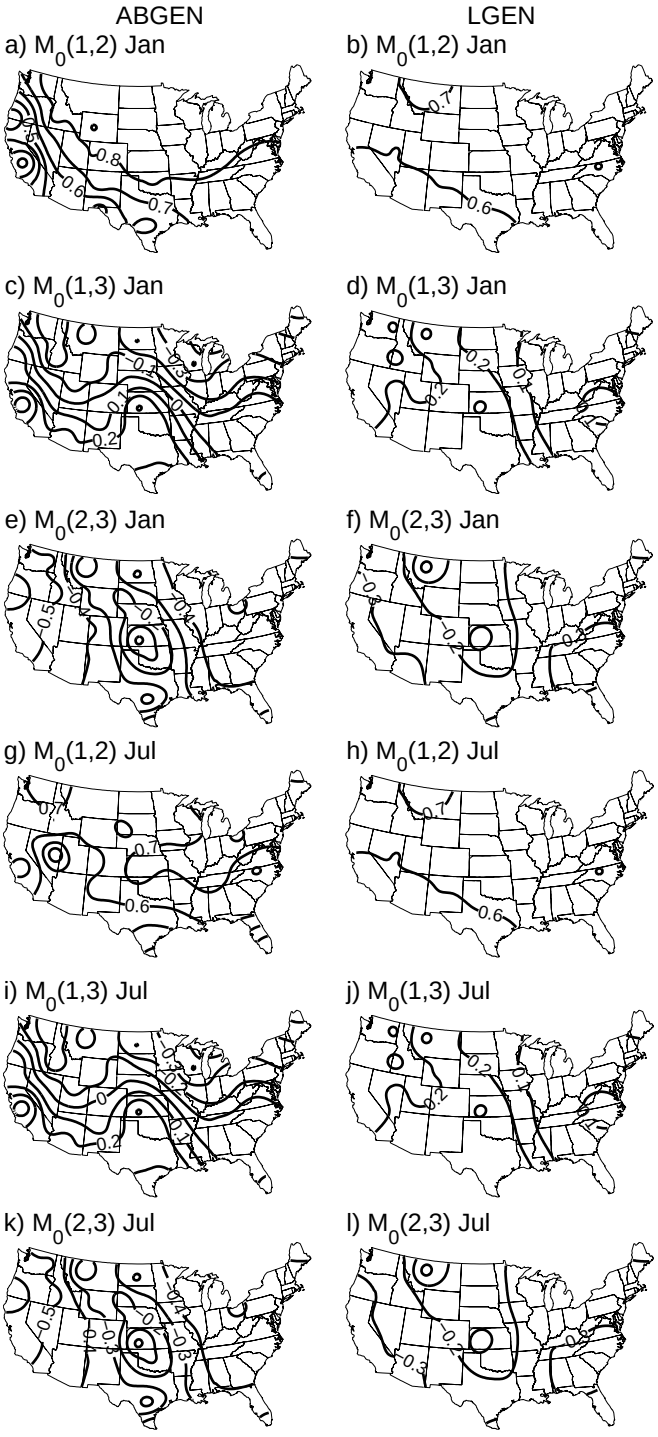

Fig. 7. Contour plots of lag-0 correlations for ABGEN (left) and LGEN (right) generated data: a) January $\mathbf{M_0}(1,2)$ ABGEN, b) January $\mathbf{M_0}(1,2)$ LGEN, c) January $\mathbf{M_0}(1,3)$ ABGEN, d) January $\mathbf{M_0}(1,3)$ LGEN, e) January $\mathbf{M_0}(2,3)$ ABGEN, f) January $\mathbf{M_0}(2,3)$ LGEN, g) July $\mathbf{M_0}(1,2)$ ABGEN, h) July $\mathbf{M_0}(1,2)$ LGEN, i) July $\mathbf{M_0}(1,3)$ ABGEN, j) July $\mathbf{M_0}(1,3)$ LGEN, k) July $\mathbf{M_0}(2,3)$ ABGEN, and l) July $\mathbf{M_0}(2,3)$ LGEN. The literature-based values are given in Eq. 4.

