January 2015

Statistical Downscaling For The Northern Great Plains: A Comparison Of Bias Correction And Redundancy Analysis

Jacob Coburn

STATISTICAL DOWNSCALING FOR THE NORTHERN GREAT PLAINS: A COMPARISON OF BIAS CORRECTION AND REDUNDANCY ANALYSIS

by

Jacob Jimmie Coburn
Bachelor of Science, University of Montana, 2012

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota
August
2015

This thesis, submitted by Jacob Jimmie Coburn in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

_____
Andrei Kirilenko

_____
Xiaodong Zhang

_____
Gretchen Mullendore

_____

_____

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

_____
Wayne Swisher
Dean of the School of Graduate Studies

_____
Date

PERMISSION

Title Statistical Downscaling for the Northern Great Plains: A Comparison of Bias Correction and Redundancy Analysis

Department Earth System Science and Policy

Degree Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Jacob Jimmie Coburn
July 1, 2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisory committee, Dr. Andrei Kirilenko, Dr. Xiaodong Zhang and Dr. Gretchen Mullendore for their continuous support and patience through this long project. I would also like to thank my friends and family, who dealt with my incessant worrying. Last but certainly not least, Nicholas, who always had a way to make it better. Thank you all!

ABSTRACT

The climate of the Earth is changing, and is primarily a result of our rampant industrialization over the past two centuries. These changes have manifested themselves in many ways over the whole of the Earth's surface and sub-systems, leading to the need to understand the changes and predict future outcomes. Coupled climate and general circulation - Earth system models (GCMs) allow for the analysis of dynamically active simulations over the whole of the planet, yet are limited by computational power. The model grids are coarse by design to perform within these computational constraints, which enables them to function and provide information at continental and larger scales, but which limit their ability to offer information for regional and local environments. Dynamical models created with higher resolutions allow for regional climate modeling yet are also limited by computational constraints and require detailed information to run. Statistical downscaling seeks to bridge the gap between coarse GCM grids by utilizing observational data and statistical models to remove the biases from the data at the local level. There have been several types of statistical methods applied to this task over many different regions with some success. The goal of this study is to utilize two methods in particular, bias-corrected spatial disaggregation (BCSD) and redundancy analysis (RDA), to downscale maximum and minimum temperature, as well as precipitation, for the Northern Great Plains (NGP) region. These methods are calibrated over the period 1950 – 1970 using a 1/8 degree gridded dataset for 17 GCMs, then applied to a verification period (1970 – 1999) and compared to observations over that period to assess the downscaled models skill in capturing local NGP variability. These methods are also applied to future model runs forced via the representative concentration pathways (RCPs) low end (2.6), median (4.5) and high end (8.5) 21st Century forcings, which provides possible outlooks for local stakeholders over the coming decades. It is found that BCSD does well in downscaling temperature and precipitation, as well as their various metrics. RDA provides more mixed success, with good skill demonstrated for temperatures but a strong wet bias in precipitation. It is noted, however, that RDA yielded better correlations to the observations. Future scenarios show broad ranges of projected outcomes that, as expected, increase with increasing forcing, though temperature shows stronger changes than precipitation, and BCSD exhibits higher sensitivity than RDA. Future research may help further constrain the results of these downscaling methods, particularly RDA, by adopting further bias correction to the results.

CHAPTER 1 – FROM GLOBAL TO LOCAL: DOWNSCALING

1.1 INTRODUCTION

It is now well accepted within the scientific community that the Earth's climate is changing, primarily as a result of industrial emissions of greenhouse gases and aerosols over the past half century as well as anthropogenic landscape modification (Bindoff et al. 2013). Various lines of evidence for these climate changes are manifesting themselves worldwide through increasing weather extremes, shrinking ice coverage, rising seas and shifting biogeophysical systems (EPA 2014; IPCC 2014). As global efforts to curb greenhouse gas emissions have fallen flat and atmospheric CO_2 concentrations have already risen by 5.8 percent above estimates made only a few years ago in the wake of the 2008 financial crisis, the effects are projected to worsen as the high emission, business-as-usual scenarios continue to play out (Peters et al. 2012). Our knowledge on how our industrialization is impacting the climate system has coevolved with the development of dozens of Global Climate Models (GCMs) over several decades which aid in the refinement of climatological science as well as investigating the changes that are underway. These models run in a similar fashion to their meteorological counterparts used for weather prediction, though they incorporate more information from other parts of the Earth system such as coupled atmospheric, oceanic and landscape interactions as well as sub-models for ice masses and biological systems. The result of the inclusion of all of these processes is a set of highly complex mathematical programs that must cover hundreds of years of model time integrated at time steps on the order of hours or shorter, all of which strain even the largest super-computers just to produce one model run, let alone several. In order to obtain results in a reasonable

amount of time, models are built at relatively coarse spatial resolutions, which average out small scale variability and necessarily leave local processes dynamically unresolved (McGuffie and Sellers 2005). Thus while model outputs are useful to diagnosing and predicting the changes in the Earth system at continental or larger scales, they have been and remain unable to provide information to the local stakeholder about what can be expected at that particular location under changing climate conditions. While increasing computational power is slowly leading to the ideal of fine-resolution global climate models, present efforts exist to more quickly bridge the gap between the coarse GCM outputs and the local scale information needed by stakeholders. This study seeks to further investigate those efforts, generally known as climate model downscaling.

1.1.2 DOWNSCALING METHODOLOGIES

Downscaling, as it is broadly defined, is the process by which coarse scale climate model data is converted to a finer resolution. The two main types of downscaling, dynamical and statistical, have very different approaches in deriving local scale climate data, and are compared in terms of their potential costs and benefits in Table 1.1. Dynamic downscaling involves the use of regional climate models (RCMs), which resolve processes on much finer grids than GCMs. They are also often referred to in the literature as limited area models and often are built to simulate specific features of the broader landscape, ranging from a whole continent down to one river basin (Wang et al. 2004). While allowing for dynamically produced information at such scales, they may include biases and require more complex information than statistical downscaling to run properly, including more advanced dynamic equations, regional forcing data and boundary conditions. While not the focus of this study, RCMs are discussed further in section 1.2 for the sake of comparison. The main focus of this work is statistical downscaling, which is the process of finding robust and efficient statistical connections between the coarse

model data and observational data that are then used to correct the biases in model output at the local level (Benestad 2008). There are many different statistical methods that have been produced, which are generally classified by their use of regression or bias correction techniques (linear methods – section 1.3.1), weather typing (nonlinear methods – section 1.3.3), or stochastic resampling (weather generators – section 1.3.4), as shown in table format in Table 1.2. This work will focus on two linear methods, Bias Corrected Spatial Disaggregation (BCSD) and Redundancy Analysis (RDA), as applied to a specific study region, the Northern Great Plains (NGP).

Table 1.1: A table comparing downscaling methodologies from Trzaska and Schnarr (2014).

	Dynamical downscaling	Statistical downscaling
Provides	• 20–50 km grid cell information • Information at sites with no observational data • Daily time-series • Monthly time-series • Scenarios for extreme events	• Any scale, down to station-level information • Daily time-series (only some methods) • Monthly time-series • Scenarios for extreme events (only some methods) • Scenarios for any consistently observed variable
Requires	• High computational resources and expertise • High volume of data inputs • Reliable GCM simulations	• Medium/low computational resources • Medium/low volume of data inputs • Sufficient amount of good quality observational data • Reliable GCM simulations
Advantages	• Based on consistent, physical mechanism • Resolves atmospheric and surface processes occurring at sub-GCM grid scale • Not constrained by historical record so that novel scenarios can be simulated • Experiments involving an ensemble of RCMs are becoming available for uncertainty analysis	• Computationally inexpensive and efficient, which allows for many different emissions scenarios and GCM pairings • Methods range from simple to elaborate and are flexible enough to tailor for specific purposes • The same method can be applied across regions or the entire globe, which facilitates comparisons across different case studies • Relies on the observed climate as a basis for driving future projections • Can provide point-scale climatic variables for GCM-scale output • Tools are freely available and easy to implement and interpret; some methods can capture extreme events
Disadvantages	• Computationally intensive • Due to computational demands, RCMs are typically driven by only one or two GMC/emission scenario simulations • Limited number of RCMs available and no model results for many parts of the globe • May require further downscaling and bias correction of RCM outputs • Results depend on RCM assumptions; different RCMs will give different results • Affected by bias of driving GCM	• High quality observed data may be unavailable for many areas or variables • Assumes that relationships between large and local-scale processes will remain the same in the future (stationarity assumptions) • The simplest methods may only provide projections at a monthly resolution
Applications	• Country or regional level (e.g., European Union) assessments with significant government support and resources • Future planning by government agencies across multiple sectors • Impact studies that involve various geographic areas	• Weather generators in widespread use for crop-yield, water, and other natural resource modeling and management • Delta or change factor method can be applied for most adaptation activities

Table 1.2: A table comparing statistical downscaling methods. Note the somewhat different classes of downscaling, with weather typing referring to non-linear techniques and regression methods referring to linear models. From Wilby et al. (2004).

Method	Strengths	Weaknesses
Weather typing (e.g. analogue method, hybrid approaches, fuzzy classification, self organizing maps, Monte Carlo methods).	• Yields physically interpretable linkages to surface climate • Versatile (e.g., can be applied to surface climate, air quality, flooding, erosion, etc.) • Compositing for analysis of extreme events	• Requires additional task of weather classification • Circulation-based schemes can be insensitive to future climate forcing • May not capture intra-type variations in surface climate
Weather generators (e.g. Markov chains, stochastic models, spell length methods, storm arrival times, mixture modelling).	• Production of large ensembles for uncertainty analysis or long simulations for extremes • Spatial interpolation of model parameters using landscape • Can generate sub-daily information	• Arbitrary adjustment of parameters for future climate • Unanticipated effects to secondary variables of changing precipitation parameters
Regression methods (e.g. linear regression, neural networks, canonical correlation analysis, kriging).	• Relatively straightforward to apply • Employs full range of available predictor variables • 'Off-the-shelf' solutions and software available	• Poor representation of observed variance • May assume linearity and/or normality of data • Poor representation of extreme events

1.2 GLOBAL MODELS AND DYNAMIC DOWNSCALING

1.2.1 GENERAL CIRCULATION MODELS

The development of climate models was not a linear process of increasing complexification. Different types of models have developed in different periods which are suited to specialized purposes (Edwards 2011; Weart 2008). Simple 0, 1 and 2 dimensional energy balance models forego much in the way of details, however they are useful for understanding the primary processes governing the climate and can be easily applied to different scenarios and long time periods. These simple models have been used since the 1970s to study the Snowball Earth events some 600 Ma, as well as the problem of latitudinal temperature gradients during hothouse periods such as the Cretaceous. Radiative-convective models (RCs) allow researchers to study the interplay between insolation, greenhouse heating and convection throughout the atmospheric column, and can be coupled to simple oceanic box models to account for the thermal lag of the system. These have been used to gauge the importance of greenhouse gases, clouds and stratospheric ozone in the temperature profile of the atmosphere. When two-dimensional movements of air are calculated as part of the RC model, simple circulation systems can be

developed and studied. Simplified versions of the GCMs have been developed which are known as Earth system models of intermediate complexity (EMICs), such as the Bern3D model that has been used to study coupled climate-CO_2 variations throughout the Phanerozoic (Eby et al. 2013). Where they sacrifice in terms of detail, they make up for by having extended temporal applicability similar to the RC models.

Three dimensional models have been in development since the 1950s. The advent of computers made the prospect for such a model much more feasible than it had been when Vilhelm Bjerknes first worked out how the primitive equations could be used to numerically simulate atmospheric movements in the 1920s (Edwards 2011). Global climate models are, in fact, the amalgamation of several sub-models that utilize input from one another and provide output back into the other components. Atmospheric general circulation models (AGCMs) run in tandem with oceanic general circulation models (OGCMs) and are coupled at the air-sea exchange (AOGCMs), though they are often run using different grids. Sub-models for lithospheric, cryospheric and biospheric interactions are also run using an input exchange. As newer models have been developed and older ones refined, further improvements in existing processes and the inclusion of new ones became feasible. Modern GCMs include biogeochemical models that mimic various elements of the carbon and other nutrient cycles, and many are now being equipped to dynamically capture chemical reactions within the atmosphere such as those involved in aerosols, clouds and ozone, prompting some to refer to these as Earth system models (Flato et al. 2013). Some models break the atmosphere and oceans into a three dimensional set of grid boxes that process the dynamic equations through the use of finite differencing techniques in which the set of governing equations is iteratively solved. Other models employ spectral methods within the atmospheric and oceanic components, with the

results transferred into the traditional Cartesian grid at each time step. These spectral methods offer more efficient representation of atmospheric waves that generate much of the variability in the atmospheric and oceanic systems (Kaper and Engler 2013). Each of the model components is then set with initial values and boundary conditions, which are then run for a long period of model time without imposing forcings to allow the model to come to equilibrium between the components, known as spin-up time. This process can often show biases in the model before any changes are imposed, such as the problem some models faced initially in which the oceans would warm or cool over time without external forcing driving an energy imbalance (Collins et al. 2011).

Global climate models have shown skill in mimicking large scale aspects of the climate system as a whole, both in the stationary statistics of particular periods and in the changes observed in both instrumental and paleoclimatic records (Knutti and Sedlácek 2013; Shukla et al. 2006). Confidence in the results is further bolstered when multiple models from differing, independent groups generally agree on some aspect of those changes, such as the polar amplification of the warming over the 20th century (Crook et al. 2011). Other variables and systems are not so well constrained by the GCMs, including precipitation and clouds, which could relate to the issues with the parameterizations as they are much more heavily affected by generalizations of small scale processes, though improvements have been noted in each successive IPCC general analysis (Flato et al. 2013; Solomon et al. 2007). Other internal systems, such as the atmospheric-oceanic oscillations and low frequency variability have also been improving in time. However, the details (such as peak dates, seasonal locking and interactions with other Earth system components) leave much to be desired, as do the teleconnections between regions (Sheffield et al. 2013). North American mean patterns are

largely captured, however particular characteristics remain elusive. Precipitation extremes are underestimated and differences between wet and dry areas tend to be underestimated. Daily temperature ranges are also not captured, nor is the apparent century-long cooling in summer temperatures in the American Southeast (the so-called warming hole). Other problems persist as well, such as the double intertropical convergence zone (ITCZ) where too much precipitation is produced in the Tropics, and the under-predicted impacts of global warming on sea ice, with larger losses observed in the Arctic than expected and a small increase observed in the Antarctic (Flato et al. 2013; Li and Xie 2013). While some have commented on the supposed recent slowdown in global surface temperature warming and the apparent lack of ability of the ensemble model mean to capture this "hiatus", there have been numerous studies (Cowtan and Way 2014; Bindoff et al. 2013; Fyfe et al. 2013; Kosaka and Xie 2013) that have concluded that the recent slowdown is really the result of some combination of internal variability (La Niña-type patterns), previously unaccounted for negative radiative forcings (tropospheric aerosols from volcanoes and reduced solar activity specifically) and surface station coverage biases within the Arctic acting on top of the upward temperature trends stemming from greenhouse gas forcing.

1.2.2 DYNAMIC DOWNSCALING

Regional climate models (RCMs) attempt to mimic the same processes as the GCMs but at within much higher resolution grids than the global models. These regional models are often nested within the grids of larger, parent models through a series of staggered-resolution grids bridging the coarse grid to the fine grid in steps. This allows the GCM to create the boundary conditions for the RCM in a stable manner that direct transmission between the model grids impedes (Rohli and Vega 2008). Previously, this interaction was one-way, with the RCM run evolving on its own without further influencing the GCM, though two-way nesting has become available. Two-way nesting is not used in downscaling as of this writing because the

computational limits of running both the GCM and RCM would severely limit the amount of detail that could be included in the RCM, as well as complicate regional to global interactions with regard to feedbacks (Laprise et al. 2012). The resolutions of the models range from 10 to 100 kilometers, and often depend upon what tasks the model is intended to serve. This means that certain variables, such as rough terrain, detailed vegetation, or human factors like pollution and land use changes must be more fully resolved where they could previously be parameterized, contributing to the much higher complexity that is inherent in the regional modeling process. In fact it is at these levels that mesoscale processes and regionalized forcings start to become important (Wang, et al. 2004). As with global models, regional model responses to known forcing events (volcanic eruptions, large ENSO excursions) are tested for fidelity to observations. While RCMs can be applied to multiple regions with relative ease, it requires that the boundary conditions be changed accordingly and, in some cases, an optimization of the parameterizations used (McGuffie and Sellers 2005). Even in the early stages of regional model development, RCMs showed promise in downscaling. In one such study, a model with a 60 by 60 kilometer grid was nested within the grids of two GCMs, one with a resolution of 4.5 by 7.5 degrees, the other 3 by 3 degrees, for the complex terrain of the Intermountain West (Giorgi 1990). The models were constrained by using only cold season forcings for the Northern Hemisphere, thus keeping the temporal span of interest to the January climatology of the region. The RCM showed a strong ability to improve the depiction of winter snowpacks and cold season precipitation. Other studies have looked at temperature, precipitation, and wind speeds over regions such as Australia, while others have discussed the prospect of dynamic downscaling over the poles (Feser et al. 2011).

Several multi-model RCM projects have been undertaken to explore their usefulness toward increasing understanding of climate changes for specific regions, particularly North America and Europe. The North American Regional Climate Change Assessment Program (NARCCAP), which started in 2006, incorporates six regional models and three boundary-forcing GCMs for an area encompassing the Contiguous United States, Canada, Mexico and parts of Alaska (Mearns et al. 2013). A fourth parent model is the Department of Energy Global Reanalysis (NCEP 2). All six RCMs output data at a resolution of 0.5° by 0.5°, though with different physical formulations and parameterizations. Two of them, the Canadian RCM and the Experimental Climate Prediction Center Regional Spectral Model (ECPC-RSM), use spectral nudging, a process in which the parent GCMs not only force the boundaries of the RCMs, but also influence the inner domain of the model. The results of the NARCCAP project show improvements over the global models, though with mixed results between the RCMs. 2-meter air temperature trends over the observational period (up to 2010) are well represented for Spring, though results are less clear for the other seasons, and sub-regional variations such as the warming hole in the Southeast mentioned earlier are still not well constrained, though the two models using GCM nudging did a better job than the rest in capturing the sub-regional cooling (Bukovsky 2012).

Sub-regional skill comparisons show that certain areas within the North American domain are better represented. When analysis of variance (ANOVA) is used to partition the contribution to sub-regional variability, GCMs captured more variability in coastal regions while inland areas and those featuring complex terrain were better captured by the RCMs (Mearns et al. 2013). The NGP, as defined by Mearns et al. (2013) was also better constrained by the RCMs, though with large uncertainties associated with both GCM input and natural variability. Fire danger, as measured by the Haines Index, has been investigated using the NARCCAP ensemble (Luo et al.

2014). By compiling model output for tropospheric stability and dryness, the RCMs showed a marked increase in fire potential throughout the United States, though model results varied in which regions would experience the largest increases in risk. Changes in precipitation are better captured by the RCMs in multiple studies, as expected by the finer resolution and the ability to better replicate convective precipitation, which is important during warmer months (Bukovsky and Karoly 2014; Kendon et al. 2012; Rummukainen 2010). These improvements, while still apparent, are not as clear for the large scale precipitation associated with cold season frontal systems. The same holds true for temperature extremes, which show mixed improvement when considered from the perspective of large scale meteorological forcings (Loikith et al. 2015).

1.3 STATISTICAL DOWNSCALING

The problem of translating the coarse scale climate model output data into finer scale data useful to local interests and impact studies has a much less computationally expensive counterpart to RCMs: empirical statistical downscaling (ESD), or just statistical downscaling (SD). Empirical statistical downscaling simplifies the downscaling process by using comparisons between observational data and model data to create mathematical relationships between them. Observational data are referred to as predictands, that is, the data that the model seeks to more closely match via the statistical method employed. The variable or set of variables used to create that model, usually from a GCM, are the predictors. These models, also called transfer functions, can relate large scale GCM data to a range of more localized grids, be they large grids of continuous data or collections of individual station time series, and are of the form:

$$Y = f(X) \qquad \text{Equation 1.1}$$

where Y is the predictand and f(X) is a function of the predictor. Once these functions are established via some comparison between observed data over some time frame and model data for the same time domain (the "training" or "calibration" period), they can be tested by using

those functions to recreate other known periods with observations, called validation or verification periods (Benestad 2008). Another method is to leave specific data points within the calibration period out of the model and then compare their downscaled values to observations. They can also be applied, with reservations, to periods without any means of verification, allowing for extrapolation at the local scale from GCM projections.

Data used in the process of statistical downscaling come from GCM outputs of specific weather variables and station-based or reprocessed spatial data of observed variables. The selection of predictands is dependent on the intended uses of the model, as certain variables lend themselves to certain applications, such as minimum temperature being used to gauge frost days. Predictors can be the GCM equivalents of the observed variables, though they can also be other variables not contained in the set of observations, which are used to infer the values of the predictands through the downscaling model. These variables are assumed to have a strong physical or statistical connection to the observed variable of interest (such as linking geopotential heights to precipitation), that they are modeled realistically at the course resolution of the GCM, and that their relationship to the predictand remains the same even in a changed climatic state, known as stationarity. This stationarity between the predictors and predictands is imperative to the applicability of the predictors to periods outside the calibration period, with different ramifications depending on the type of statistical method employed.

1.3.1 LINEAR METHODS I: BIAS CORRECTION

One form of statistical downscaling is bias correction, known as bias corrected spatial disaggregation, or BCSD (Maurer and Hidalgo 2008). BCSD entails the use of the observed distribution of the predictand to correct the distribution of the modeled predictor. The advantage of this method its simplicity of implementation, however it relies on the assumption that the climate model adequately represents all the necessary processes and interactions that work to

produce the predictand in question and that it does so with a consistent, or stationary bias that is able to be removed from the data. BCSD utilizes quantile mapping to transform the model data distribution to more closely fit that of the observed distribution and then return the new model values corrected for the distribution shift. Cumulative distribution functions (cdf) are constructed for both observed and modeled data, with the model data then being shifted to match the observed data over said quantiles. The quantile to quantile relationship can be estimated in a number of ways, be it a comparison of empirical quantiles, through the transform of a theoretical distribution, or through a regression of quantiles onto one another.

For this study, empirical cumulative distributions (ecdf) were chosen so as to allow the model data to conform as closely as possible to the observed data without the downscaled data becoming biased by being forced to conform to a parametric distribution, which is useful if the data come from a distribution that has no easily defined parametric function, such as precipitation. This process is shown by:

$$P_{t,i} = ecdf_{mod}(X_{t,i}^{Cal}) \qquad \text{Equation 1.2,}$$

$$CF_{t,i} = ecdf_{obs}^{-1}(P_{t,i}) - ecdf_{mod}^{-1}(P_{t,i}) \qquad \text{Equation 1.3,}$$

$$Y_{t,i}^{Val} = X_{t,i}^{Val} + CF_{t,i} \qquad \text{Equation 1.4,}$$

$$\Delta P(0) = [ecdf_{mod}(0) - ecdf_{obs}(0)] / ecdf_{mod}(0) \qquad \text{Equation 1.5.}$$

The ecdf of the predictor data ($X_{t,i}^{Cal}$) over the calibration period is calculated in Equation 1.2. This yields the cumulative probabilities (P) of each data point from lowest to highest for each location (i) over each day in the calibration period (t) for the GCM data. The probabilities are calculated in evenly spaced intervals between 0 and 1 and interpolated for points between the calculated probabilities. The data points for each of the probabilities P is then returned from both the predictor and observed data through the inverse ecdf, with the correction or bias factor,

CF, being calculated by subtracting the GCM data from the observed data for each probability in the distribution for each location over the calibration period (Equation 1.3). The correction factors are assumed constant for all time periods and are applied to the distributions of predictor data outside the calibration interval, ($X_{t,i}^{Val}$) , as in Equation 1.4, to yield the validation period data $Y_{t,i}^{Val}$ (Ahmed et al. 2013; Themebl et al. 2011; Wilks 2011). Some variables, such as precipitation, exhibit stepwise behavior in the real world in which it is either precipitating or not, with precipitating days varying in rate. As noted in Section 1.2.1, frequencies of precipitating days are often missed by climate models, which tend to produce too much light precipitation (drizzle) and too little extreme precipitation (Gutowski et al. 2003). The bias correction method outlined in Equations 1.2 through 1.4 above may not fully correct the frequency of precipitating events, even if the overall rates of precipitation match well with observations. Equation 1.5 gives the ratio of the frequency of dry days between the observed and modeled distributions, where the modeled distribution is then modified to bring the ratio closer to 1 through a wet day correction.

Some studies have looked into the relationship between GCM variables and their observational counterparts, such as Schmidili et al. (2006). This study focused on the bias corrected downscaling of precipitation over the European Alps. In this case, the use of the model output for precipitation was used as the predictor. A further bias correction was applied in that the scaling factor for the precipitation was not held constant, but allowed to vary via a numerical indicator of broad scale mid-tropospheric flow, taking a rudimentary account of the type of pattern at each time step analogous to the non-linear methods shown in Section 1.3.3. By comparing 34 years of reanalysis precipitation data, they were able to correct both the frequency and intensity of precipitation, including at the highest intensities. The enhanced bias correction

using a simple circulation index did show extra skill over that of the simple bias correction. Other studies have applied bias correction to temperature, precipitation and humidity at both daily and monthly time scales and found that these downscaling methods resulted in value added over raw GCM data for multiple regions, and that they have comparable skill to regression and non-linear downscaling strategies overall (Abatzoglou and Brown 2012; Ahmed et al. 2013; Themeßl et al. 2011).

1.3.2 LINEAR METHODS II: REGRESSION ANALYSIS

Linear regression techniques involve finding numerical functions relating the fine scale data and the coarse data, be it individual variables or fields of data. This class of methods has the benefit of being easily extrapolated, though with the assumption that the function relating the predictor and predictand holds in a changed climate. Regression methods range from relatively simple multi linear regression (MLR) and multivariate regression (MVR) models to complex methods involving canonical correlation (CA), principal components analysis (PCA), or some other form of dimensionality reduction, in which large scale spatio-temporal patterns are used as the predictands in the downscaling models. While advanced methods of regression analysis try to account for as much variability within the captured patterns (constrained variability), there is always some component of the data that is left over (unconstrained variability). Regression models must overcome the inherent reduction in variability that occurs due to the unconstrained portion of data variance, with some studies looking into post-processing variance inflation (Storch 1999). It was not recommended, however, due to its implicit assumption that the unaccounted for variance was a result of the coarse scale dynamics guiding the downscaling model and not a result of local processes, as well as that the added inflation detracted from the skill of the model. It was recommended that including a stochastic term to generate "noise"

values based on the spread of the unconstrained variability would work to yield better results, though few studies make use of such a term.

Most studies focus on complex methods that are more capable of capturing the higher variability of regional climates, though some have used simple regression models as a means of comparing them to the complex methods. One such study looked into the monthly precipitation patterns for the Lake Karla watershed in Thessaly, Greece (Vasiliades et al. 2009). They utilized sea level pressure, precipitation, surface temperatures, geopotential and wind speeds and regressed them against precipitation data to attain the regression model, and used the downscaled precipitation to calculate a standardized drought index (SDI) for an assessment of the drought conditions over the region. It was found that the simple regression method worked well for monthly time scales, though the authors cautioned the use of such methods on daily precipitation, in which more complex methods would be needed to fully capture the non-Gaussian distribution of wet day intensities. This conclusion was echoed in Themeßl et al, (2011), where simple regression techniques were applied to daily precipitation and failed to significantly improve the output of the GCMs analyzed, though correlations between regression-based models and observations were greater than those from bias corrected models.

1.3.2a CANONICAL CORRELATION ANALYSIS (CCA)

One of the most common linear techniques used in downscaling is canonical correlation analysis, or CCA (Wilby et al, 2004). This process involves relating local scale variables to a broad field of course scale data with the largest possible temporal correlation. CCA seeks to extract useful model estimates of observed variables by finding those combinations of independent and dependent variables that result in the highest correlations, usually by means of a reduction in the dimensions of the data. This allows for the simultaneous comparison of many different locations at different times, and can be implemented via raw observational and model

values or by combinations of their principle components (called empirical orthogonal functions (EOFs) in geophysical research). The dimensions of the data here refer to the spatial and temporal variability, which CCA reduces by finding those spatio-temporal patterns (the EOFs) that account for the maximum amount of the variations in the data, with subsequent orthogonal patterns accounting for further portions of the variance. Values of the downscaled model data may be obtained from a direct canonical correlation analysis of the variable in question (such as observed tmax to modeled tmax), or they can be constructed using other model parameters which act as explanatory predictor variables, such as atmospheric temperature at height or wind speeds.

CCA has been applied to various problems ranging from precipitation in Sweden (Busuioc et al. 2000) and the Iberian Peninsula (Storch 1992) and in South Africa during the Austral Winter (Landman et al. 2000), as well as Central European temperatures (Huth et al. 2003; Huth 1999) and sea level pressures over the Baltic Sea (Heyen et al. 1996). In all cases, multiple variables were combined from the models to produce the fields needed to compare to the local variable. Surface temperature is often calculated from surface energy balance data but was found to be better represented by temperatures at the 850 hPa level than conditions than those extracted from combinations of heating terms and wind patterns (Huth 1999). Precipitation was often compared to geopotential thicknesses and moisture variables, which often yielded good results in downscaling. The investigation into Iberian precipitation also yielded the interesting result that the GCM used did not match the local scale results when simpler regression methods were applied and compared, indicating the sensitivity of the results to the methods and predictors used.

A general sense of the computational process can be seen in the equations describing the "brute force" method where no dimensional reduction takes place (Benestad et al. 2008). The

first step is to find the interrelated patterns between X and Y. The covariance matrices C_{YY}, C_{XX}, and C_{YX} are calculated from centered data matrices X and Y:

$$C_{YY} = Y^{Cal}Y^{[Cal]T} \quad \text{Equation 1.6,}$$

$$C_{XX} = X^{Cal}X^{[Cal]T} \quad \text{Equation 1.7,}$$

$$C_{YX} = Y^{Cal}X^{[Cal]T} \quad \text{Equation 1.8.}$$

Multiplying the inverses of the C_{YY} and C_{XX} matrices by the C_{YX} matrix yields the normalized covariance matrix, C, or the interrelated patterns between X and Y, shown in Equation 1.9. Performing a singular value decomposition on C results in the right side of Equation 1.10. The left matrix (L) and right matrix (R) are the rotation matrices containing the eigenvectors of the covariance matrix, while M is the canonical correlation matrix, with the correlations in the diagonal from greatest to least. Multiplying L by the identity matrix of the predictors ($C_{XX}C_{XX}^{-1}$) yields the map, H, of highest spatial correlation for X and multiplying R by the identity matrix of the predictand ($C_{YY}C_{YY}^{-1}$) yields the map, G, of highest spatial correlation for Y, as shown in Equation 1.11 and 1.12. The canonical variates, U and V, are calculated according to Equations 1.13 and 1.14, and contain, in descending order of importance, the time series that best capture the variability of the data:

$$C = C_{YY}^{-1} C_{YX} C_{XX}^{-1} \quad \text{Equation 1.9,}$$

$$C = LMR^{T} \quad \text{Equation 1.10,}$$

$$G = C_{YY}C_{YY}^{-1}R \quad \text{Equation 1.11,}$$

$$H = C_{XX}C_{XX}^{-1}L \quad \text{Equation 1.12,}$$

$$U = C_{YY}^{-1}LY \quad \text{Equation 1.13,}$$

$$V = C_{XX}^{-1}RX \quad \text{Equation 1.14.}$$

Finally, new output data can be calculated by applying the prediction matrix to the new predictor data. Relating $X = HV^T$ and $Y = GU^T$ and using Equations 1.9 through 1.14, Equation 1.15 results, where $\Psi = GM(H^TH)^{-1}H^T$:

$$Y^{Val} = \Psi X^{Val} \qquad \text{Equation 1.15,}$$

The set of equations describe the process of canonical correlation analysis, however they are of limited practical use for gridded climate datasets. This stems from two issues; the problem of collinearity of the locations in the matrices and the sparseness of the centered data, which inhibits the inversion of the matrices in Equation 1.9. One possible solution is to use empirical orthogonal functions (EOFs) which are the spatial-temporal equivalents of principle components (Benestad 2008). This process would reduce the number of dimensions of the datasets and thus reduce the collinearity and the overall sparseness of the data, and has been deployed in climate and Earth system science applications, such as downscaling for wind energy generation in the Gaspé region of Quebec, Canada (Yosvani et al, 2015) or the downscaling precipitation from the East Asian Monsoon (Simon et al, 2013). Yosvani et al, (2015) used EOF analysis on reanalysis data from the NCEP-NCAR group to create regional time series and spatial maps of wind patterns for the Gaspé region. Simon et al, (2013) use EOF patterns of relative vorticity at 850 hPa and vertical velocity at 500 hPa to capture rainfall patterns for various stations throughout the Poyang Lake catchment within the Yangtze River system. Both studies, as well as others, concluded that the reduction in dimensions of the predictors aided in making the downscaling process less computationally demanding. This method is more limited, however, when the predictor and predictand are the same variable. Another means by which to accomplish this task, and the tactic employed in this work, is through the use of redundancy analysis.

1.3.2b REDUNDANCY ANALYSIS (RDA)

Redundancy analysis is very similar to CCA, though it differs in one assumption (Legendre and Legendre 2012). CCA seeks out those concurrent patterns of highest correlation between the GCM and observational datasets. Redundancy analysis does not seek out such a relationship, but rather assumes such a relationship between the independent and dependent variables. CCA makes no distinction between the datasets in terms of causation, however RDA does. In doing so, this method acts as a mid-point between multiple linear regression and canonical correlation, with the regressions being modified by the eigenvalues from the canonical correlations. This helps avoid the problem of inversion by only needing the inverse of the independent predictor matrix. It was used as part of a broader CCA process for temperatures over Turkey and to reconstruct wave heights for the North Atlantic from wind forcing data from GCMs (Tatli et al. 2004; WASA 1998). RDA was developed over time from the same basic formulae of CCA, but brought into its own theoretical grounding by the late 1970s, having been applied to various problems in the mathematical ecology (Tyler 1982).

The loadings of the multiple regression model B are calculated according to Equation 1.17, where each specific loading is a constant based upon the regression of each observed data location against every GCM location over the calibration period, which is essentially multivariate regression. Equation 1.18 gives the fitted responses to the predictors X and the loadings B. The covariance matrix of the fitted values C_{YY} is calculated according to Equation 1.19, where C_{YX} and C_{XX} correspond to the covariance matrices in Equation 1.7 and Equation 1.8. Principle components analysis is done on the covariance matrix C_{YY} in Equation 1.20. Substituting Equation 1.19 into 1.20 gives Equation 1.21, which is the equation for redundancy analysis. Solving Equation 1.21 results in the eigenvectors in U_K, which give the contributions of

the loadings in B, where $\lambda_K I$ are the eigenvalues in matrix form similar to M in equation 1.10 and the canonical coefficients are calculated in Equation 1.22 (Legendre and Legendre 2003):

$$B = [X^{[Cal]T} X^{Cal}]^{-1} X^{[Cal]T} Y^{Cal} \quad \text{Equation 1.17,}$$

$$Y^{Val} = X^{Val} B \quad \text{Equation 1.18,}$$

$$C_{YY} = C_{YX} C_{XX}^{-1} C_{YX}^{T} \quad \text{Equation 1.19,}$$

$$[C_{YY} - \lambda_K I] U_K = 0 \quad \text{Equation 1.20,}$$

$$[C_{YX} (C_{XX})^{-1} C_{YX}^{T} - \lambda_K I] U_K = 0 \quad \text{Equation 1.21,}$$

$$C = B U_K \quad \text{Equation 1.22.}$$

1.3.3 NON-LINEAR METHODS

Non-linear techniques refer to downscaling methods that seek to categorize similar data points into groups by their values and extract individual values from within those groups by either some form of group-based regression or nearest neighbor interpolation. The groups that are used can be designed around the values of the predictands themselves (such as cluster analysis) or can be created from spatial maps of predictor variables (weather classification or typing). Cluster analysis seeks out patterns, called clusters, in the predictor-predictand bi-variate space and assigns all data points to those clusters that they are closest too. For weather classification, predictor maps are classified by similar patterns, with local values of the predictand grouped with their respective pattern. When a new pattern is obtained, it is compared to the groups and assigned to those that are most closely correlated, with the local variable values then assigned based on that pattern. Weather classification schemes have been in use for a number of decades and were initially used to aid in weather prediction (Lutgens and Tarbuck 2007). This methodology relies on the stationarity of the relationship between these classifications and the local values assigned to them, though involving some form of regression

within the classifications can lessen this restraint. The very assumption that must hold in order for this method to work may also prevent it from being applicable to unobserved climate states, as the stationarity of the patterns between climate states would necessitate that the method is insensitive to changing climate conditions.

Various studies have utilized these methods with some mix of success at wringing out fine resolution data that was then put to use in impact studies and model validation. One study mentioned in Section 1.3.1 looked into the use of BCSD and multivariate adapted constructed analogs (MACA) to model fire danger in the Rocky Mountain West (Abatzoglou and Brown 2012). The MACA method used temperature, humidity, precipitation and wind field variables to construct analog maps by which to compare with the local data and categorized them accordingly. While both methods succeeded in giving useful localized data, MACA showed better skill than BCSD, retrieving data with higher accuracy when compared to non-calibration time periods. These results were then extrapolated using the model outputs for the coming century, with projections of increased fire risk in most places in the Rocky Mountains. Another similar study looked into temperature and precipitation values for the same region using BCSD and constructed analogs and found that while both methods were able to match seasonal and monthly values fairly well, the constructed analogs did better at capturing some of the extremes of temperature, particularly summer highs and winter lows (Maurer and Hidalgo 2007). However, both methods showed limited skill in recreating observed wet and dry extremes for daily precipitation, though BCSD continues to be used frequently in hydrological studies. Still other similar methods have been applied to current climates to assess possible past climates, such as European temperatures and precipitation during the Last Glacial Maximum (Vrac et al. 2007).

The artificial neural networks (ANN) techniques are a "black-box" method of downscaling, meaning that, aside from the inputs and outputs, the specific steps within the program are unknown. This method involves the transformation of model predictors into predictands using step-wise conditional pathways that take data points through a series of nodes, which have conditional linkages between them and the input and output nodes which seek the combination of input variables and weights for those variables that best explain the data as a whole (Wilby et al, 2004). While multiple linear regression models do this in a linear fashion, the ANN nodes can manipulate the weights of the inputs in non-linear ways as well. This type of model construction creates pathways through the nodes that do not offer a straightforward physical interpretation and are known to dampen the variance of time series, which renders it less useful in the analysis of extreme values (Benestad 2008). ANN has been applied to a diverse set of downscaling projects, including stream flows in British Columbia, Canada (Cannon and Whitfield 2002) and wind power viability in changing climates (Sailor et al. 1998). In both cases, the neural network approach was able to effectively downscale the GCM data to that of the applications in question, though with the same caveats previously mentioned. Another comparative study looked into the relative strengths of ANN type models and linear regression models in comparing temperature and precipitation over the Midwest (Schoof and Pryor 2001). While both model types arrived at similar results, there were some key differences. As expected, the models did better with temperature compared to precipitation. In this case, however, the linear models had less success than the ANN models, though the differences were less extreme when an autoregressive term was added to the temperature downscaling. The precipitation remained an issue in both models, and was harder to constrain based on the complex factors that pertain to capturing the physical processes most responsible for any rain or snowfall that occurs.

1.3.4 WEATHER GENERATORS

Weather generators utilize randomness simulators that are able to recreate daily weather variables, usually precipitation, minimum and maximum temperatures, and sometimes solar radiation. They take in observed data for a specific location for a specified time period (usually at least 30 years-worth of data) and use the distributions in that data to create variable values that can be programmed to follow a seasonal cycle and to include extreme events, with the caveat that the generator will not take interannual or longer term trends in climate into account, and thus simulate only stable regimes. While trends cannot be simulated, different climate states at different times can be, as long as the trend during the time frame used is not so large as to change the model statistics over that time span. Often, the goal of a weather generator is to provide for temporal downscaling, usually with monthly data supplied and daily data output desired that match the monthly statistical parameters. Another application is the production of multiple runs of the daily weather within the same time span for the same average climate state to ascertain how variable the local climate is for the time and location specified. One study compared the effectiveness of a weather generator WGEN to two separate ANN models in simulating precipitation across six regions of the United States (Wilby et al. 1998). They concluded that WGEN did a better job constraining the true variability and mean statistics, whereas the ANN networks, while able to model the means adequately, were not able to capture the extremes, particularly the wet-day extremes.