6.3 Diurnal Temperature Range (DTR)

Diurnal temperature range (DTR ≡ T_{max} - T_{min}) also can be an effective evaluation tool for SWGs. Given that DTR is a function of both T_{max} and T_{min}, its accurate simulation requires that the relationships between these two variables be preserved. Using the 100-year simulations described above, LGEN (with constant **A** and **B**) produces many more days with negative DTR (i.e., a fundamental simulation error) than ABGEN at most stations used in this study (Fig. 8). While both weather generators (LGEN and ABGEN) simulated monthly means and standard deviations of generated variables well, the frequency distribution of DTR is not simulated as well by the literature-based generator. Although monthly mean DTR is similar in both models, the station-based generator produces much better agreement between observed and simulated standard deviation of DTR (Fig. 9), especially during the winter months. The cause for these simulation errors in LGEN can be traced to $\mathbf{M_0}(1,2)$, the element of the correlation matrix that relates the current day's T_{max} and T_{min}. Errors in the LGEN standard deviation of DTR are highly correlated with differences in literature-based and observed values of $\mathbf{M_0}(1,2)$ (monthly correlations range from –0.92 to –0.70, significant at the 99% level).

The station-specific generator also reproduces the relationships between temperature and radiation more accurately. Because DTR is closely linked to cloud cover and precipitation (Leathers et al. 1998), and radiation is a reasonable surrogate for cloud cover, allowing the relationships between temperature and radiation to vary by location and time of year helps to improve the simulation of temporal variability in DTR.

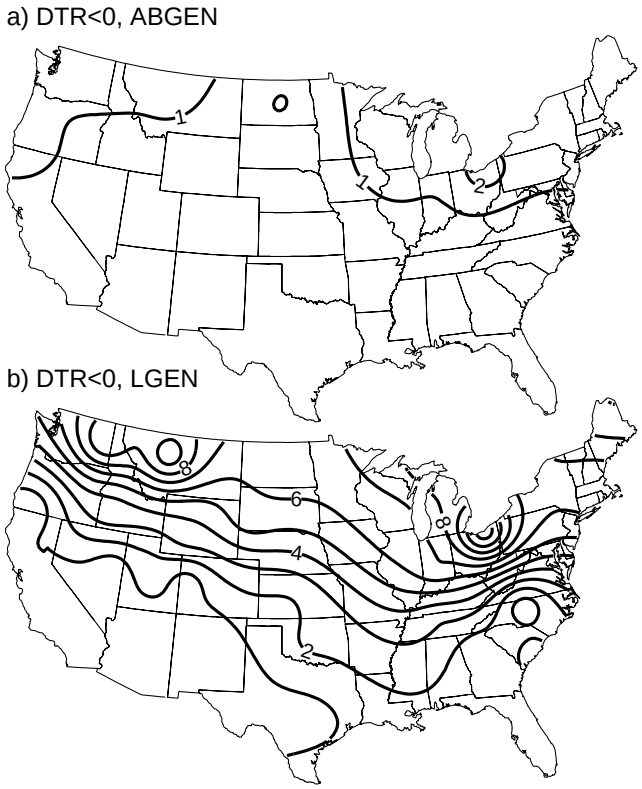

Fig. 8. Contour plots of the number of times per year (in a 100-year simulation) the weather generator simulates negative diurnal temperature range for a) ABGEN and b) LGEN.