1.3.5 STATISTICAL DOWNSCALING COMPARED TO RCMs

Statistical downscaling techniques have also been compared and applied to dynamical methods. When compared in ability to downscale European temperatures and precipitation, it was observed that the statistical model did better in reproducing summer temperatures, whilst the regional model did better with winter precipitation (Murphy 1998). Statistical downscaling

methods applied to the Animas River Basin in Colorado were compared to RCM output for the same region and were found to have comparable skill in recreating the hydrological conditions for the basin, though statistical downscaling is often preferred for hydrological studies due to its less computationally intensive requirements (Wilby et al. 2000). It is also possible to apply statistical downscaling methods to RCM output by first using the RCM to dynamically model a region and then use its output as a predictor or set of predictors in a statistical model. When GCM and RCM output were corrected via linear interpolation or bias correction along the lines outlines in Section 1.1, the BCSD method applied to RCM output did the best in capturing the true local features of precipitation climatology in Columbia River Basin (Wood et al. 2004). Linear interpolation was better when applied to RCM output than GCM data, though not to the extent of bias correction. These results suggest that new studies may benefit from using statistical methods on RCM data rather than on GCM data directly.

CHAPTER 2 – STUDY AREA, DATA AND METHODOLOGY

2.1 NORTHERN GREAT PLAINS

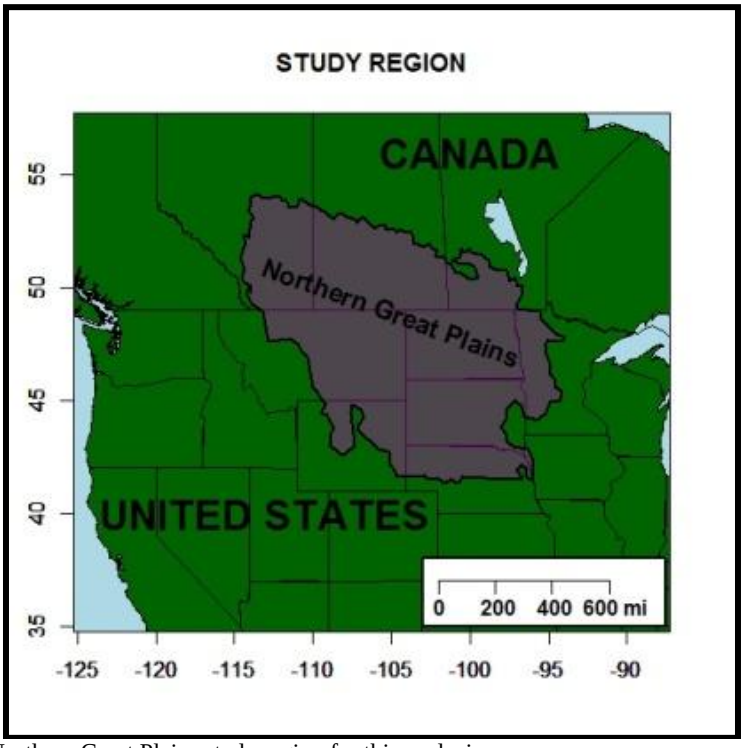

Figure 2.1: A map of the Northern Great Plains study region for this analysis.

The Northern Great Plains is an important, though often overlooked region for the world. It encompasses a continental interior climate regime that challenges local ecological systems through relatively large seasonal changes in temperature and interdecadal fluctuations in precipitation. For much of its post-1800 history, it has been an agricultural center, contributing greatly to global food supplies. While some level of fuel and mineral extraction has always been present in the NGP, there has been a recent spike in energy extraction over the Bakken oil formation made possible by technological advancements in horizontal drilling and hydraulic fracturing. The slow conversion of land use from natural grassland to intensive agriculture,

combined with the recent rapid rise in population and economic development have fragmented local ecosystems and diminished their adaptive capacity (Lemons et al. 2012). These changes have made the region more sensitive to climate change, though the effects of the global shift in temperatures on the NGP are not well constrained by the GCMs (Karl et al. 2009). This is due in part to the NGP acting as a transition zone between the warm and dry Rocky Mountain West and the wetter areas near the Great Lakes and Appalachia. While the trends in precipitation have differed between sub-regions within the NGP, average temperatures have tended to follow global trends, warming at an average rate of 0.14°C per decade since 1880, with 8 of the last 10 summers being warmer than the 20th Century average (Ojima et al. 2012). It is generally expected, as stated in Ojima et al (2012), that the change in variability within the region is the most important to diagnosing the impacts of climate change, rather than the mean changes. This is especially true for precipitation, as the erratic decadal patterns and limited resources in the drier portions of the region, as well as the prospect of increased flooding in the Devils Lake and Red River basins, make information on local changes all the more important.

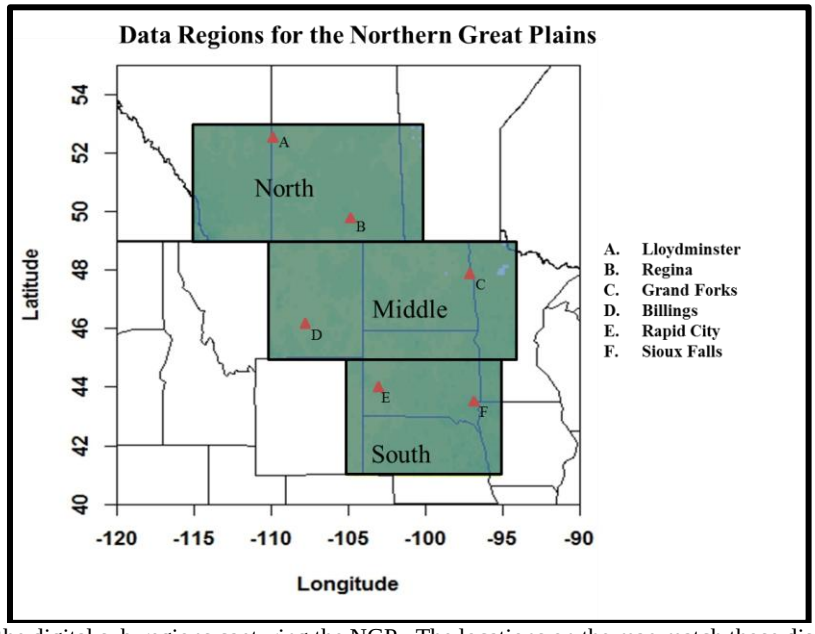

Figure 2.2: A map of the digital sub-regions capturing the NGP. The locations on the map match those discussed in Section 3.3.

The regional focus of this study is on the Northern Great Plains ecoregion, as outlined in Figure 2.1, and captured through the use of three sub-regions shown in Figure 2.2. The boundaries extend from 42°N to 55°N and from 95°W to 115°W. It includes all of North and South Dakota, as well as portions of Minnesota, Nebraska, Wyoming, Eastern Montana and southern areas of the Canadian provinces Manitoba, Saskatchewan, and Alberta. Elevations decrease from west to east as one moves away from the mountainous West. While there is some climatological homogeneity within the region, important variations do exist due to the differences in localized geography. The annual mean temperature tends to hover near 0°C with hot summers (> 30°C) and cold winters (< -15°C), with a large annual temperature range in excess of 45°C and a mean daily range of 13.6°C (see Figure 2.3). This range differs slightly from colder to warmer as one goes from north to south. Precipitation increases from west to east (see right panel of Figure 2.4) across the region, ranging from less than 1 mm/day to over 2.5 mm/day, due to the interaction between the dry air coming off of the Rocky Mountains and the moist air traveling up from the Gulf of Mexico (Rohli and Vega 2008). Figure 2.3 also portrays how the seasonal cycle of daily average precipitation reaches a peak in May or June, while the temperatures peak in late July or early August and reach their lowest levels in January. Differences in seasonal heating and pressure patterns cause the polar front and its annual cycle, bringing seasonal storms, summer convective activity and the winter polar lows that can often result in blizzards that sweep across the region. Rain and snowfall can vary significantly from year to year, with dry periods and drought conditions giving way to wet periods that result in excessive flooding along many of the major rivers. While landlocked within North America, the NGP is influenced by some teleconnections and regional patterns, including the El Nino-

Southern Oscillation (ENSO), the Arctic Oscillation (AO), and the Pacific-North American pattern (PNA).

2.2 DATA

2.2.1 OBSERVATIONS

Observational data come from Maurer et al, (2002) downloaded from the University of Santa Clara (http://www.engr.scu.edu/~emaurer/data.shtml). It consists of gridded station data in which precipitation and temperatures from stations within the continental United States (COOP network), as well as northern Mexico (Servico Meteorologico National) and southern Canada (Environment Canada) were used to fill a 0.125° by 0.125° (1/8 x 1/8, or about 13 km to a side) grid via the SYMAP interpolation used in Sheperd (1984) for the period from January 1, 1950 to December 31, 1999 in daily time steps from averaged 3-hourly data. Other data, such as dew-point temperature, were derived from the station variables however were not used in this analysis. The variables selected as predictands from this analysis were minimum and maximum temperature (Tmax and Tmin in °C) and precipitation (Pr in mm/day). Figure 2.4 shows the mean daily values for the region for the period 1970 – 1999. The total area covered ranged from 25 - 53°N and 67 – 125°W. Grid squares heavily covered by water and without adequate station coverage nearby were left unfilled, including the Red Lake and Lake of the Woods areas in Northern Minnesota. Precipitation data from the interpolated station grid were then scaled to match the long term output of the parameter-elevation regressions on independent slopes model (PRISM). This produced two advantages, in that the precipitation in this data set is station based (not produced through model dynamics, as in the ECMWF-ERA 40 reanalysis) and is optimized to take into account some aspects contained in the PRISM analysis, such as elevation.

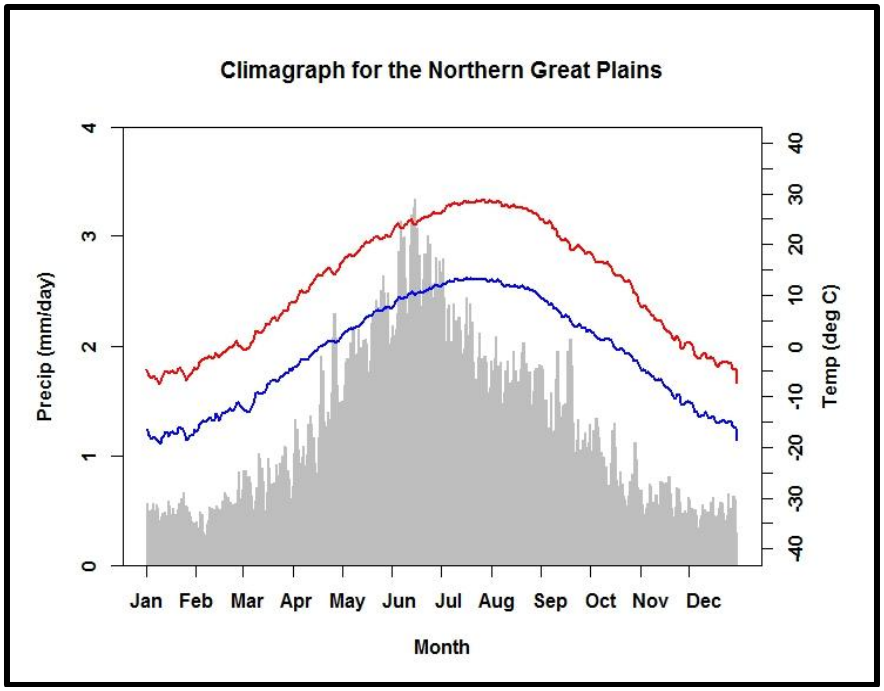

Figure 2.3: A climagraph for the NGP as a whole, based on the Maurer dataset (section 2.2.1). The red line is max temperature, the blue min temperature, and the grey bars are the precipitation amounts in mean daily mm.

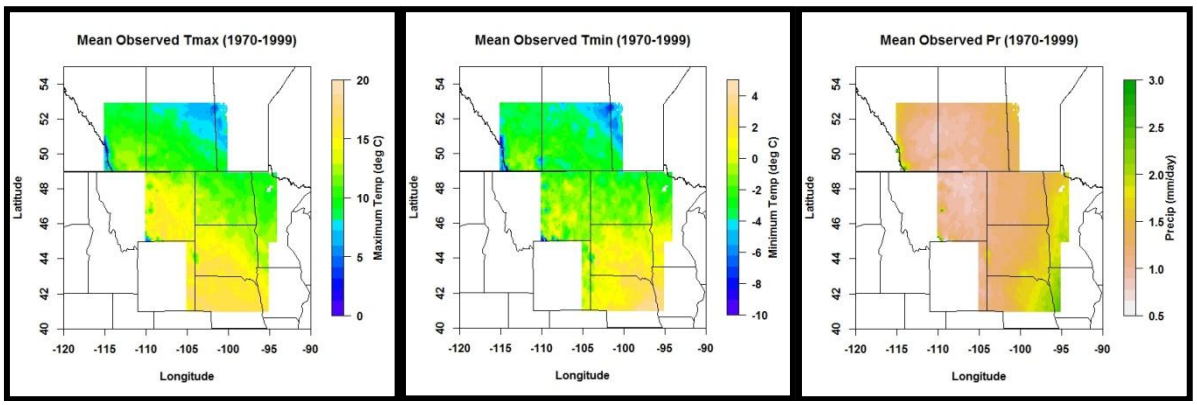

Figure 2.4: Observed maximum and minimum temperature and precipitation maps (from left to right), averaged over the whole verification period (1970 – 1999).

2.2.2 GENERAL CIRCULATION MODELS

Data from 17 GCMs or earth system models were downloaded from the CMIP5 online database (http://pcmdi9.llnl.gov/esgf-web-fe/; see Table 2.1). They range in resolution from about 1 to 4° (mean of 2.5 by 2.5°) in latitude and longitude and from 2.5 to 4.8°C (mean of 3.6°C) in sensitivity to a doubling of CO2. The predictor variables sought from the models corresponded to the observed variables, namely maximum and minimum temperature (in degrees

kelvin, K) and precipitation (in $kg \cdot m^{-2} \cdot s^{-1}$). These were converted to °C (Tmax – 273.14; Tmin – 273.14) and to mm/day (Pr · 86400, as $kg \cdot m^{-2} \cdot s^{-1}$ is equivalent to mm per second) to match the units of the observed variables. These variables were extracted for the NGP-specific area as well as for the North American region (25 - 75°N, 75 - 125°W), for the period 1950 - 1999. The NGP is represented by only about 20 to 30 grid squares in most of the models, and the North American region by only 250 – 450 grid squares. This is in contrast to the high resolution of the gridded observations which encapsulate the NGP with a total of 10,496 grid squares.

Model projections up to 2100 were also downloaded from CMIP5 for the GCMs for which projected data was available (see Table 2.1), with projected data available and used to downscale for projected scenarios forced by the so called representative concentration pathways or RCPs. The same variables and areas were subset as in the verification matrices. There are four possible RCPs to use, each representing a different possible outcome dependent upon assumed human activities and their impact on the radiative forcing of the system (Van Vuuren et al. 2011). Three of the RCPs were chosen for this analysis: RCP 8.5, RCP 4.5 and RCP 2.6, where 8.5, 4.5 and 2.6 represent the external climate forcing by 2100 in Wm^{-2} relative to 1850. For comparison, a doubling of CO_2 ($2xCO_2$) is estimated to cause a radiative imbalance of approximately 3.7 Wm^{-2}. 8.5 represents business as usual, with ever increasing populations, economic growth and technological change leading to growing emissions rates through 2100 and quite possibly quadrupled CO2 levels. 4.5 can be approximated as the roughly doubled CO2, where the annual atmospheric loading rate is held constant or slows somewhat. 2.6 represents a scenario with prompt and heavy action to reduce emissions and to keep CO2 to lower levels, possibly below 500 ppm. While there are multiple possible development scenarios for how to reach each of

these forcing levels, including how anthropogenic greenhouse gas and aerosol emissions change, they were chosen based upon their rough encapsulation of high, middle and low end projections.

Table 2.1: A table of the GCMs used in this study for historical and projected (RCP) time periods. RCP indicators with a * denote that only RCP 8.5 data was available during this project. The equilibrium climate sensitivities (ECS) are calculated on a 2xCO2 forcing of 3.7 Wm^{-2}. The resolution is reported in what was output in CMIP5. The underlined models are used for examining the impact statistics and the RCPs in more detail. Adapted from Flato et al. 2013.

General Circulation Models from CMIP5					
Model Name	**Abbreviation**	**Institution**	**Resolution**	**ECS**	**RCPs Used**
BCC-CSM1.1	**BCC**	*Beijing Climate Center, China Meterological Administration*	2.81 x 2.81	3.3	yes
BNU-ESM	**BNU**	*Beijing Normal University*	2.81 x 2.81	4.1	no
CanESM2	**CAN**	*Canadian Center for Climate Modeling and Anlaysis*	2.81 x 2.81	3.7	yes
CMCC-CESM	**CMCCCESM**	*Centro Euro-Mediterraneo per I Cambiamenti Climatici*	3.75 x 3.75	NA	yes*
FGOALS-g2	**FGOALS**	*LASG (Institute of Atmospheric Physics)- CESS*	2.81 x 2.81	NA	yes
GFDL-ESM2G	**GFDL2G**	*NOAA Geophysical Fluid Dynamics Laboratory*	2.5 x 2	2.6	yes
GFDL-ESM2M	**GFDL2M**	*NOAA Geophysical Fluid Dynamics Laboratory*	2.5 x 2	3	yes*
GFDL-CM3	**GFDLCM3**	*NOAA Geophysical Fluid Dynamics Laboratory*	2.5 x 2	4.8	yes
IPSL-CM5A-LR	**IPSLLR**	*Institut Pierre Simon Laplace*	3.75 x 1.9	4.8	yes
IPSL-CM5A-MR	**IPSLMR**	*Institut Pierre Simon Laplace*	2.5 x 1.25	NA	yes
MIROC5	**MIROC5**	*University of Tokyo, National Institute - Environmental Studies*	1.41 x 1.41	2.6	yes
MIROC-ESM-CHEM	**MIROCCHEM**	*University of Tokyo, National Institute - Environmental Studies*	2.81 x 2.81	NA	yes
MIROC-ESM	**MIROCESM**	*University of Tokyo, National Institute - Environmental Studies*	2.81 x 2.81	4.1	yes
MPI-ESM-LR	**MPILR**	*Max Planck Institute for Meteorology*	1.8 x 1.8	3.3	yes
MPI-ESM-MR	**MPIMR**	*Max Planck Institute for Meteorology*	1.8 x 1.8	NA	yes
MRI-CGM3	**MRI**	*Meteorological Research Institute*	1.125 x 1.125	3	yes
NorESM1-M	**NORESM**	*Norwegian Climate Centre*	2.5 x 1.9	3.3	yes

2.2.3 DATA PROCESSING AND STORAGE

The data files for this project were stored as matrices in CSV files on the N-drive server for the Earth System Science and Policy department at the University of North Dakota. The CSV data format was used for its applicability in the R programming interface. The file names denote the variable and sub-region (North, Middle or South, see Figure 2.2). These data files include the observed fields for maximum (tmax) and minimum (tmin) daily temperature (°C) and daily precipitation (pr, in millimeters), the GCM data for said variables, including NGP-specific data used in BCSD and the North American data for RDA, in the units provided from CMIP5 (temperatures in kelvin and precipitation in $kgm^{-2}s^{-1}$). The downscaled files contain the same

variables in the same units as the observations. Downscaled file names denote the GCM used to downscale, the variable, and the time period or projection, where *Validation* refers to the verification period (1970 - 1999) or the RCP is noted by R (*R2.6*, *R4.5* or *R8.5*) and the decade noted as D (*D20*, *D50* or *D80*). Observed data are reported in the standard calendar, while most of the GCMs used report data without leap years (set 365 days per year). The leap years were removed from the observed data during downscaling, as well as from those GCMs that did include them. This process resulted in downscaled files that do not contain leap years. The observations and GCM data were stored in a 50 year block from 1950 to 1999, resulting in 18250 reported days. Downscaled data are stored in 30 year blocks with 10950 days each. The downscaled matrices report the data with rows denoting locations, with the first two rows denoting the latitude and longitude coordinates of the locations, and dates in month-day-year format along the columns (see Table 2.2). The locations are listed by latitude then longitude, increasing from the southwest corner to the northeast corner of each sub-region. Observed and GCM data are reported in data blocks that report the data with time along the rows and locations along the columns, though the files contain the data matrices only.

Table 2.2: An example of the downscaled data stored as discussed. This example includes the first week of readings for the first ten locations for the maximum temperature forced by the BCC GCM for the southern sub-region. The filename would thus be bccTmaxSouthValidation.csv.

Lat	Lon	1/1/1970	1/2/1970	1/3/1970	1/4/1970	1/5/1970	1/6/1970	1/7/1970	1/8/1970
41	-105	3.83	2.47	0.86	1.07	1.18	-1.8	-3.44	-2.23
41.125	-105	4.26	2.9	1.28	1.47	1.59	-1.38	-3.02	-1.8
41.25	-105	4.79	3.44	1.78	1.96	2.07	-0.84	-2.52	-1.23
41.375	-105	5.33	3.89	2.28	2.45	2.56	-0.43	-2.05	-0.81
41.5	-105	5.7	4.15	2.47	2.68	2.79	-0.19	-1.89	-0.55
41.625	-105	6.07	4.43	2.87	3.05	3.16	0.04	-1.52	-0.3
41.75	-105	6.13	4.51	2.8	2.99	3.1	0.23	-1.47	-0.03
41.875	-105	5.65	4.04	2.43	2.61	2.72	-0.19	-1.93	-0.47
42	-105	5.64	4.06	2.4	2.6	2.71	-0.3	-1.98	-0.58
42.125	-105	5.78	4.26	2.55	2.75	2.86	-0.21	-1.82	-0.46

2.3 METHODS

2.3.1 R – A STATISTICAL PROGRAMMING LANGUAGE

The R software environment provided the means by which to downscale the data from the 17 GCMs. R is a free, open source statistical programming language that is descended from older statistical packages such as S and SAS, and is available for Windows, Macs and Linux operating systems (Team 2000). The source code for R is most often in C or Fortran, and several of the add on packages, including the ones used in this study, *Qmap* and *Vegan*, are sourced to those languages as well. The primary and secondary components are found in the CRAN repository, which is where most packages containing non-base functions and data are stored and able to be easily installed and implemented in the R interface. The openness of the R interface has resulted in many additional packages and uses throughout the sciences, with most being documented in various online forums or able to be accessed through the R *help* function. Many packages are actually coded in the R interface, and are thus able to be manipulated in the package itself or brought directly into the interface for even more maneuverability. The down side to the openness of R, as with any other open source language, is the lack of official support when bugs or other issues arise. However, new versions of R are released every few months or so, with many of the bugs found in previous versions having been patched. The current version as of this writing is R3.2.1.

2.3.2 BIAS CORRECTED SPATIAL DISAGGREGATION

BCSD is accomplished in R by means of quantile mapping, as outlined in Section 1.3. This quantile mapping is done in the two functions *fitQmap* and *doQmap*, found in the function package *Qmap* downloaded from the package repository for R called CRAN. The matrices involved are imported as the observed data, the GCM data for the historical period (1950 – 1999)

and the GCM matrices for the projection period (2010 – 2099 for RCP 2.6, 4.5 and 8.5). First, the historical data files are split between the calibration period (1950 – 1979) and the validation period (1970 – 1999). The use of 30 year timespans ensures that both time periods, even at the expense of having a ten year overlap (1970 – 1979), constitute the traditional climatological-normal that is widely used in climate research (Rohli and Vega 2008). In order to use these functions, the matrices of the GCM data must be interpolated to match the resolution of the observed data. This was done using the *interp2grid* function from the package *Climates*, which is used with the internal R function *apply* on the GCM data to interpolate the latitude and longitude at each time slice. *fitQmap* sets up the model of correction factors for each location, with four parameters input to the function; the observed matrix, the GCM matrix, the specific method of quantile mapping, and the wet day correction. The specific method used is the empirical quantile method, denoted as *QUANT*, which applies Equations 1.2 and 1.3. Other possible methods to choose from include those based upon specific statistical distributions (*DIST*), those using parametric distribution equations (*PTF*) or a form of quantile regression (*RQUANT* or *SSPLIN*). The empirical quantiles were preferred to the distribution-based methods due to the need for a-priori information on what the best distribution would be for each variable, while regression of quantiles using either linear regression or splines resulted in less accurate downscaled data than the empirical quantile mapping. The wet day correction is applied only to the precipitation model by calculating the empirical probability of non-dry days ($Pr > 0$) in the observations and setting all values below that corresponding probability in the modeled data to 0, bringing the ratio in Equation 1.5 as close to 1 as possible. The resulting model from *fitQmap* is then applied to the function *doQmap*, with the input parameters being the downscaling model and the matrix of new GCM data to downscale, and is shown in graphical form in Figure 2.5.

The downscaling function is applied to the verification period data and the RCPs, with the resulting RCPs broken down into the three decadal periods for the 2020s, 2050s and 2080s. These downscaled data files were then output in CSV format and stored according to Section 2.2.3 above.

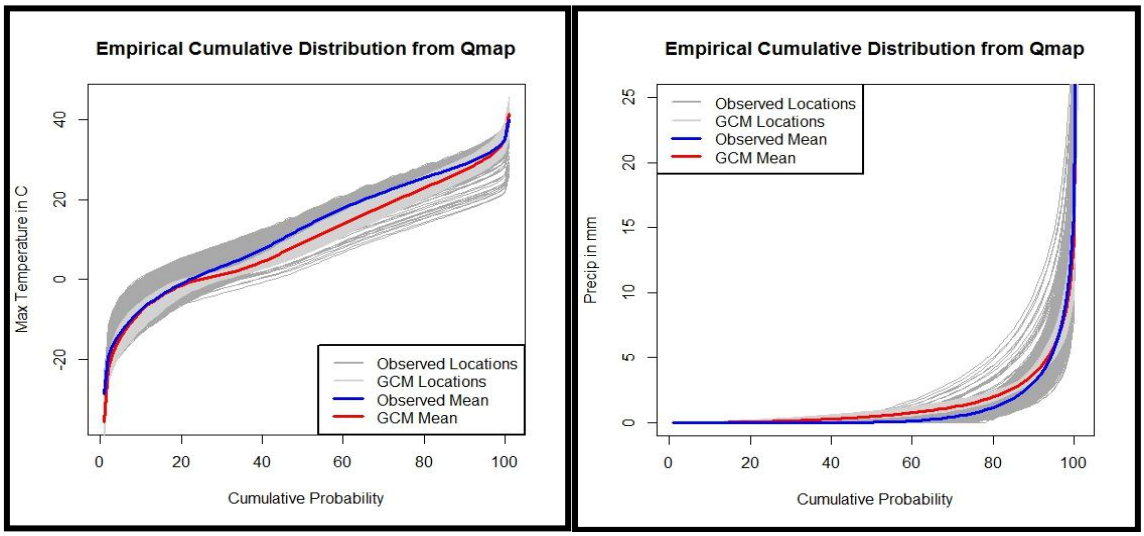

Figure 2.5: Empirical cumulative distributions for Tmax (left) and Precipitation (right) for the BCC model for the middle portion of the NGP region. The *fitQmap* function creates the correction factor using the observed data (blue line) that is then applied to the GCM data (light grey lines) to obtain the downscaled data (dark grey lines). The red line is the average GCM cumulative distribution curve, for reference.

2.3.3 REDUNDANCY ANALYSIS

RDA in R is most often applied through functions pertaining to statistical ecology, which meant great care had to be taken to ensure the proper use of the R package used. The package *Vegan* from the CRAN repository provided the framework function needed for this downscaling method. As in the BCSD section above, the historical data is split into the calibration and verification periods. There are two functions that work to implement this method; *rda* and *predict*. The *rda* function uses the observed and GCM matrices much like *fitQmap* does, in which is constructs the downscaling model using Equations 1.17 – 1.22 in Section 1.3.1. The inputs for this function are the two data matrices and the scaling command. Scaling refers to the standardization of the data by subtracting the mean from the data before the regression is applied

and is done automatically unless otherwise commanded. This scaling is avoided for this study, as the correction imposed by the scaling is not corrected for and results in non-physical values, particularly for precipitation, where initial downscaling attempts resulted in negative daily values. The model derived from the *rda* function is applied to new GCM predictor data in the *predict* function, where the model and GCM data are inputs, along with the specification that the type of output from the function be the full response. Other types of results available from *predict* include location-specific loading scores or statistical ecology metrics, though these are not applicable to this study. As applied here, the process used by the *rda* function simplifies the steps described in Section 1.3.2 down to a straightforward application of Equation 1.17, which is most comparable to multi-variate regression. The model created tries to explain the day to day observations of each predictand at each location by using a combination of data from all locations in the GCM data, in the order that best captures the variations observed. The bulk of the other equations used in the RDA process are best described as further analysis of the relationships derived by the model, which could be useful in a model using combinations of GCM predictors to calibrate the model values, but which is not relevant here due to the purely statistical connection between the GCM predictors and their observed counterparts.

This regression model relied on the ability of the variance and patterns in the predictor data to capture the patterns in the observed data, though it was found that utilizing the observed and GCM data as stored according to Section 2.2.3 above created a couple of problems. One problem was the severe lack of variance in the data resulting from a mass application of the *rda* and *predict* functions to all the data at once. This was fixed by applying the functions to smaller temporal portions of the data, with the observations and GCM data all broken into separate monthly time series (all observed January data regressed against GCM January data, for

example). It was also found that the spatial domain of the predictor data set did make a difference to how well the *rda* function was able to capture the true distribution of the data. When supplied with only the GCM data for the NGP, the downscaled data exhibited significant excursions from the observations over the verification period. This may stem from the lack of possible model loadings (see section 1.3.2B). There are often fewer than 10 or so GCM grid squares covering the NGP. As shown in Figure 2.6, 50 to 100 loadings are about how many explanatory locations it takes for the explained variance to asymptotically approach its maximum. The use of GCM data covering all of North America accounted for this problem by offering more loadings for the regression model. Thus it is clear that the inclusion of more explanatory variables for each observed location increases the amount of variance explained by the model, with this increased skill coming as a result of the increased variance in the larger GCM dataset.

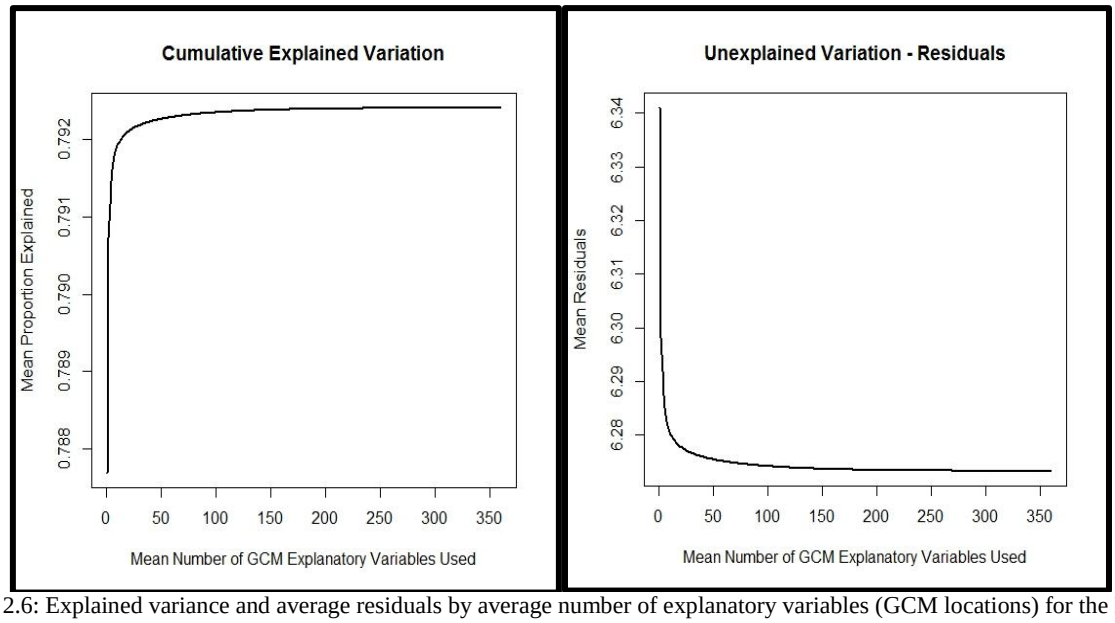

Figure 2.6: Explained variance and average residuals by average number of explanatory variables (GCM locations) for the BCC model over the middle portion of the NGP region. As expected of the RDA model, increasing the explanatory variable loadings increased the explained variation in the left panel and decreased the residuals shown in the right panel.

2.3.4 VERIFICATION PROCESS

In order to test the ability of the downscaled models to capture the general characteristics of the local and regional climate, four metrics were calculated that encapsulate the means and variations and test their differences between the downscaled models and the observations over the verification period. The general distributions of the datasets are compared to assess the relative skill that the downscaled models have in reproducing the data values and their spread. Spatial patterns are also discussed, where in the metrics described below are calculated for six individual grid squares (denoted by a notable town or city located there – see Figure 2.2), as well as spatial maps for the difference of means and the correlations. The metrics include the difference of means, the difference of standard deviations, correlations, and the root mean square error (RMSE), which were all calculated for each location and averaged to obtain the regional NGP average. Two of the metrics, the difference of means and the correlations (Pearson product coefficient), were calculated over monthly time steps, with maximum and minimum temperature being averaged in °C and precipitation totaled in millimeters over each month. The difference of means measures the average difference between the downscaled monthly values and the observed values. Both the difference of means and the correlations were calculated for monthly values due to the chaotic nature of the daily time series. Daily values are expected to vary considerably and daily GCM predictions, even when downscaled, are not expected to match the observed day to day values explicitly (as evidenced by the low correlations between the daily values, especially for precipitation), though the variation in those values should be better encapsulated by downscaling.

The other two metrics, RMSE and the difference of standard deviations, were calculated from the daily data. The difference of standard deviations measures the difference in daily

variability between the downscaled models and the observations. RMSE captures the average error in daily values (in the same units as the monthly means) for the downscaled datasets only, giving a sense of the relative skill in capturing daily values by both downscaling regimes. It is expected that the downscaled models will exhibit lower mean and standard deviation differences while also exhibiting higher correlations compared to the GCM data. Previous studies have indicated that BCSD and RDA produce similar levels of skill at capturing observed data values and improvements over raw GCM data, though the simplification of the RDA method employed (see Section 2.3.3) in this study may alter that somewhat (Tyler 1982). It is expected that the RDA data will exhibit a lower RMSE than the BCSD data, however, due to the regression involved.

2.3.5 IMPACT DIAGNOSTIC ANALYSIS

Impact diagnostics for this study are statistics calculated to assess some of the extreme elements of the NGP climate and how they may change in the 21st Century RCP projections. The impact diagnostics are taken, in part, from the STARDEX project on European weather extremes but modified for the NGP, and are calculated for the region as a whole by calculating the statistics for each location, then averaging among all locations (Goodess 2005). For temperature these include heatwave (maximum consecutive annual days with Tmin > 90th percentile) and coldwave (maximum consecutive annual days with Tmax < 10th percentile) indices, as well as frost days (mean annual total of days in which Tmin < 0°C). Growing degree days (GDD) are calculated according to Miller et al (2001), where an average of daily Tmax and Tmin is calculated and a baseline temperature (growing threshold) is subtracted, with the resulting daily totals added together to obtain annual growing degree days. The baseline chosen

here was 10°C, which is close to the threshold for many crops, including many grown in the NGP (Ojima et al. 2012). Precipitation indices include the maximum consecutive dry day lengths (indicative of excessive drought) as well as the percentage of precipitation from heavy events (Pr > 90^{th} percentile) and the number of said events (possible precursors to flooding). The percentiles for the temperature and precipitation indices are taken from the observations. These impact diagnostics are compared over the verification period to further test the downscaled models ability to capture observed climate characteristics before being applied to the projections over the 21^{st} Century to assess their changes over time.

CHAPTER 3 – RESULTS

3.1 ASSESSING STATIONARITY

The relationship between the coarse scale predictors and the local scale predictands must remain the same, even under changing climate conditions (Benestad 2008). Put another way, a robust downscaling method must be able to capture the distributional patterns in the data while accounting for the shifting temporal mean. Linear methods can account for such shifts in the mean and show the relevant trends provided the predictors used to force the model exhibit strong enough connections to the predictands and are broadly influenced by shifting boundary conditions. Non-linear models that utilize weather pattern classification are prone to break the assumption of stationarity due to the complicated ways in which changing mean climate conditions interact with the broad scale circulation, which are not well captured by the limited number of pattern classifications available from historical data (Wilby et al. 2004; Wilby 1997). In this study, the coarse GCM temperature and precipitation data act as the predictors. Given that the aim of the GCMs is to capture the real world as accurately as possible, it is reasonable to assume that there is some relationship between the coarse data and the predictands. As the GCMs are forced by external parameters and allowed to dynamically change, it is also reasonable to expect that the coarse variables will exhibit changes due to the external factors and thus communicate those changes to the predictands through the derived statistical models.

One way stationarity is tested in downscaling is by the use of so called “perfect proxy” models. The basic premise of this method is that given that there are no observational values beyond the present, model runs are employed to act as observations for both the historical and

projected periods, with one set of GCM data acting as the predictor and the other GCM or RCM data the predictand, thus creating "observations" at times outside of the historical record (Driouech et al. 2010). One simplified mechanism involved the use of a high resolution, 25 km GCM run for two periods, 1979 – 2008 and 2086 – 2095, under two different warming scenarios for the 21st Century (Dixon et al. 2013). Predictor data was created by coarsening the GCM data to 1/64th the resolution for both the historical and projected periods. The high resolution data was then related to the coarse data through a statistical model (which method is not elucidated in the presentation), then the model was applied to the coarse data in the projection to achieve statistically downscaled version of the data for the late 21st Century. These data were then compared to the original high resolution GCM data, which in this case showed that the statistical model was not stationary, as the spatial and temporal residual patterns did not hold between the historical and projected periods. Adding to that conclusion, the breakdown of stationarity in time was amplified in the model run exhibiting greater warming, indicating a serious problem in the statistical model relating the predictor to the predictand.

A simple method not requiring the use of GCMs or RCMs involves the calculation of the residuals of the downscaling models and the associated slopes, which elucidates the evolving predictor-predictand relationship. In Schmith 2008, a non-linear regression was used to capture temperature and precipitation data from stations across Europe using geopotential height data from the NCEP-NCAR reanalysis. The residuals for each station were calculated as deviations of the downscaled values from the observations over a non-calibration period. The residuals were subjected to a Monte Carlo simulation to ascertain the uncertainty of the linear slope calculated from said residuals. While a slope was evident in the residuals for all the stations, most were not significant at the 0.95 level, though a few were. Other regression methods were

also tested and yielded varying results, with some showing good correlations to the observations and relatively stationary model-observation relationships, where others also had high correlations yet showed significant trends in the residuals in over half of the station time series. Other tests of the stationarity were also discussed, such as the Theil-Sen non-parametric trend estimate and the test of split intercepts, where trends for different portions of the data are calculated to see if the end points line up. Schmith describes the possible role predictors play in generating the differing levels of stationarity, noting that the principle component patterns of geopotential yielded higher stationarity than the raw predictor values. The advantage of this type of check on stationarity is the data and methods used are more available.