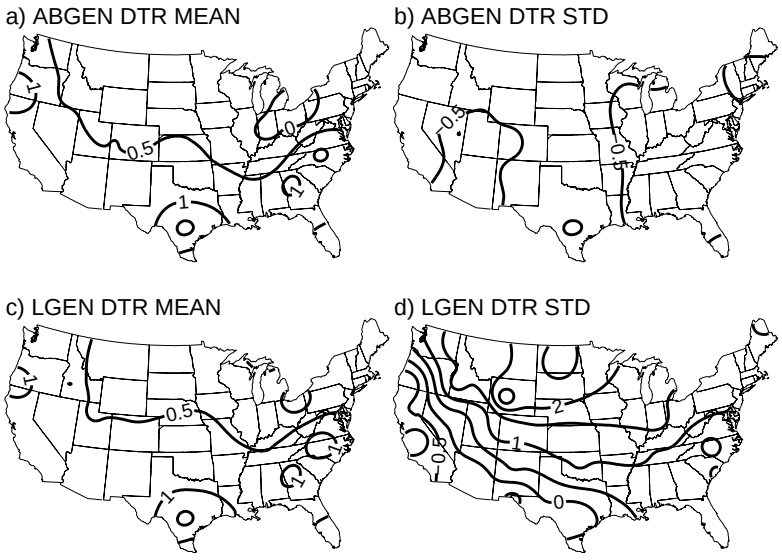

Fig. 9. Contour plots of the errors in simulating the mean (left) and standard deviation (right) of diurnal temperature range (DTR) during January: a) ABGEN mean DTR minus observed mean DTR, and b) ABGEN standard deviation of DTR minus observed standard deviation of DTR, c) LGEN mean DTR minus observed mean DTR, and d) LGEN standard deviation of DTR minus observed standard deviation of DTR.

7 Discussion and Conclusions

In this study, the effects of SWG parameterizations have been investigated. Using historical data from 29 stations in the United States, we examined the spatial and seasonal differences in the lag-0 and lag-1 cross-correlations between T_{max}, T_{min}, and R. These correlations ultimately determine the nature of the **A** and **B** matrices used in the SWG, and were found to have profound spatial and seasonal variations.

To investigate the impacts of the seasonal and spatial variability in the elements of these matrices, 100-year simulations for 29 stations were undertaken with (1) **A** and **B** assumed constant (values from Richardson 1982), and (2) **A** and **B** computed for each individual station on a monthly basis.

The simulations were compared to observed data using statistical and graphical methods. The results suggested that monthly means and standard deviations of each simulated variable agree with observed values for both simulations; however, the literature-based generator failed to preserve relationships between variables. This shortcoming is evident in both the simulated diurnal temperature range (DTR) and in the correlations between simulated variables.

These results suggest that literature-based values may be appropriate for applications where monthly values of the means and standard deviations of generated variables are of interest. For applications that require proper simulation of relationships between variables, station-specific parameterizations are more appropriate. In addition, because SWGs are now being used in climate-change studies (e.g., GCM downscaling research; Semenov & Barrow 1997, Wilks 1999), additional caution is warranted. While weather-

generator parameters will certainly change as climate changes, the magnitude of changes will vary seasonally and spatially.

Acknowledgement. This manuscript was substantially improved by the comments of the editor and two anonymous reviewers.

Literature Cited

Greene WH (2000) Econometric Analysis. Prentice Hall, Upper Saddle River, NJ.

Harmel RD, Richardson CW, Hanson CL, Johnson GL (2002) Evaluating the adequacy of simulating maximum and minimum daily air temperature with the Normal distribution. J Appl Meteorol 41: 744-753.

Hayhoe HN (2000) Improvements in stochastic weather generators for diverse climates. Clim Res 14: 75-87

Janis MJ (2002) Observation-time-dependent biases and departures for daily minimum and maximum air temperatures. J Appl Meteorol 41: 588-603.

Johnson GL, Hanson CL, Hardegree SP, Ballard EB (1996) Stochastic weather simulation: overview and analysis of two commonly used models. J Appl Meteorol 35: 1878-1896.

Katz RW (1996) Use of conditional stochastic models to generate climate change scenarios. Clim Change 32: 237-255.

Leathers DJ, Palecki MA, Yarnal B (1998) Climatology of the daily temperature range annual cycle in the United States. Clim Res 9: 197-211.

Matalas NC (1967) Mathematic assessment of synthetic hydrology. Water Resour Res 9: 937-945.

NCDC/NREL (1993) Solar and Meteorological Surface Observation Network (SAMSON, available from NCDC). Available at http://www.ncdc.noaa.gov.

Nicks AD, Gander GA (1993) Using CLIGEN to stochastically generate climate data inputs to WEPP and other water resource models. USGS Water-Resources Investigations Report 93-4018, USGS, Fort Collins, CO.