For this study, the method of residual slope is used to test the assumption of stationarity for the downscaling employed for the NGP. The residuals are calculated by subtracting the downscaled daily values from the observed daily values over the verification period. The slope is calculated by ordinary least squares regression for each location and stored as time series. As locations clustered so close together are not expected to be independent (highly collinear), the slopes for each of the 10000 runs are averaged over the region and summarized in Figures 3.1 and 3.2, which show the results for maximum temperature and precipitation for downscaled BCC model values. Spatial variations in slope were small for most of the region, though the area in the North region near the base of the Rocky Mountains showed some increased deviations. The confidence interval of the slopes is calculated by means of a bootstrap method, wherein random orders of the residuals are sampled without replacement to create 10000 time series of possible regionally-averaged residuals. In each case, for both temperature and precipitation and for BCSD and RDA, the slopes clustered close to 0, with all mean values and confidence intervals on the order of 10^{-7} or so in magnitude. The p-values for each test were 0.51, 0.55 for the

temperature tests and 0.74 and 0.28 for precipitation. As none of the values were below 0.05 ($p < 0.05$), none of the slopes calculated were statistically significant, and thus very likely indistinguishable from 0. Thus it is shown that the relationships derived by the models to relate the biases in the GCM data to the observed data are robust in time, at least over the verification period.

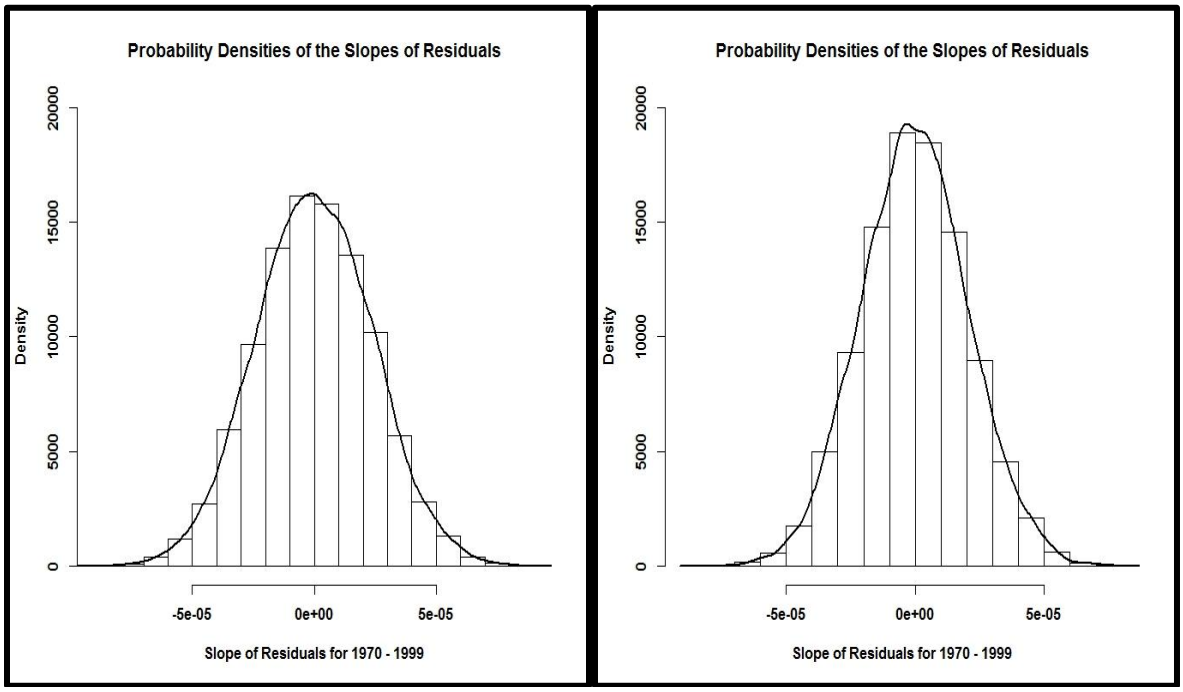

Figure 3.1: The distribution of slopes created from the resampling of residuals for maximum temperature for one GCM. BCSD is shown in the left panel, RDA in the right panel.

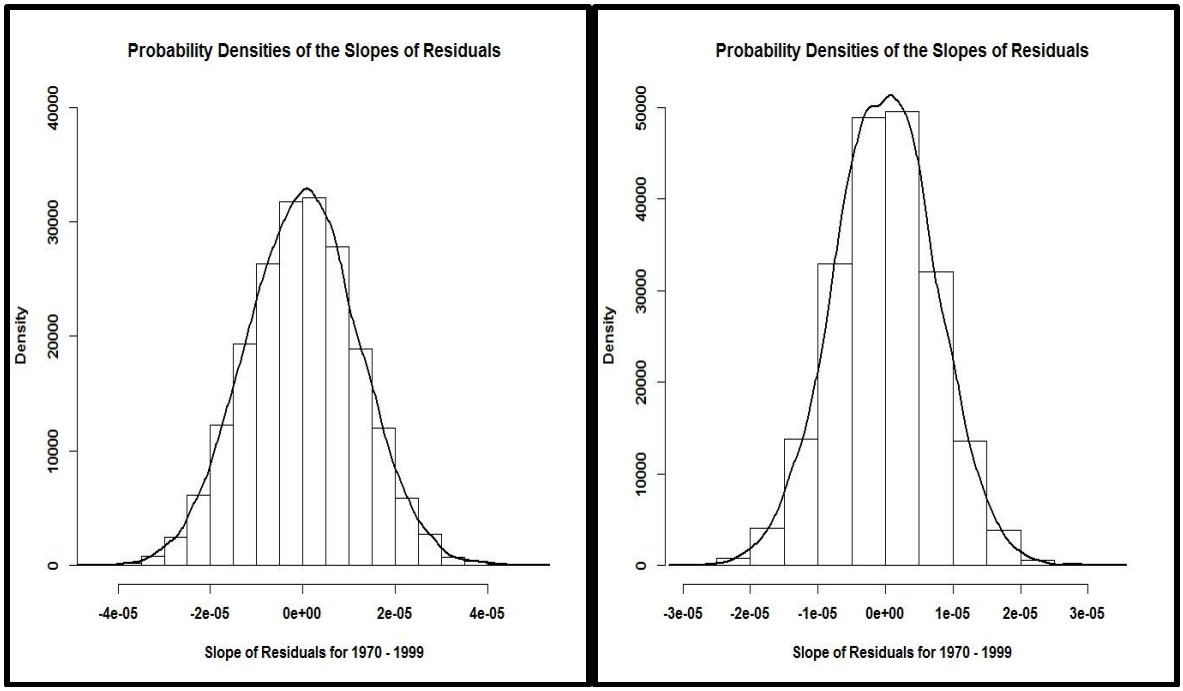

Figure 3.2: The distribution of slopes created from the resampling of residuals for precipitation for one GCM. BCSD is shown in the left panel, RDA in the right panel.

3.2 STATISTICAL VERIFICATION: SPATIAL AND TEMPORAL

Tests of the downscaled models skill at reproducing variations in the observed data showed unanimous improvement over the GCMs, though to varying levels of statistical significance. Table 3.1 summarizes these findings for all of the downscaled models. In general, the correlations between the downscaled data and the observed data are higher than those between the GCM data and observations, with RDA producing slightly higher correlations than BCSD. As in previous studies, the correlations for temperature (about 0.9) were higher than precipitation (0.3), though RDA models improved temperature correlations by almost 0.05 and brought precipitation values above 0.5 compared to those of BCSD which were similar to GCMs for temperatures and 0.38 for precipitation. The RMSE scores are smaller for RDA than BCSD for minimum temperature and precipitation and comparable for max temperature. Mean and standard deviation differences are smaller (by 50 – 90 percent) for downscaled temperature data compared to the GCMs. BCSD shows an improvement in precipitation over GCMs as well with

regard to mean (>2 mm versus >13 mm), and standard deviation (<1 mm versus >5) differences, however RDA models do slightly worse at capturing the absolute values of precipitation and its variations than the GCMs. This pattern in the skill of the downscaled models (or lack thereof) in capturing the climatological characteristics of the three variables is robust throughout the verification period and for every downscaled GCM.

Table 3.1: The table of validation statistics used in this study. Mean Diff refers to the difference of means at the monthly time scale, while the St. Dev Diff refers to the difference of standard deviations at the daily scale. The values represent the averages of all the downscaled models, with the confidence intervals, calculated at 95 percent, coming from the spread of those models.

Validation Statistics (1970 - 1999)			
Max Temperature	**GCM**	**BCSD**	**RDA**
Correlation	0.919	0.925	0.949
confidence interval	0.005	0.004	0.006
RMSE	NA	9.061	9.163
confidence interval	NA	0.179	0.428
Mean Diff.	-1.677	0.195	-0.189
confidence interval	1.098	0.175	0.101
St. Dev Diff.	-0.923	0.227	0.241
confidence interval	0.751	0.099	0.307
Min Temperature	**GCM**	**BCSD**	**RDA**
Correlation	0.917	0.927	0.952
confidence interval	0.005	0.004	0.004
RMSE	NA	8.31	7.762
confidence interval	NA	0.133	0.354
Mean Diff.	2.916	-0.261	-0.397
confidence interval	0.836	0.164	0.075
St. Dev Diff.	-0.695	0.324	0.225
confidence interval	0.537	0.1	0.246
Precipitation	**GCM**	**BCSD**	**RDA**
Correlation	0.184	0.378	0.591
confidence interval	0.053	0.035	0.01
RMSE	NA	5.097	4.504
confidence interval	NA	0.061	0.169
Mean Diff.	13.824	-1.14	14.287
confidence interval	5.055	0.871	4.031
St. Dev Diff.	-0.016	0.02	-0.846
confidence interval	0.285	0.078	0.26

The distributions shown in Figures 3.3 and 3.4 for the predictands illustrate how well the downscaled models captured the mean and spread of the data. The maximum and minimum temperature distributions are relatively well reproduced by both methods, as shown in Figure 3.3. Both exhibit the characteristic bimodal structure, with observed peaks around 15°C (warm peak) and -2°C (cold peak) for minimum temperatures and 25°C and 0°C for maximum temperatures. Both methods show a drastic improvement over the GCM distributions. General circulation models tended toward smaller variations in the seasonal range and daily ranges (minimum temperatures tend to be too warm and maximum temperatures too cool). This pattern of reduced variation is also seen in Table 3.1, where GCM standard deviations are -0.9 (tmax) and -0.7 (tmin) compared to observations. The BCSD models clustered somewhat more reliably about the observed distribution, and had smaller monthly mean and standard deviation differences than GCMs. RDA models also capture both temperature distributions, though the cold peak in minimum temperatures is not as well captured, with two minor peaks produced instead.

The precipitation distributions in Figure 3.4 highlight the stark differences in skill between the BCSD and RDA methods. Observations follow a log-normal or gamma-like distribution. GCMs reproduce this pattern, though with fewer dry days (lower peak near 0) and slightly higher levels of low-rate precipitation. BCSD models corrected this pattern, though with somewhat too many dry days leading to a slight dry bias, as seen in Table 3.1, where monthly precipitation values averaged about 1.14 mm less than observed over the verification period. RDA models did little to correct the GCM biases and by some measures, made them worse. Where the BCSD models tended to over correct the wet bias in the GCMs slightly, RDA heightened the differences by producing too many occurrences of slight to moderate precipitation at the expense of dry days. Both methods, however, captured the extreme precipitation days

relatively well, though BCSD did better at capturing the standard deviation (-0.02 versus -0.85 for RDA – see Table 3.1). The large biases in RDA precipitation are discussed further in Section 3.3.

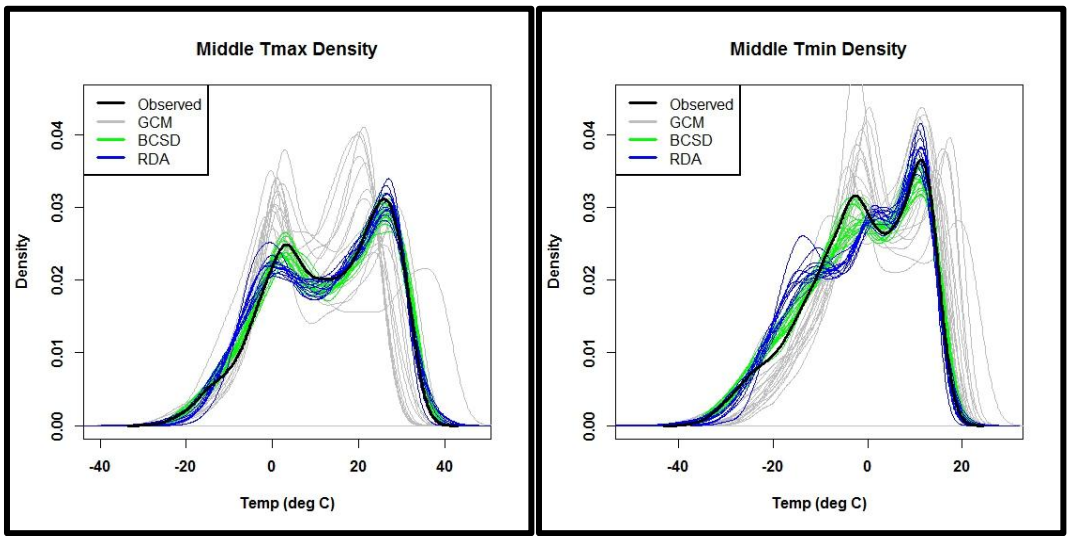

Figure 3.3: Ensemble density distributions of middle-portion NGP maximum and minimum temperatures for the verification period 1970 – 1999, as an example. The other sub-regional temperature distributions (not shown) were similar.

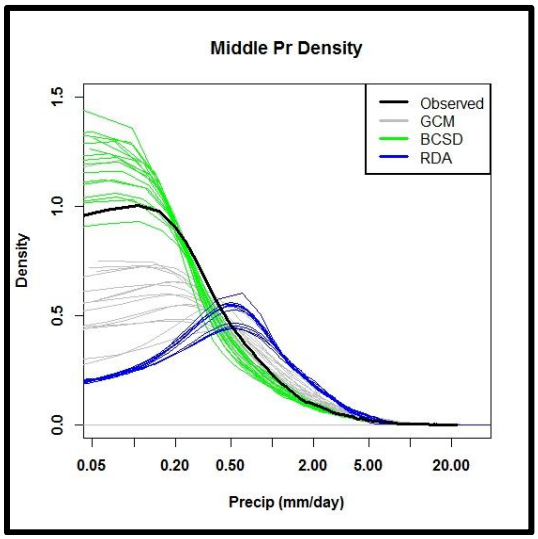

Figure 3.4: Ensemble density distribution of middle-portion NGP precipitation for the verification period 1970 -1999. The other sub-regional precipitation distributions (not shown) were similar.

The spatial patterns of the difference of means and the correlations between the downscaled and observed data are shown in Figures 3.5 through 3.7. The meridional gradient in

both tmax and tmin is well replicated both downscaling methods. As shown in Figure 3.5 and Figure 3.6, BCSD produced striations in the bias pattern, with warm biases along the edges of the data ranges and where many of the boundaries of the GCM grid squares occur within the regions and are most pronounced in the middle and south portions of the NGP. Biases in the RDA models exhibit no pattern, though there does seem to be a shift to slightly warmer biases in the southeast section of the NGP, particularly for maximum temperature. Both BCSD and RDA exhibited non-patterned cool biases in the north data region. This cool bias in the northern portions of the NGP may be related to the change in station networks used to fill in the observed data grids between the United States (COOP network) and Canada (Environment Canada). The average bias is small for both maximum and minimum temperature, as shown in Table 3.1.

The mean bias in precipitation shown in Figure 3.7 shows much less of a spatial pattern than the temperature biases. Both methods did capture the zonal distribution of precipitation, with lower values in the West and higher values in the East, though RDA did produce values in excess of 2 mm/day much further West than is observed. BCSD precipitation has very small biases that tend toward zero. The only striation is visible in the middle data region as a small wet bias that occurs in the same place as the most prominent biases in the BCSD downscaled temperatures. The RDA method produced no pattern in the mean biases for precipitation, though the magnitude, as mentioned above, is larger than BCSD, with much of the area averaging 0.5 – 1 mm/day too wet compared to observations, which amounted to over 14 mm per month as per Table 3.1.

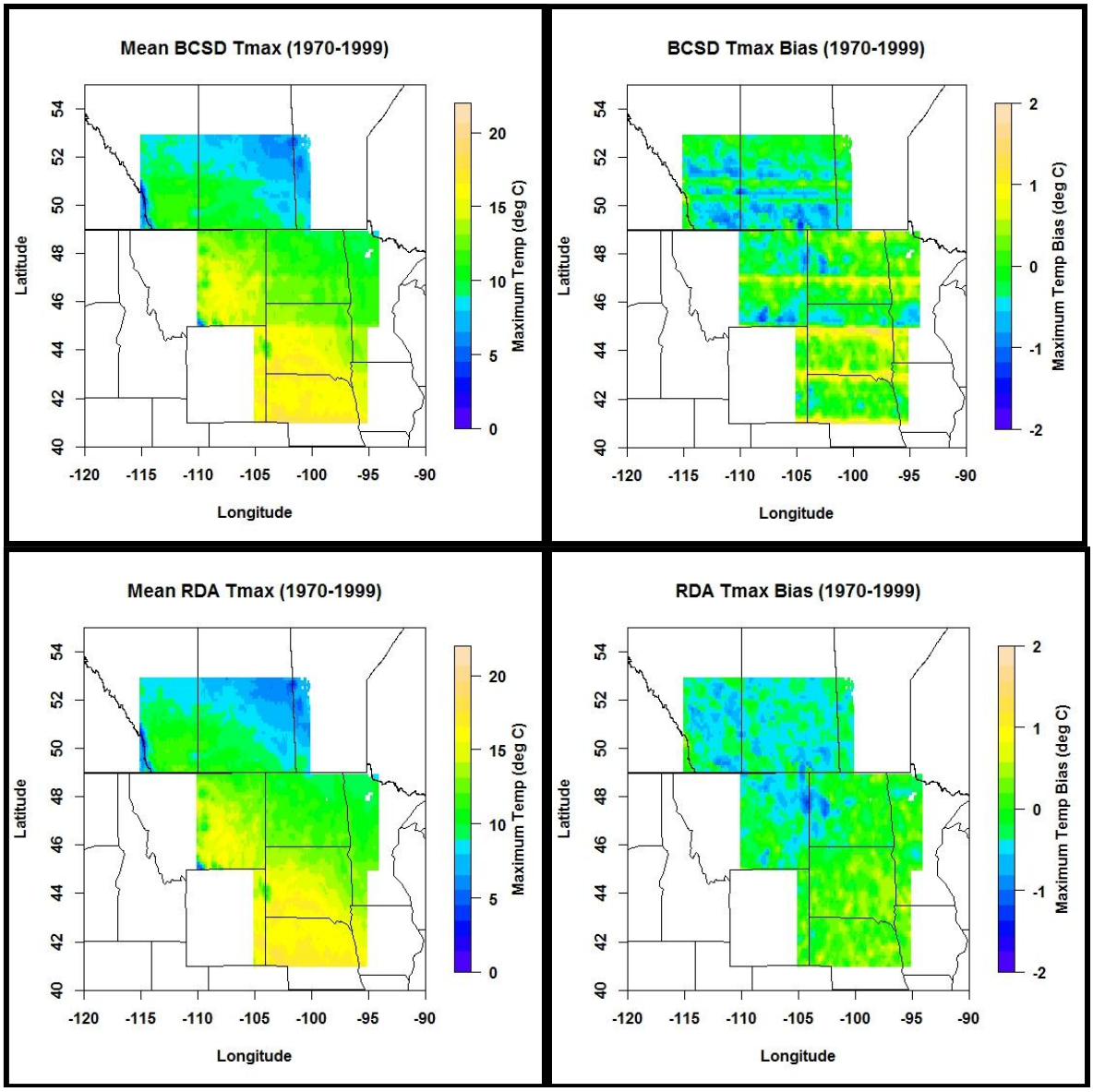

Figure 3.5: Ensemble mean downscaled models for Tmax, left are the mean values over the verification period, right are the biases compared to the observations. Note the stripes in the biases from BCSD, discussed in text.