Nicks AD, Gander GA (1994) CLIGEN: a SWG for climate inputs to water resource and other models. In: Watson DG, Zarzueta FS, Harrison TV (eds). Proceedings of the 5th International Conference on Computers in Agriculture, Orlando, FL. American Society of Agricultural Engineers, St. Joseph, MI, p 903-909.

Richardson CW (1982) Dependence structure of daily temperature and solar radiation. Trans Am Soc Agric Eng 25: 735-739.

Richardson CW, Wright DA (1984) WGEN: A model for generating daily weather variables. US Dept. of Agriculture, Agricultural Research Service, Publ. No. ARS-8, Washington, DC.

Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. Clim Change 35: 397-414.

Semenov MA, Brooks RJ (1999) Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. Clim Res 11: 137-148.

Wilks DS (1999) Multisite downscaling of daily precipitation with a stochastic weather generator. Clim Res 11: 125-136.

Wilks DS, Wilby RL (1999) The weather generator game: a review of stochastic weather models. Prog Phys Geog 23: 329-357.

Justin T. Schoof

Atmospheric Science Program
Department of Geography
Indiana University
Bloomington, IN 47405
Phone: (812) 855-7956
Fax: (812) 855-1661
E-mail: jschoof@indiana.edu
http://mypage.iu.edu/~jschoof

EDUCATION

Ph.D., Indiana University, Atmospheric Science Program, Department of Geography, September 2004.
Minor: SPEA (School of Public and Environmental Affairs)
Dissertation title: *Generation of regional climate change scenarios using general circulation models and empirical downscaling*
Graduate Advisor: Dr. Sara C. Pryor
Graduate Committee: Dr. Rebecca Barthelmie, Dr. Scott Robeson
External Committee Member: Dr. David F. Parkhurst, SPEA

M.Sc., Indiana University, Atmospheric Science Program, Department of Geography, December 1999.
Thesis title: *Synoptic circulation classification and downscaling for the Midwest United States*

B.A., Indiana University, May 1997
Majors: Geography/Mathematics
1.1.

PROFESSIONAL EXPERIENCE

2003-2004	Research Assistant for Dr. Sara C. Pryor
2003	Instructor (2002-2003 Spring Semester): Indiana University - Bloomington *G109: Weather and Climate* (3 cr) 165 students

2000-2002 Research Assistant for Dr. Sara C. Pryor; *Nitrogen deposition to a mid-latitude deciduous forest, modeling of atmosphere-ocean exchange in the marine boundary layer, synoptic weather type classification and empirical downscaling for the Midwest United States*

2002 Instructor (2002-2003 Fall Semester): Indiana University – Purdue University at Indianapolis Department of Geography
G111: Hurricanes (1 cr)
30 students

1999 Associate Instructor (1999-2000 Fall Semester): Indiana University Department of Geography:
Weather and Climate
Supervisor: Dr. Hans Peter Schmid
Duties: Instruction of 90 students in 3 laboratory sessions

1997-1999 Research Assistant for Dr. Sue Grimmond, Dr. Scott Robeson, and Dr. Hans Peter Schmid; *Installation and maintenance of climatological instrumentation, collection and analysis of data, and publication of results*

FELLOWSHIPS, GRANTS, AND AWARDS

2003-2004 *Indiana University College of Arts and Sciences Dissertation Year Research Fellowship*. $13,500

2003 *Indiana University James H. Coon Science Prize.* Awarded a student who shows promise in one of the sciences). $1,500

2003 *Indiana University Graduate Student Travel Support Grant*. Indiana University College of Arts and Sciences. Awarded to attend American Meteorological Society 2004 Annual Meeting, Seattle, WA, 11-15 January, 2004. $300

2002 *Indiana University Department of Geography Chairman's Graduate Student Recognition Award*. Awarded for outstanding academic performance.

2001 *Indiana University Department of Geography Departmental Graduate Fellowship Award*. Awarded for academic excellence. $1,500

2001 *Indiana University Department of Geography Chairman's Graduate Student Recognition Award*. Awarded for outstanding academic performance.

2000 *Indiana University Esther L. Kinsley Master's Thesis Award*. Awarded for outstanding theses at Indiana University. $500

2000 *Indiana University Department of Geography Chairman's Graduate Student Recognition Award*. Awarded for outstanding academic performance.

2000 *American Meteorological Society Global Change Travel Scholarship*. Awarded to attend the American Meteorological Society 81st Annual Meeting, Albuquerque, NM, 14-19 January, 2000. Approximately $500.

1999 *Indiana University Graduate Student Travel Support Grant*. Indiana University College of Arts and Sciences. Awarded to attend American Geophysical Union 1999 Spring Meeting, Boston, MA, 1-4 June, 1999. $200

1999 *Indiana University Department of Geography Steven S. Visher Award for Outstanding Paper in Climatology*. $500

1999 *Indiana University Department of Geography Chairman's Graduate Student Recognition Award*. Awarded for outstanding academic performance.

REFEREED JOURNAL PUBLICATIONS

In Review:

Pryor SC, Barthelmie RJ, Schoof JT. 2004. The impact of non-stationarities in the climate system on the definition of 'a normal wind year': A case study from the Baltic. *International Journal of Climatology*

Published:

Schoof JT, Pryor SC. 2003. Evaluation of the NCEP/NCAR reanalysis in terms of synoptic scale phenomena: A case study from the Midwestern USA. *International Journal of Climatology* **23**: 1725-1741.

Schoof JT, Robeson SM. 2003. Seasonal and spatial variations of cross-correlation matrices used by stochastic weather generators. *Climate Research* **24**, 95-102.

Schoof JT, Pryor SC. 2001. Local estimates of temperature and precipitation: A comparison of two circulation-based downscaling methods. *International Journal of Climatology* **21**, 773-790.

Pryor SC, Barthelmie RJ, Carreiro M, Davis ML, Hartley A, Jensen B, Oliphant A, Randolph JC, Schoof JT. 2001. Nitrogen deposition to and cycling in a deciduous forest. In Optimizing Nitrogen Management in Food and Energy Production and Environmental Protection: *Proceedings of the 2nd International Nitrogen Conference on Science and Policy. The Scientific World* **1**(S2), 245-254.

Pryor SC, Barthelmie RJ, Schoof JT, Sorensen LL, Erickson DJ. 2001. Implications of heterogeneous chemistry for nitrogen deposition to marine ecosystems: Observations and modeling. *Water, Air, and Soil Pollution: Focus* **1***:* 99-107.

Grimmond CSB, Robeson SM, Schoof JT. 2000. Spatial variability of micro-climatic conditions within a mid-latitude deciduous forest. *Climate Research* **15**, 137-149.

PUBLISHED CONFERENCE PAPERS

Pryor S, Barthelmie R, Schoof J. How coherent is inter-annual variability of wind energy indices over Europe and what are the implications for large scale penetration by wind energy of electricity markets? *Proceedings of European Wind Energy Conference*, London, UK, November 2004 (refereed).

Pryor SC, Barthelmie RJ, Schoof JT. Wind energy prognoses for the Baltic region. *Proceedings for the 4th Study Conference on BALTEX (Baltic Sea Experiment).* Bornholm, Denmark, May 2004.

Schoof JT, Pryor SC. An evaluation of two GCMs: North American teleconnections and synoptic phenomena. *Proceedings of the 15th AMS Symposium on Global Change and Climate Variations*, Seattle, WA, January, 2004.

Pryor SC, Schoof JT, Barthelmie RJ. Near-surface flow regimes: Recent changes and tools for prognoses. *Proceedings of the 15th AMS Symposium on Global Change and Climate Variations*, Seattle, WA, January, 2004.

Barthelmie RJ, Pryor SC, Schoof JT. 2003. Evidence of trends in near-surface wind speeds over the Baltic. *Proceedings of Offshore Wind Energy in the Mediterranean and other European Seas (OWEMES) 2003, Naples, April 2003,* 35-49.

Schoof JT, Robeson SM. 2002. Seasonal and spatial variability of serial and cross-correlation matrices used by stochastic weather generators. *Proceedings of the American Meteorological Society 13th Conference on Applied Climatology*. 4pp.