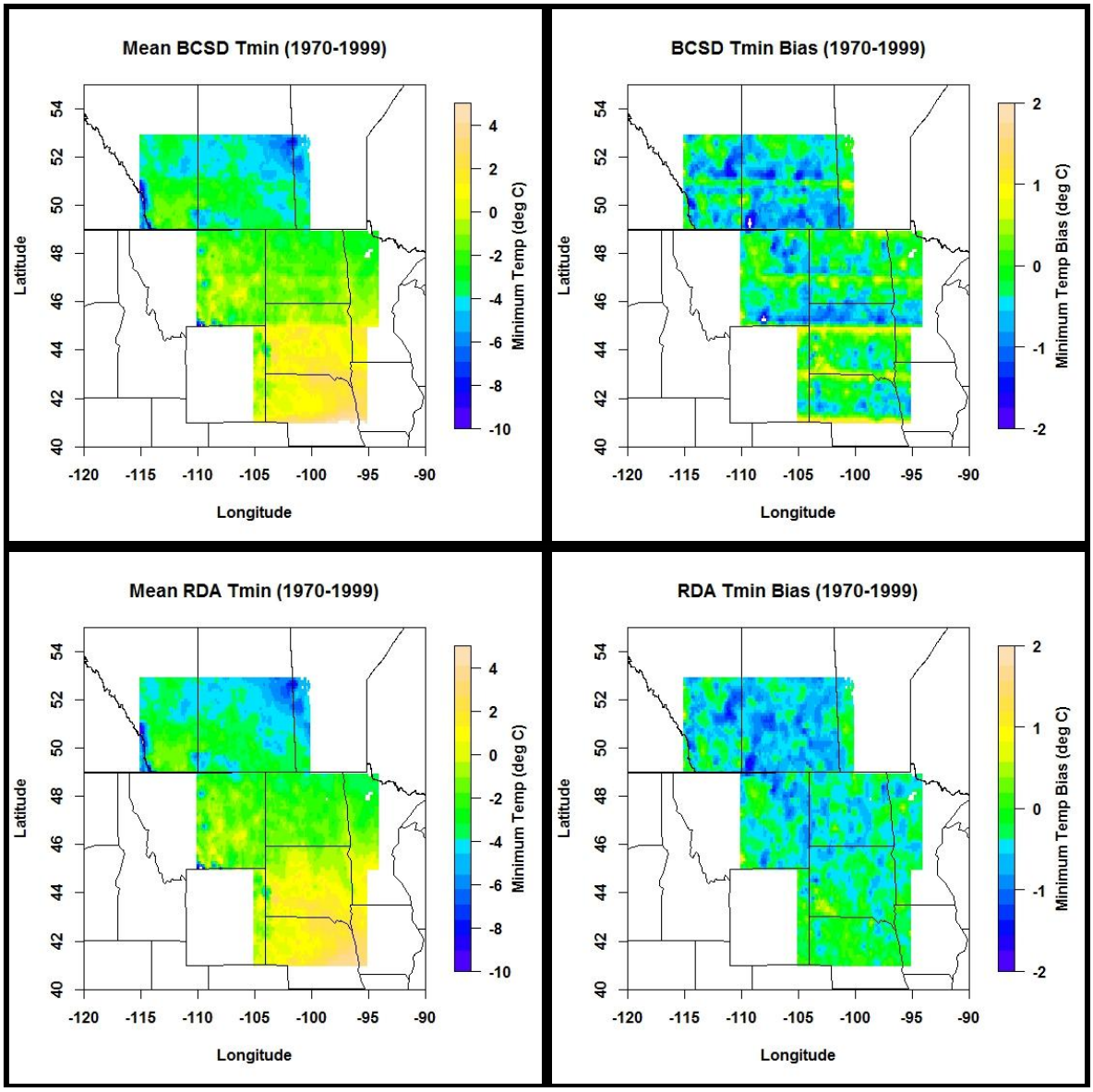

Figure 3.6: Ensemble mean downscaled models for Tmin, left are the mean values over the verification period, right are the biases compared to the observations. Note the stripes in the biases from BCSD, discussed in text.

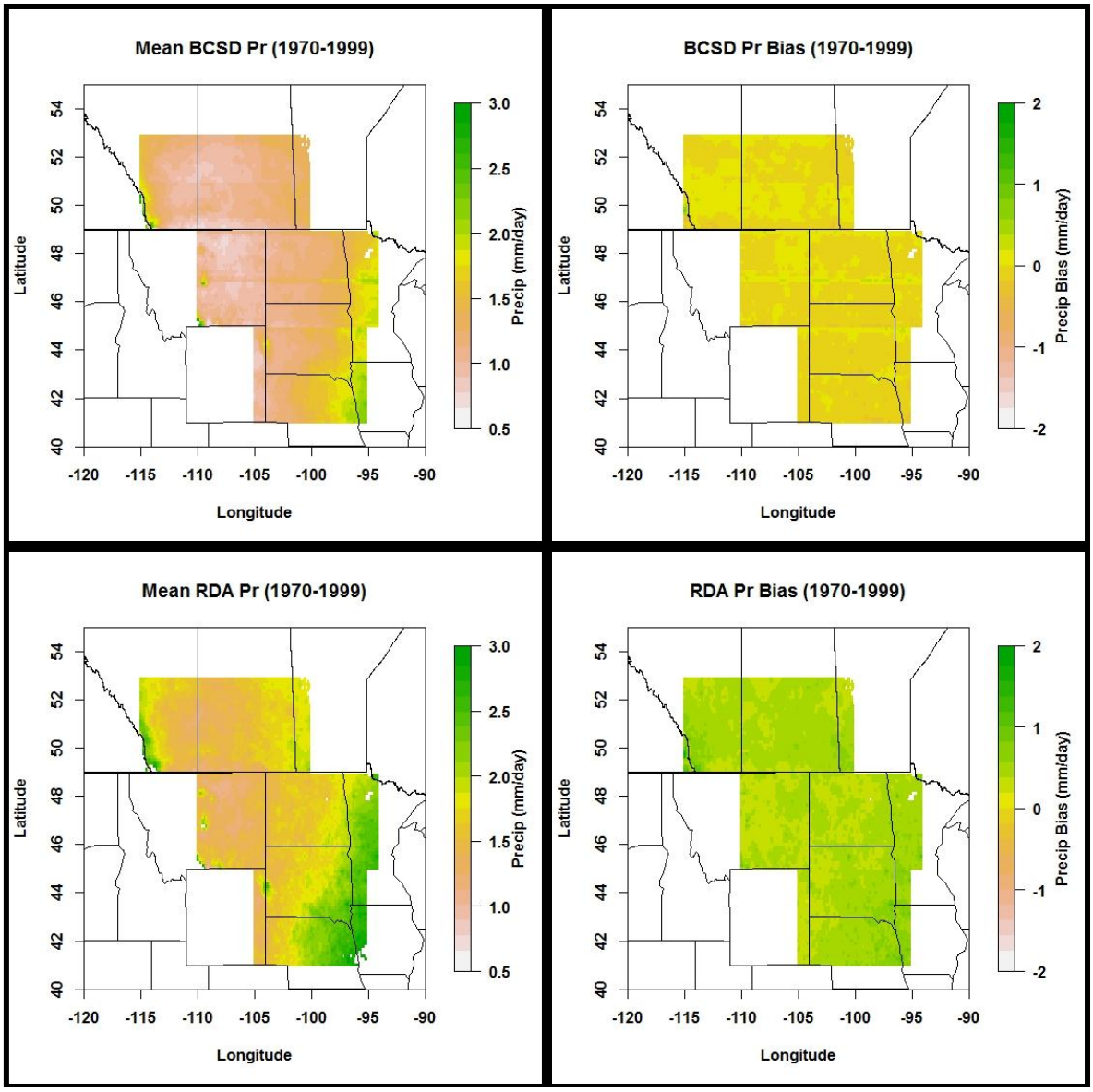

Figure 3.7: Ensemble mean downscaled models for Pr, left are the mean values over the verification period, right are the biases compared to the observations. The stripes are less apparent in the precipitation spatial averages than in temperature, though a small wet bias occurs in the central part of the middle portion of the NGP.

Figure 3.8 depicts the correlations for the between the observed and downscaled values for the three predictands. Correlations for maximum and minimum temperature range from 0.85 to 0.97, with RDA showing higher correlations than BCSD. The striations evident in the BCSD mean difference maps are also present in the correlations. There are anomalously low correlations between BCSD downscaled values and observations at the edge of the Rocky

Mountains in the southwest corner of the north data region, especially for minimum temperature, though the low correlations only fall to between 0.85 and 0.9, compared to the mean value of 0.93 (BCSD) and 0.95 (RDA). The RDA temperature correlations exhibit a meridional gradient, with higher values in the north and east, and lower in the south and west. Precipitation correlations showed little pattern for either BCSD or RDA. There is, however, an area of low or even slightly inversely correlated values for BCSD in the same area that exhibited somewhat lower temperature correlations.

3.3 LOCAL VERIFICATION

The six locations selected from within the NGP region offer a representation of the skill of the downscaling methods at the local level. This level of analysis shows that the skill of the methods at individual locations (at the highest possible resolution for this dataset, at 1/8° latitude and longitude) mirror the results of the regional analysis described above. Figures 3.9 and 3.10 show an annual (at daily resolution) and 5-year (at monthly resolution) time series for maximum temperature (3.9) and precipitation (3.10) for two of the six locations. Table 3.2 gives the same statistics as Table 3.1 at the scale of the six individual grid squares. As with the NGP as a whole, temperature correlations are higher than precipitation, while RDA produced better correlations than BCSD, as well as generally smaller RMSEs. However the wet bias in the RDA precipitation models is also clearly evident in Figure 3.10, as the total monthly values average 12 – 15 mm too high (see Table 3.2) and show no improvement compared to the GCMs, which also exhibit a wet bias. RDA models also under-represent the variation in precipitation, as the standard deviation differences are higher than for BCSD models.

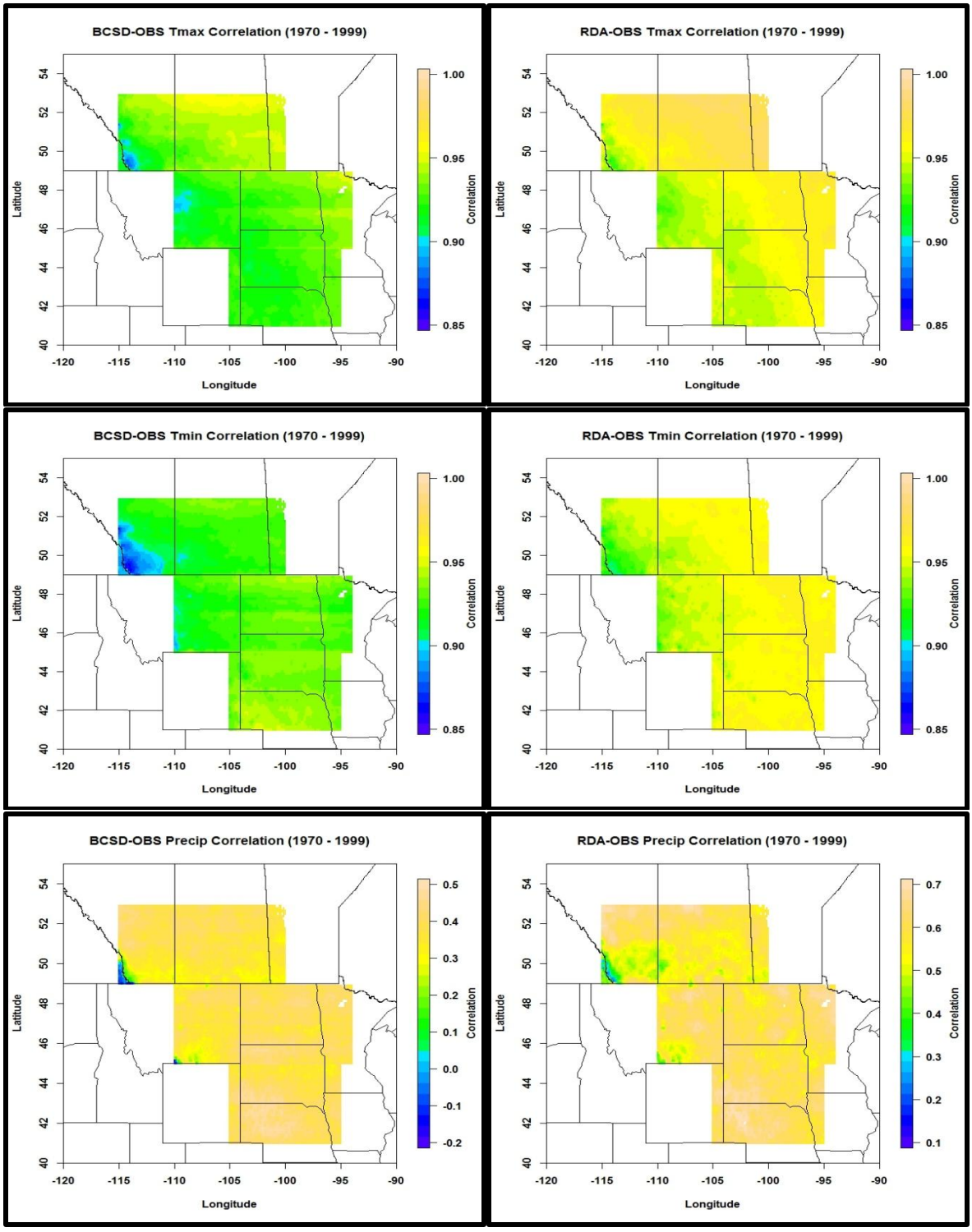

Figure 3.8: The spatial correlations for the verification period for Tmax, Tmin and Pr. The left column of panels are for BCSD, the right column for RDA. Note the slightly different scales for the precipitation correlations at the bottom, which were used so the patterns would be clearer.

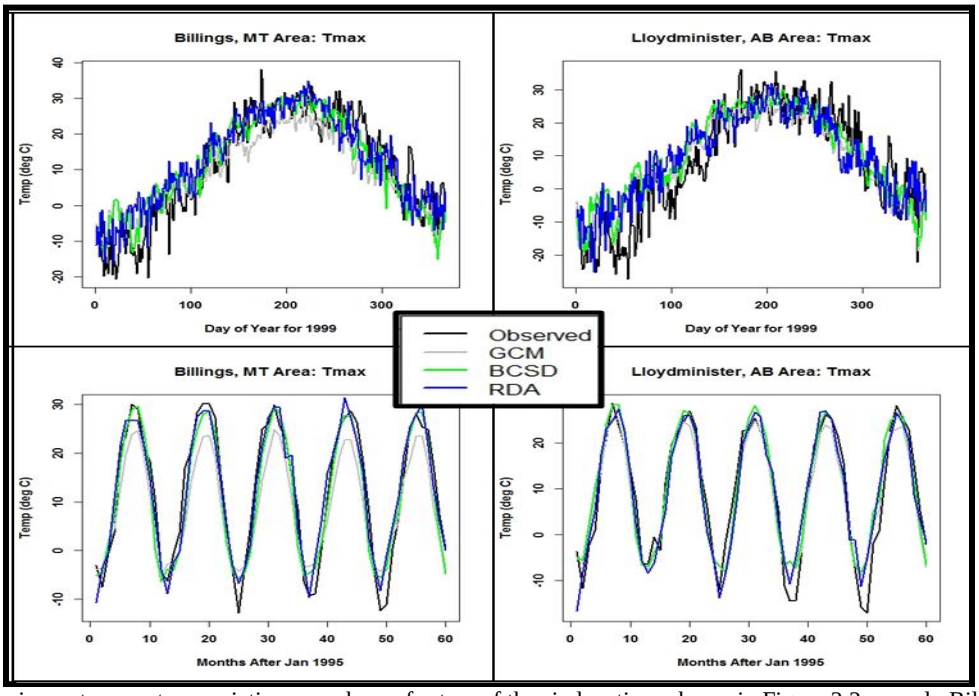

Figure 3.9: Maximum temperature variations are shown for two of the six locations shown in Figure 2.2, namely Billings, Montana and Lloydminster, Canada. The top row shows the average daily temperature, in degrees Celsius, for a year averaged over the verification period. The bottom row shows 5 years of monthly data in the same units starting from January 1995.

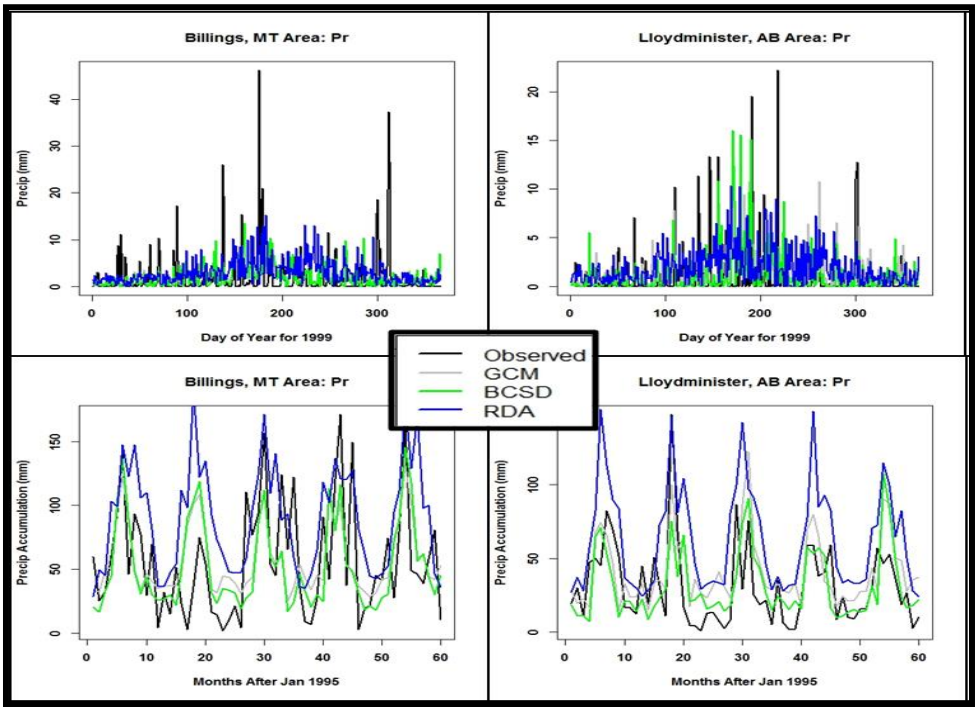

Figure 3.10: Precipitation variations are shown for two of the six locations shown in Figure 2.2, namely Billings, Montana and Lloydminster, Canada. The top row shows the average daily precipitation, in millimeters, for a year averaged over the verification period. The bottom row shows 5 years of monthly data, also in millimeters, starting from January 1995.

Table 3.2: A table of the same values as Table 3.1 but for the six individual grid squares shown in Figure 2.2. The top row (left column on this page) shows the average values of all the BCSD models for these locations, the bottom row (right column on this page) the RDA values.