Pryor S, Barthelmie R, Carreiro M, Davis M, Hartley A, Jensen B, Oliphant A, Randolph J, Schoof J. 2001. Nitrogen deposition to a mid-latitude deciduous forest and ecosystem response. *Proceedings of 2nd International Nitrogen Conference, ESA.*

Pryor SC, Barthelmie RJ, Jensen B, Davis ML, Hirzy KC, Schoof JT, Sorensen LL. 2001. Bidirectionality of ammonia fluxes: Observations over a deciduous forest, *Proceedings of the Sixth International Conference on Air-Surface Exchange of Gases and Particles*, CEH, Edinburgh, 36-41.

Schoof JT, Pryor SC. 2000. Synoptic circulation classification and downscaling for the Midwestern United States. *Proceedings of the American Meteorological Society 15th Conference on Probability and Statistics in the Atmospheric Sciences*, J4-J7.

Pryor SC, Barthelmie RJ Schoof JT, Sørensen LL, Erickson DJ. 2000. Modeling heterogeneous chemistry of nitrogen gases in/on sea spray. *Journal of Aerosol Science Supplement. Proceedings of European Aerosol Conference* 1033-1034.

Grimmond CSB, Robeson SM, Schoof J. 1999. Variability in below-canopy climatic conditions during the growing season within an eastern North American deciduous forest. R.J. de Dear and J.C. Potter (eds) *Proceedings of the International Conference on Biometeorology (ICB)*, Sydney, Australia, November 1999. ICBP03.05. 4pp.

Robeson SM, Grimmond CSB, Schoof J. 1998. Comparison of open-site and below-canopy climatic conditions within an eastern North American deciduous forest. *Proceedings of the American Meteorological Society 23rd Conference on Agricultural and Forest Meteorology*, 184-187.

CONFERENCE PRESENTATIONS (presenter underlined)

Pryor SC, Barthelmie RJ, Schoof JT. How coherent is inter-annual variability of wind indices across Europe and what are the implications for large scale penetration by wind energy of electricity markets? *European Wind Energy Conference*, London, November 2004.

Pryor SC, Barthelmie RJ, Schoof JT. Wind energy prognoses for the Baltic region. *4th Study Conference on BALTEX (Baltic Sea Experiment)*. Gudhjem, Bornholm, Denmark, May 2004.

Pryor SC, Barthelmie RJ, Schoof JT. Historical and prognostic changes in 'a normal wind year': A case study from the Baltic. *European Wind Energy Association Special Topic Conference: The Science of Making Torque From Wind.* Delft University of Technology, The Netherlands, April, 2004.

Schoof JT, Pryor SC. An evaluation of two GCMs: North American teleconnections and synoptic phenomena. *15th AMS Symposium on Global Change and Climate Variations,* Seattle, WA, January, 2004.

Pryor SC, Schoof JT, Barthelmie RJ. Near-surface flow regimes: Recent changes and tools for prognoses. *15th AMS Symposium on Global Change and Climate Variations,*Seattle, WA, January, 2004 (poster).

Pryor SC, Barthelmie RJ, Schoof JT. Observed and predicted flow variability over the Baltic Region: Implications of climate change for wind energy viability. *AGU/EGS Joint Assembly*, Nice, France, April 2003.

Barthelmie RJ, Pryor SC, Schoof JT. Evidence of trends in near-surface wind speeds over the Baltic. *Offshore Wind Energy in the Mediterranean and other European Seas (OWEMES), 2003,* Sicily, April 2003.

Schoof JT, Robeson SM. Seasonal and spatial variability of serial and cross-correlation matrices used by stochastic weather generators. *American Meteorological Society 13th Conference on Applied Climatology*, Portland, OR, May 2002.

Pryor S, Barthelmie R, Carreiro M, Davis M, Hartley A, Jensen B, Oliphant A, Randolph J, Schoof J. Forest canopy uptake of atmospheric nitrogen at a Midwestern US mixed hardwood site and possible implications for carbon storage. *American Geophysical Union,* San Francisco, CA. December 2001.

Pryor SC, Barthemlie R, Carreiro M, Davis M, Hartley A, Jensen B, Oliphant A, Randolph J, Schoof J. Nitrogen deposition to a mid-Latitude deciduous forest and ecosystem response. *2nd International Nitrogen Conference*. Washington, D.C., October 2001

Schoof JT A comparison of two synoptic circulation classifications for the Midwestern United States. *Annual Meeting of the Association of American Geographers*, New York City, NY, March 2001.

Pryor SC, Barthelmie RJ, Davis ML, Schoof JT, Hirzy KC, Hartley A, Carreiro M, Jensen B. Nitrogen deposition to and cycling in a forest ecosystem. *Annual Meeting of the Association of American Geographers*, New York City, NY, March 2001.

Pryor SC, Barthelmie RJ, Schoof J, Erickson D. Modeling heterogeneous chemistry on sea spray: implications for nitrogen deposition. *American Association for Aerosol Research Annual Conference*, St. Louis, MO, November 2000.

Pryor SC, Barthelmie RJ, Jensen B, Hirzy K, Schoof J, Davis M. An investigation of the role of particles in observations of the bidirectionality of ammonia fluxes. *American Association for Aerosol Research Annual Conference*, St. Louis, MO, November 2000 (poster).

Pryor SC, Barthelmie RJ, Schoof J, Erickson D. Modeling heterogeneous chemistry in sea salt droplets. *European Aerosol Conference*, Dublin, Ireland, September 2000. (invited)

Pryor SC, Barthelmie RJ, Schoof JT, Sorensen LL, Erickson III DJ. Implications of heterogeneous chemistry for nitrogen deposition to marine ecosystems: Observations and modeling, *Sixth International Conference on Air-Surface Exchange of Gases and Particles*, Edinburgh, UK, July 2000.

Schoof JT, Pryor SC. Synoptic circulation classification and downscaling for the Midwestern United States, *American Meteorological Society 15th Conference on Probability and Statistics in the Atmospheric Sciences*, Asheville, NC, May 2000.

Schoof JT, Pryor SC. Synoptic circulation classification and statistical downscaling for the midwestern United States, *American Geophysical Union Spring Meeting*, Boston, MA, June 1999 (poster).

Grimmond S, Zutter H, Potter S, Schoof J, Souch C. Evaluation and application of automated methods for measuring sky view factors in urban areas, *International Conference on Urban Climate*, 1999, Sydney Australia, November 1999. (poster).

Grimmond CSB, Robeson SM, Schoof J. Variability in below-canopy climatic conditions during the growing season within an eastern North American deciduous forest. *International Conference on Biometeorology (ICB)*, Sydney, Australia, November 1999.

Robeson SM, Grimmond CSB, Schoof J. Comparison of open-site and below-canopy climatic conditions within an eastern North American deciduous forest. *American Meteorological Society 23rd Conference on Agricultural and Forest Meteorology*, Albuquerque, NM, November 1998.

INVITED PRESENTATIONS

Schoof JT. 21st Century Climate Projections. Indiana University Dean's Advisory Board, September 2004.

Schoof JT. Regional climate change. E162 Environment and People, School of Public and Environmental Affairs, Indiana University, May 2004.

Schoof JT. 20th century climate variability. G475/575 Climate Change, Geography, Indiana University, October 2003.

Schoof JT. Mid-latitude cyclones. G304 Physical Meteorology and Climatology, Geography, Indiana University, October 2003.

Schoof JT. Thunderstorms. G304 Physical Meteorology and Climatology, Geography, Indiana University, April 2003.

Schoof JT. Regional climate change implications for the Midwestern United States. E162 Environment and People, School of Public and Environmental Affairs, Indiana University, February 2002.

Schoof JT. Synoptic climatological classification. G433 Synoptic Meteorology and Climatology, Geography, Indiana University, March 2001.

Schoof JT. Regional climate change implications for the Midwestern United States. E162 Environment and People, School of Public and Environmental Affairs, Indiana University, February 2001.

MEMBERSHIP IN PROFESSIONAL SOCIETIES

American Geophysical Union (AGU)
American Meteorological Society (AMS)
Association of American Geographers (AAG)