Grid Square Verification Statistics															
BCSD					Max Temperature			Min Temperature			Precipitation				
Area	**State/Country**	**Region**	**Lat**	**Lon**	**Corr**	**RMSE**	**Mn Diff**	**Sd Diff**	**Corr**	**RMSE**	**Mn Diff**	**Sd Diff**	**Corr**	**RMSE**	**Mn Diff**	**Sd Diff**
Regina	Saskatchewan/Canada	North	50.5	104.6	0.92	9.01	0.25	0.16	0.89	8.07	0.37	0.28	0.34	4.32	-0.32	0.08
Lloydminster	Alberta/Canada	North	52.5	110.1	0.88	8.94	0.14	0.23	0.88	8.01	0.2	0.29	0.36	4.45	-0.45	0.07
Grand Forks	North Dakota/USA	Middle	47.9	97.1	0.91	9.23	-0.05	0.14	0.9	8.29	0.08	0.27	0.35	4.67	-0.67	-0.02
Billings	Montana/USA	Middle	45.8	108.5	0.93	8.81	0.1	0.09	0.92	7.87	0.22	0.21	0.38	5.01	-0.81	0.09
Sioux Falls	South Dakota/USA	South	43.5	96.7	0.92	8.86	0.22	-0.03	0.91	7.92	0.28	0.03	0.41	4.01	-1.2	-0.05
Rapid City	South Dakota/USA	South	44.1	103.2	0.91	8.79	-0.1	-0.13	0.93	7.85	0.03	-0.05	0.43	4.61	-1.7	-0.08
RDA																
Area	**State/Country**	**Region**	**Lat**	**Lon**	**Corr**	**RMSE**	**Mn Diff**	**Sd Diff**	**Corr**	**RMSE**	**Mn Diff**	**Sd Diff**	**Corr**	**RMSE**	**Mn Diff**	**Sd Diff**
Regina	Saskatchewan/Canada	North	50.5	104.6	0.92	9.32	0.35	0.18	0.91	8.36	0.47	0.3	0.57	4.42	13.5	-0.89
Lloydminster	Alberta/Canada	North	52.5	110.1	0.9	9.41	0.24	0.16	0.93	8.45	0.3	0.22	0.55	5.22	12.8	-0.98
Grand Forks	North Dakota/USA	Middle	47.9	97.1	0.93	9.03	0.1	0.2	0.92	8.07	0.23	0.33	0.58	4.77	16.7	-0.76
Billings	Montana/USA	Middle	45.8	108.5	0.94	9.12	-0.12	0.15	0.93	8.16	0	0.27	0.58	5.11	12.3	-1.02
Sioux Falls	South Dakota/USA	South	43.5	96.7	0.95	8.74	0.11	0.11	0.94	7.78	0.17	0.17	0.62	5.05	14.5	-0.87
Rapid City	South Dakota/USA	South	44.1	103.2	0.95	8.63	0.08	0.12	0.96	7.67	0.21	0.25	0.6	4.85	14.3	-1.12

The verification statistics described above for the region both temporally and spatially, as well as for a subset of individual locations, indicate a significant difference in skill between the BCSD and RDA method, particularly with regard to precipitation, and a general improvement over the data produced by GCMs for the NGP. BCSD was able to reduce the monthly mean biases for maximum and minimum temperature, as well as precipitation, while improving the representation of the variation in the data as shown by the differences in standard deviation between the downscaled and observed data. The correlations and RMSE were similar to or slightly improved compared to the GCMs, though precipitation did show an increase in correlation from less than 0.2 to over 0.35. There is a slight dry bias in the BCSD data, as well as some spatial patterns in the mean monthly bias and correlations that suggest some problems for the method at the meeting points of GCM grid squares or in regions with significant variations in the data from the regional average, such as at the foot of the Rockies. RDA showed comparable skill to BCSD with regard to maximum and minimum temperatures while increasing the correlations further than BCSD at the monthly time scale. There is a significant wet bias in the RDA downscaled precipitation data, with daily wet biases giving rise to acute monthly wet biases, though the correlations are over 0.55 and the RMSE is comparable to BCSD at the daily time scale. Thus RDA shows significant improvement in correcting GCM biases for temperatures, however shows no improvement, and by some metrics a degradation, in daily precipitation, which is discussed in the next section.

3.4 RDA PRECIPITATION BIAS

The failure of the RDA method to adequately capture the precipitation over the NGP is a curious issue, though not unforeseeable given the problems outlined in the literature regarding the use of such regression-based methods to model daily precipitation. The biggest problem with

using multivariate linear regression either within or outside of the auspices of redundancy analysis to capture daily precipitation is the incompatibility of fitting a Gaussian or normally distributed model to non-Gaussian data distributions (Wilby et al. 1998). Daily precipitation is better captured by a log-normal or gamma distribution, with a large number of dry days and an exponentially decreasing frequency of increasing precipitation rates. The application of simple regression models to daily precipitation tends to produce either too many or too few wet days, often with reduced variance compared to the observations. The ability of the RDA method to capture the maximum and minimum temperature data with a level of skill similar to BCSD also shows this to be the likely reason for its failure, as the temperature data do exhibit a near normal distribution, if bimodal in nature.

The bias in low-rate precipitation days can be observed in Figure 3.2, where the distribution of downscaled precipitation rates produces too few dry days and too many occurrences of values ranging from 1 – 7 mm/day. In effect, the transition from dry days to wet days is a problem for simple linear regression based methods, while previous studies have noted a marked improvement in non-linear downscaling methods, such as step-wise regression or cluster analysis (Maurer and Hidalgo 2008; Wilby et al. 2000). Figure 3.11 shows a comparison of precipitation intensity days for three of the downscaled GCMs. The RDA bars show similar or even greater frequencies of low precipitation day counts compared to GCMs, though higher values tend to better match observations. This pattern is seen in the dry day counts in Figure 3.12 as well. Both downscaling methods correct the GCM pattern of dry days to match those of the observations (more dry days in Winter, fewer in Summer), however the RDA counts are only about 60 percent of what they should be. This lack of dry days in favor of slightly wet days does little to affect the skill of the method on a day to day basis (with the RMSE being similar to

BCSD), however the wet bias adds up and causes significant differences at the monthly time scale, as evidenced by Figure 3.13, where the average total monthly precipitation over the verification period shows positive deviations from the observed values that often reach as high as 60 mm, with higher wet biases in Summer.

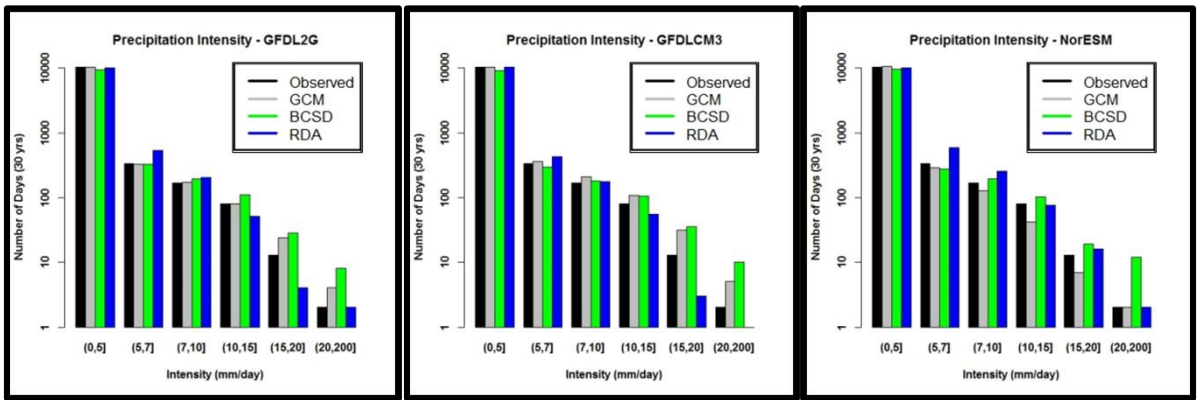

Figure 3.11: Distribution of precipitation intensities for three representative GCMs (GFDL2G, GFDLCM3 and NorESM) out of the full range of 17 downscaled for this study.

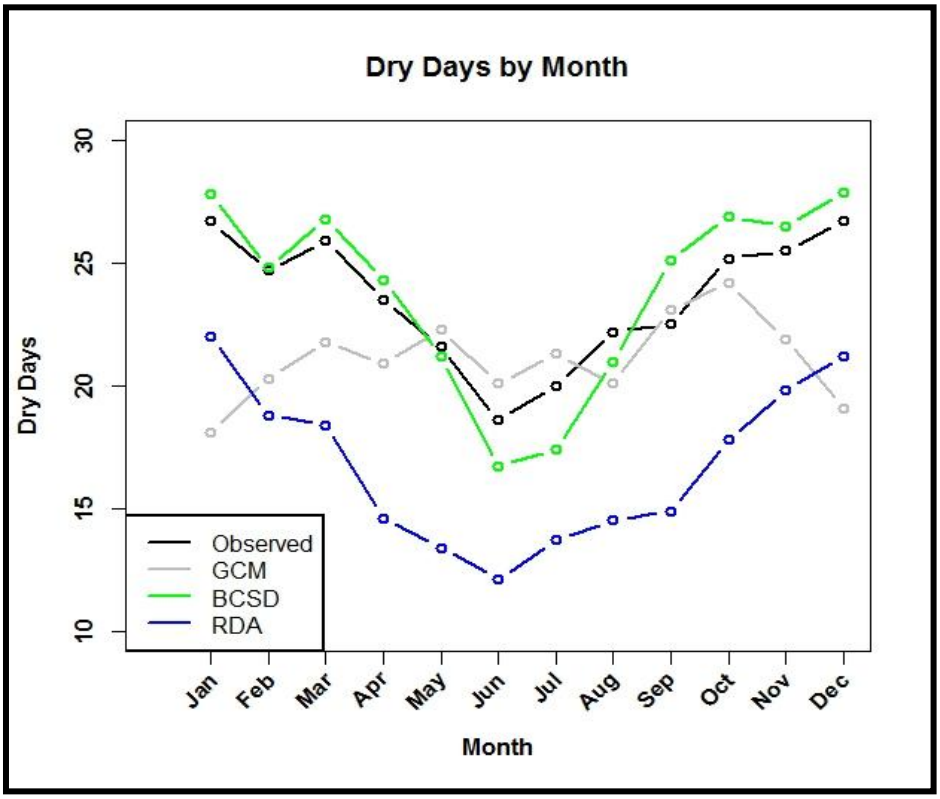

Figure 3.12: Monthly averaged dry days, averaged over all downscaled models for the verification period 1970 - 1999, as well as the observed and non-downscaled GCM dry days.

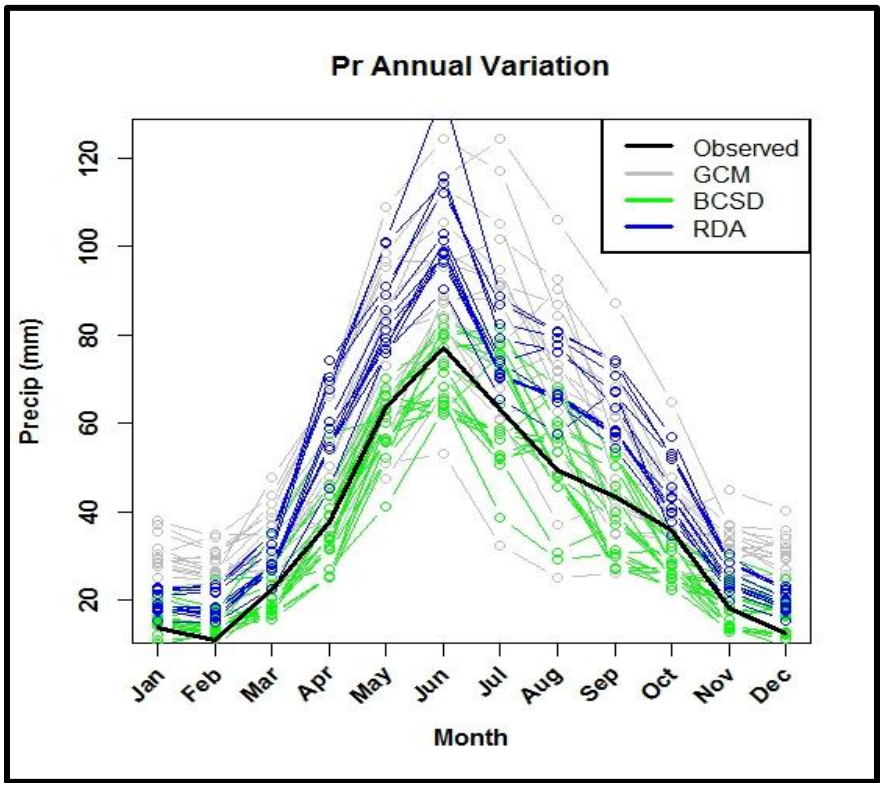

Figure 3.13: NGP averaged monthly total precipitation over the verification period 1970 – 1999 for observed, downscaled and GCM data.

3.5 IMPACT DIAGNOSTICS AND THE RCPS

One important application of downscaling is the ability of the downscaled data to effectively capture the impacts of climate change at local and regional spatial scales. These impacts range from extreme temperatures to prolonged periods without precipitation. In order to test whether these impacts are skillfully replicated, the statistics discussed in Section 2.3.5 are calculated for the downscaled and observational data and compared over the verification period. These same impact diagnostics are also calculated for the RCPs to ascertain how regional climate characteristics are expected to change in the 21st Century. As shown in the previous sections, the downscaled models show an improvement over the GCMs that force them, with the important

exception of RDA downscaled precipitation. For this reason, beyond comparing the verification results, RDA precipitation data will not be used to downscale the RCP projections.

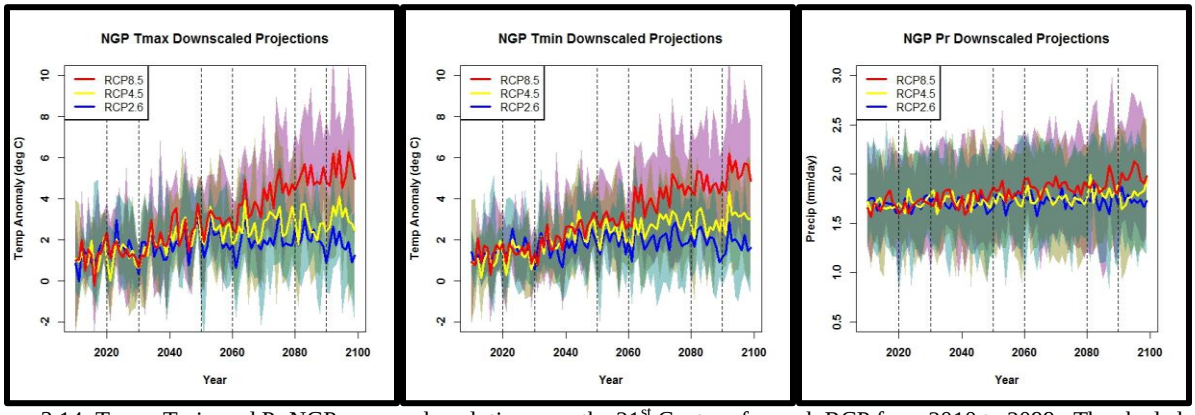

Figure 3.14: Tmax, Tmin and Pr NGP averaged evolution over the 21st Century for each RCP from 2010 to 2099. The shaded regions are the total model spread for each model, the colored lines are the means.

Figure 3.14 illustrates the overall changes to each of the predictands for the NGP as a whole over the full RCP period (2010 – 2099). Both maximum and minimum temperature increase by an average of 4°C by 2100 under the 8.5 RCP projection, with the l low estimate 2.6 projection resulting in less than 1°C rise for the NGP. Different models show different levels of regional warming, however, with the highest warmings of up to 8°C in the models with the highest sensitivities. As with the global projections, the highest RCP projection has the widest error margin due to the increasing uncertainties in forcings and feedbacks at higher temperature changes (IPCC 2014). It is noteworthy that the three RCPs warm at a similar rate, on average, until the 2060's, at which point they begin to diverge markedly. Precipitation shows little change over the 21st Century, with all three RCPs depicting little difference except for RCP 8.5 from the 2080s onward, where the daily average rate of precipitation shows a slight increase compared to the other projections. Table 3.3 shows the average of each of the predictands in the 2080s for BCSD and for temperature downscaled via RDA, as well as the average value of the observed and downscaled values for the verification period (1970 – 1999). The standard

deviations are also calculated. The variation in daily data values is very similar between the verification period and each of the RCPs for BCSD data. RDA data show increases in the standard deviation in the latter 21st Century of about 2°C for maximum temperature and 1°C for minimum temperature, though these are small changes at the daily time scale.

Table 3.3: A table showing the change in the mean and standard deviation for each of the predictands for both downscaling methods averaged over the 2080s, as well as the observed and downscaled average values for the verification period. RDA-based precipitation values were not included here due to the extreme wet bias, however the impact of those biases on the impact diagnostics can be seen in Table 3.4.

Summary Statistics					
	Verification Period	**RC Pathways**			
	OBS	**BCSD**	**2.6**	**4.5**	**8.5**
	Max Temperature				
mean	12.35	12.35	15.6	15.8	16.3
st. dev	13.61	13.93	13.7	14.3	13.9
	Min Temperature				
mean	-1.09	-1.38	1	1.8	2.4
st. dev	11.92	12.22	11.9	11.7	11.9
	Precipitation				
mean	1.3	1.25	1.31	1.32	1.33
st. dev	3.68	3.65	3.74	3.75	3.81
	OBS	**RDA**	**2.6**	**4.5**	**8.5**
	Max Temperature				
mean	12.35	12.25	14	14.3	14.9
st. dev	13.61	13.72	15	15.1	15.8
	Min Temperature				
mean	-1.09	-1.52	0	0.2	0.6
st. dev	11.92	12.14	12.7	12.9	13.2

Table 3.4 shows the average values of the impact diagnostics calculated from all the downscaled models and the parent GCMs over the verification period. The precipitation diagnostics for RDA models are included for comparison. Overall the downscaling brought the estimates of the impact diagnostics more in line with observed values, though to varying degrees of accuracy and only rarely within the envelope of the model derived 95 percent confidence intervals. For the data centered on the 1980s, extremely cold days outnumbered extremely warm

days, and frost days usually exceeded 50 percent of the year (> 182 days). Growing degree days with a base of 10°C varied between models, but averaged to just over 1100. The maximum length of dry days approached 3 months, while the percent of precipitation stemming from extreme rainfall or snowfall events was about 36 percent. GCMs produced more warm extremes than cold, and simulated too few frost days, maximum dry day lengths (underestimated by 50 percent), and wet extremes (both the number of days and the percent of total precipitation). BCSD models did relatively well in bringing these GCM estimates closer to observed values, though the value of cold extremes went up (to 28 days) rather than down as was needed. RDA proved more skillful at capturing temperature extremes, bringing the values very close to the observed values and, for heatwaves, within the 95 percent confidence interval. The wet bias in precipitation carried over into the RDA impact diagnostics, with the maximum dry days being less than 50 percent of observed and the percent of precipitation coming from extreme events calculated at 6 percent higher than observed, though the bias in the number of extreme events was similar to that of the BCSD models, as expected from the precipitation intensity distributions in Figure 3.11.

Table 3.4: A table of the impact diagnostics described in the text calculated for the verification period for the observations, GCM data and the downscaled models. The values are the averages of all downscaled models in which RCP data were also available, while the 95 percent confidence intervals are calculated from the same model spreads, as in Table 3.1. GDD refers to growing degree days while the wet events count the number of extreme precipitation events and the wet percentage is the percent of precipitation out of all precipitation coming from those exceptional events.

Impact Diagnostics: Verification Period							
	Heatwave	**Coldwave**	**Frost Days**	**GDD**	**Dry Day Max**	**Wet Event**	**Wet %**
OBS	16	23	199	1147	81	9	0.36
GCM	38.3	25.3	162.7	1011.7	44	4.3	0.32
Confidence Interval	17.7	3.6	1.2	99.7	0.9	0.3	0.3
BCSD	22.7	27.7	187.3	1167	72.3	8	0.33
Confidence Interval	4.6	1.9	1.2	14.4	1.5	0.5	0.2
RDA	14.3	21.7	184.7	1140.7	42.7	10.7	0.42
Confidence Interval	2.5	0.3	1.2	4.3	2.6	1.1	0.2

The impact diagnostics are shown to be better captured for the region as a whole by downscaling. These downscaling methods are then applied to the RCPs and compared on a decadal basis in Table 3.5. Each of the impacts changed as expected given the warming climate projected through the 21st Century. Heatwaves increased, with the consecutive hot days calculated at over three months under the highest RCP projection by the 2080s. Coldwaves decreased, reducing to only half the length of verification period coldwaves by the latter part of the century. Frost days are shown to go down by about 25 percent under the highest projections, meaning that, even being a colder interior continental climate, the annual percentage of frost days will go under 50 percent even under modest warming. This inexorably aids the upward trend in growing degree days, which increase by 13 – 63 percent under modest to high warming scenarios. RDA models show less temperature and temperature-based impact change than BCSD models, as is evident from Table 3.3 and Table 3.5. This may be a result of the increase in variability noted above, which would obscure the changes as well as indicate a relative lack of persistent weather systems that result in extremes such as heatwaves and cold snaps.

The NGP lies in a unique position geographically in that GCMs disagree on what trend the precipitation will take going into the latter 21st Century, however a slight majority favor a wetting of the Northern Plains and a drying of the South (Ojima et al. 2012). As stated previously, given the extreme wet bias of the RDA precipitation data, only the BCSD method was applied to RCP precipitation projections. The BCSD results are also cataloged in Table 3.5. Overall the wet events are shown to increase under high RCP projections by up to 5 such days, along with a corresponding increase in the percentage of precipitation coming from heavy events (44 percent compared to the modern 36 percent). This pattern of increasing precipitation extremes mirrors that of the global projections (IPCC 2014). The maximum consecutive dry

days are expected to, on average for the region, decrease. This decrease could mean that drought conditions will become less severe for the NGP, however such a conclusion does not take into account increased potential evapo-transpiration induced by increased warming.

The apparent divergence in the temperature impacts by the two methods is clear in the data, as shown above. These differences are also influenced by the sensitivity of the GCMs that are downscaled, in that the data from those GCMs that is used to drive the BCSD and RDA models over the RCP projection period change with the sensitivity of the GCM. This is shown in Table 3.6, where the model that is least sensitive all the downscaled models, the GFDL2G GCM downscaled via RDA, is compared to the model exhibiting the highest sensitivity, GFDLCM3 downscaled via BCSD. The changes in the temperature impacts are shown for the early and late 21st Century for the RCP 8.5 projection. It is clear that the high sensitivity model produces the largest changes in the diagnostics, with consecutive hot days reaching a full four months in the GFDLCM3 versus the small increase seen in the GFDL2G. Cold days, frost days and growing degree days also show the increased changes that occur in higher sensitivity models, with the GDD increasing by 70 percent versus only 27 percent under low sensitivity. This shows that the sensitivity of the GCMs that are downscaled plays a large role in what changes are shown to occur under climate change projections. This increases the uncertainty in model projections, even having undergone downscaling.

Table 3.5: The impact diagnostics calculated in Table 3.4, but now extended to the projected decades (2020s, 2050s and 2080s) under the three RCP scenarios (2.6, 4.5 and 8.5), averaged over both downscaling methods. Historical impact diagnostics come from the observations. As in Table 3.3, RDA precipitation is left out of the calculations due to its extreme wet bias.

Impact Diagnostics										
	Historical	**RCP 2.6**			**RCP 4.5**			**RCP 8.5**		
	1980s	***2020s***	***2050s***	***2080s***	***2020s***	***2050s***	***2080s***	***2020s***	***2050s***	***2080s***
	Bias Corrected Spatial Disaggregation									
Heatwave	16	41	46	43	46	64	65	52	85	102
Coldwave	23	17	16	10	16	12	11	15	12	10
Frost Days	186	175	168	172	175	164	161	172	161	142
GDD	1167	1305	1370	1367	1278	1443	1549	1303	1567	1944
	Redundancy Analysis									
Heatwave	16	26	29	27	29	40	41	33	54	64
Coldwave	23	15	12	12	14	12	11	16	12	10
Frost Days	186	172	165	169	172	161	158	169	158	139
GDD	1167	1279	1343	1340	1252	1414	1518	1277	1536	1905
	Bias Corrected Spatial Disaggregation									
Max Dry Days	81	77	73	74	78	74	72	75	73	68
Wet Event	9	11	12	11	11	12	12	12	13	14
Wet %	0.36	0.35	0.35	0.36	0.34	0.36	0.37	0.39	0.42	0.44

Table 3.6: A table comparing the early 21st Century temperature impact diagnostics to those in the late 21st Century to show the impact of GCM sensitivity on the results. The BCSD downscaled GFDLCM3 model was the most sensitive of the downscaled models, with the largest temperature changes, while the RDA downscaled GFDL2G showed the least.

Impact Diagnostic Sensitivity				
	GFDLCM3 - BCSD		**GFDL2G - RDA**	
	2020s	***2080s***	***2020s***	***2080s***
Heatwaves	58	120	12	21
Coldwaves	16	14	15	13
Frost Days	168	117	172	150
Growing Degree Days	1458	2479	1256	1590

CHAPTER 4 – DISCUSSION AND CONCLUSIONS

4.1 DOWNSCALING FOR THE NGP

Downscaling global climate model data for the Northern Great Plains is a necessary objective because of the increasing importance of the region in the global economy, particularly with regard to energy and agriculture. GCMs, by the constraints of available computational resources, must use low resolution grids that average out large amounts of information at the local level, making the output from those models only marginally useful to local stakeholders. Just as well, GCMs greatly reduce the variation in daily and annual temperatures, as well as between dry and wet extremes of precipitation, resulting in significant biases. Both dynamic downscaling and statistical downscaling seek to remedy these biases and provide information about local changes in climate. While dynamic downscaling allows for more detailed analysis of the causes and consequences of the forcings and feedbacks operating at the global and regional level, statistical methods for downscaling allow for easier implementation and are much less computationally expensive. This study focused on the use of two such statistical methods and compared them by testing their skill in representing independent data for the historical period. The appraisal of model skill in representing the climatic state and changes of the NGP in Chapter 3 showed that both Bias Corrected Spatial Disaggregation and Redundancy Analysis had the potential to offer useful local and regional information. Both methods showed comparable skill in capturing the temporal and spatial patterns in maximum and minimum temperature. BCSD also did well in capturing precipitation, though RDA exhibited a significant wet bias. The skill of the models was represented both in the set of validation statistics and impact diagnostics

chosen to assess the impact of climate changes on the weather extremes. The set of diagnostics was then applied to RCP projections through the 21st Century, shedding light on the potential impacts that different environmental sensitivities and emissions scenarios could cause. A general summary of the strengths and weaknesses of each method is shown in Table 4.1. Thus the main thrust of this work has been realized; the testing and application of downscaling models to the NGP region, which allowed for local scale changes important to local stake holders to become clearer and more accessible compared to what has been available from coarse GCMs.

Table 4.1: A table of the general conclusions regarding the use of BCSD and RDA in downscaling for the NGP, noting the advantages, disadvantages and possible improvements for both.

Final Results		
	Bias Corrected Spatial Disaggregation	**Redundancy Analysis**
Advantages:	*ease of implementation/adapatability*	*higher correlations (>0.9 for tmax/tmin, >0.3 for precip)*
	low monthly mean differnces (< 10% of GCMs)	*slightly lower RMSEs*
	higher skill in all 3 predictands	
	better representation of impact diagnostics	
Disadvantages:	*requires prior interpolation (greater uncertainty)*	*less adaptable*
	striations in spatial patterns/discontinuities	*significant reduction in st. dev. (-0.84 ± 0.26 mm)*
	slight dry bias in precipitation (negative mean diff.)	*significant wet bias (60% of obs. dry days)*
		large monthly mean diff. (14 ± 4 mm)
		associated wet bias in impact diagnostics
Possible Improvements:	*better ways of quantile mapping*	*better predictors - particularly for precip*
	better wet day adjustments	*new geographical zones for predictors*
		normalization of predictors/predictands

The skill of the BCSD models in capturing local climate variations and patterns followed closely to the bulk of previous work in which BCSD was applied to other regions and climatological situations. The slight dry bias the existed in the BCSD models was noted in previous studies (Ahmed et al. 2013; Abatzoglou and Brown 2012). The success of BCSD

compared to RDA is a departure from the results of other studies at first glance, wherein the more complex regression models better captured the predictands of interest in those studies (Trzaska and Schnarr 2014; Murphy 1999; Wilby et al. 1998). It should be noted, however, that those studies modified the regression models by using optimized predictors rather than the same GCM variables used for the BCSD method. The model used in the RDA process used the same predictors as the BCSD method, which combined with the limitations discussed in Chapter 3 to produce lower skill, overall, than BCSD.

The Redundancy Analysis models showed mixed results between the temperature and precipitation predictands. Both maximum and minimum temperatures were well represented by the RDA method, with comparable distributions to and higher correlations than BCSD models. Precipitation contained a significant wet bias in which extra low to moderate rate precipitation days were produced at the expense of dry days. This wet bias is largely the result of the regression model being ill-suited to capturing the non-Gaussian distribution of daily precipitation, as evidenced by the dichotomous ability of the method to capture the temperature predictands relatively well. This deviation from observations in RDA downscaled precipitation is broadly similar to the results of Themeßl et al. (2010) where similar regression methods failed to capture the values of precipitation. RDA models did, however, still maintain higher correlations with the observations and better constrained the extreme values of precipitation, also in accordance to previous studies (Themeßl et al. 2010; Huth 2004).

Downscaling of the RCPs showed many of the characteristics expected of broader scale climate change. More sensitive models produced, on average, more change in the impact statistics for the region. Many of the changes may not become noticeable until later on in the 21st Century, as the RCP temperatures do not diverge markedly until 2060 – 2070. Heatwave

lengths are projected to increase, growing degree days are expected to rise, and coldwaves and frost days are expected to decline dramatically, as expected under warming scenarios (Flato et al. 2013). Precipitation changes are not projected to be large for the region, with a small decrease in dry days projected, as documented in previous studies (Ojima et al. 2012). BCSD models show less change in heavy precipitation (both in percentage of precipitation and the number of events) than RDA, though both show non-negligible increases, particularly in the late 21st Century under high forcing scenarios.

4.2 FURTHER RESEARCH OPPORTUNITIES

The conclusions above are all made under the assumption that the models showed skill in recreating climate conditions at the local level. While the ability of the downscaling models to reproduce observed changes during the verification period bolsters this assumption, it is by no means given that the statistical relationships derived in the BCSD and RDA models would hold under a changing climate. RDA in particular is prone to the problem of stationarity described in Section 1.3, as it assumes the regression-based relationships it is based upon will stay the same, however there is little to back such an assumption that the GCM variables, particularly precipitation, will maintain a given statistical relationship to the observed values. There is also the problem, as faced directly by dynamical downscaling (see Section 1.2.2), of regional forcings. These include land use changes and aerosols, which are much more concentrated over the regions in which they are emitted than the well-mixed greenhouse gases. These regional forcings are assumed to change in a certain way by the global models that feed into the statistical ones used here, but which may vary in a different way than expected within the NGP region, resulting in changes unable to be captured in the downscaled models. As is clear, these methods,

while promising, are only a first step toward local scale climate information for the Northern Great Plains.

This work has shown that there is skill in using observations to correct GCM output for the Northern Great Plains, and further studies could show how best to improve these results. New research may focus on other methods than BCSD or RDA, particularly with regard to precipitation, or they may focus on improving the two methods applied here. As for other types of statistical models, non-linear methods have already been known to better constrain precipitation patterns. Dynamical downscaling is also an option for the region, as it has been applied to multiple other domains. BCSD could possibly be improved through the use of other types of cumulative distribution functions, though they often require a priori assumptions about the possible distribution of the variable sought. RDA, as applied in this study, was fairly simple and lacked a physical basis for its prediction. Making use of other predictors that have some physical connection to the variables, such as geopotential height or upper tropospheric humidity to predict precipitation, may make a difference in improving the skill of the model. It may also be possible to use BCSD and RDA in conjunction by bias correcting the output of the redundancy model.

APPENDICES

APPENDIX A
Data Distribution Table

Table 5.1: A table of the distributional data averaged for the observations, GCM data and for both downscaling methods. This provides the basis of Figures 3.1 and 3.2.

MODEL ENSEMBLE AVERAGE DISTRIBUTIONS														
Maximum Temperature Summary					**Minimum Temperature Summary**					**Precipitation Summary**				
North	***Obs***	***GCM***	***BCSD***	***RDA***	**North**	***Obs***	***GCM***	***BCSD***	***RDA***	**North**	***Obs***	***GCM***	***BCSD***	***RDA***
mean	9.18	8.13	8.83	8.8	***mean***	-3.32	-0.69	-3.86	-3.94	***mean***	1.17	1.44	1.14	1.71
stdv	14.15	13.1	14.65	14.28	***stdv***	12.39	11.92	12.81	12.69	***stdv***	3.58	3	3.6	2.63
min	-39.86	-37.27	-40.23	-48.49	***min***	-53.24	-45.87	-51.91	-63.81	***min***	0	0	0	0
10%	-10.76	-8.91	-12.18	-11.16	***10%***	-21.88	-17.78	-23.27	-22.42	***10%***	0	0	0	0
median	10.58	8.35	10.33	10.6	***median***	-1.38	0.55	-1.61	-1.48	***median***	0	0.29	0	0.78
90%	26.26	24.69	26.6	26.1	***90%***	10.99	13.45	10.93	10.73	***90%***	3.3	4.14	3.22	4.67
max	43.44	40.11	45.83	50.39	***max***	28.32	26.22	29.45	32.91	***max***	193.82	71.1	195.93	164.11
Middle					**Middle**					**Middle**				
mean	12.5	10.7	12.47	12.35	***mean***	-1.26	1.52	-1.56	-1.71	***mean***	1.23	1.71	1.19	1.64
stdv	13.87	12.83	14.13	13.94	***stdv***	12.07	11.34	12.34	12.23	***stdv***	3.4	3.61	3.38	2.65
min	-40.82	-34.74	-35.59	-45.06	***min***	-48.64	-42.74	-50.18	-60.08	***min***	0	0	0	0
10%	-6.47	-5.72	-7.09	-7.03	***10%***	-18.74	-14.18	-19.7	-19.17	***10%***	0	0	0	0
median	13.78	11.04	13.77	13.86	***median***	-0.21	2.08	-0.31	0	***median***	0	0.31	0	0.67
90%	29.47	27.09	29.87	29.49	***90%***	13.24	15.44	13.25	12.93	***90%***	3.5	5.01	3.37	4.61
max	45.73	43.05	47.98	52.02	***max***	32.81	28.41	33.3	31.86	***max***	183.8	82.55	278.39	179.66
South					**South**					**South**				
mean	15.38	13.5	15.76	15.6	***mean***	1.3	3.99	1.29	1.09	***mean***	1.51	2.12	1.42	1.95
stdv	12.81	12	13.01	12.94	***stdv***	11.3	10.52	11.52	11.5	***stdv***	4.06	4.41	3.98	3
min	-31.82	-29.05	-28.3	-36.96	***min***	-44.73	-36.9	-43.63	-51.17	***min***	0	0	0	0
10%	-1.85	-1.91	-1.85	-2.25	***10%***	-13.74	-9.72	-14.35	-14.66	***10%***	0	0	0	0
median	16.55	13.81	17.04	16.88	***median***	1.36	3.94	1.46	1.85	***median***	0	0.32	0	0.81
90%	31.11	29.05	31.81	31.58	***90%***	15.61	17.52	15.89	15.49	***90%***	4.55	6.41	4.25	5.51
max	45.39	44.35	50.41	55.48	***max***	34.3	29.56	34.21	33.74	***max***	233.52	86.6	212.2	193.91
NGP					**NGP**					**NGP**				
mean	12.35	10.78	12.35	12.25	***mean***	-1.09	1.61	-1.38	-1.52	***mean***	1.3	1.76	1.25	1.77
stdv	13.61	12.64	13.93	13.72	***stdv***	11.92	11.26	12.22	12.14	***stdv***	3.68	3.67	3.65	2.76
min	-37.5	-33.69	-34.71	-43.5	***min***	-48.87	-41.84	-48.57	-58.35	***min***	0	0	0	0
10%	-6.36	-5.51	-7.04	-6.81	***10%***	-18.12	-13.89	-19.11	-18.75	***10%***	0	0	0	0
median	13.64	11.07	13.71	13.78	***median***	-0.08	2.19	-0.15	0.12	***median***	0	0.31	0	0.75
90%	28.95	26.94	29.43	29.06	***90%***	13.28	15.47	13.36	13.05	***90%***	3.78	5.19	3.61	4.93
max	44.85	42.5	48.07	52.63	***max***	31.81	28.06	32.32	32.84	***max***	203.71	80.08	228.84	179.23

APPENDIX B
Verification Data Table

Table 5.2: A table of the regionally averaged verification statistic values, showing the raw GCM, BCSD and RDA downscaled values for all 17 GCMs used in this study. This is the basis of Table 3.1.

NGP Averaged Verification Data

Model	General Circulation Models			Bias Corrected Spatial Disaggregation				Redundancy Analysis			
Precipitation	*Corr*	*Mn Diff*	*St. Dev Diff*	*Corr*	*Mn Diff*	*St. Dev Diff*	*RMSE*	*Corr*	*Mn Diff*	*St. Dev Diff*	*RMSE*
BCC	0.145	6.062	-0.613	0.353	-1.695	-0.075	5.017	0.598	11.387	-1.206	4.272
BNU	0.238	22.355	-0.436	0.397	2.069	0.251	5.246	0.564	-1.818	-0.653	4.611
CAN	0.111	-7.369	-0.535	0.239	-0.439	0.021	5.142	0.607	11.693	-1.231	4.252
CMCCCESM	0.068	34.749	0.292	0.438	-0.735	0.17	5.187	0.575	6.146	-0.955	4.431
FGOALS	0.098	12.432	-0.327	0.395	-0.375	-0.03	5.073	0.609	10.733	-1.304	4.206
GFDL2G	0.2	15.283	0.122	0.423	-1.844	-0.008	5.056	0.583	23.328	-0.327	4.834
GFDL2M	0.189	19.626	0.28	0.408	-2.922	-0.125	4.981	0.574	23.229	-0.301	4.859
GFDLCM3	0.193	22.283	0.188	0.437	-2.195	-0.042	5.036	0.575	20.511	-0.523	4.709
IPSLLR	0.12	7.338	-0.338	0.415	-1.265	0.021	5.094	0.585	15.658	-0.862	4.474
IPSLMR	0.153	4.835	-0.058	0.313	-1.15	0.016	5.114	0.61	11.071	-1.241	4.246
MIROC5	0.247	15.897	0.699	0.383	-0.653	0.034	5.128	0.604	17.549	-0.782	4.551
MIROCCHEM	0.163	10.466	-0.423	0.289	-1.124	-0.003	5.125	0.617	10.474	-1.334	4.2
MIROCESM	0.154	9.121	-0.494	0.379	-3.244	-0.126	5.001	0.615	11.847	-1.267	4.226
MPILR	0.215	17.972	0.875	0.359	-1.132	0.193	5.229	0.579	12.296	-0.693	4.606
MPIMR	0.259	15.217	0.683	0.392	-0.799	-0.002	5.078	0.602	13.972	-1.066	4.367
MRI	0.351	18.991	0.618	0.381	0.305	0.107	5.142	0.585	17.165	-0.734	4.584
NorESM	0.229	9.746	-0.798	0.42	-2.18	-0.061	5.003	0.571	27.636	0.095	5.14
Max Temperature	*Corr*	*Mn Diff*	*St. Dev Diff*	*Corr*	*Mn Diff*	*St. Dev Diff*	*RMSE*	*Corr*	*Mn Diff*	*St. Dev Diff*	*RMSE*
BCC	0.931	-1.79	0.611	0.933	0.268	0.342	9.006	0.936	0.007	0.076	8.502
BNU	0.94	-1.53	0.189	0.937	0.298	0.094	8.616	0.954	0.117	-0.126	8.366
CAN	0.933	4.161	2.461	0.934	0.286	0.425	8.657	0.957	0.122	-0.131	8.287
CMCCCESM	0.916	-3.339	-1.614	0.936	-0.043	0.268	8.884	0.966	0.065	-0.796	7.41
FGOALS	0.926	-5.953	-0.064	0.93	0.66	0.065	8.705	0.962	0.105	-0.25	8.118
GFDL2G	0.905	-2.6	-1.492	0.917	0.256	0.11	9.439	0.944	-0.087	1.053	9.706
GFDL2M	0.914	-1.935	-1.998	0.921	0.46	0.164	9.339	0.943	0.238	0.827	9.751
GFDLCM3	0.917	-2.002	-2.908	0.921	0.376	-0.173	9.311	0.939	-0.22	0.831	9.643
IPSLLR	0.91	-3.574	-1.641	0.913	0.24	0.357	9.274	0.956	0.052	0.236	8.746
IPSLMR	0.926	-2.257	-0.754	0.923	-0.019	0.4	9.098	0.951	0.004	0.643	9.233
MIROC5	0.931	-0.042	0.803	0.936	-0.609	0.421	8.512	0.948	0.011	-0.161	8.332
MIROCCHEM	0.907	0.276	-0.478	0.915	0.295	0.401	8.978	0.952	-0.02	-0.089	8.354
MIROCESM	0.917	0.514	-0.489	0.926	-0.206	0.449	8.8	0.949	-0.087	-0.041	8.385
MPILR	0.907	-1.798	-2.568	0.913	0.793	0.026	9.669	0.95	-0.052	0.053	8.405
MPIMR	0.91	-1.839	-2.563	0.919	0.064	0.276	9.532	0.955	-0.028	-0.285	8.234
MRI	0.911	-2.68	-2.53	0.92	0.297	-0.053	9.315	0.953	0.157	0.037	8.593
NorESM	0.924	-2.113	-0.664	0.93	-0.103	0.295	8.907	0.951	0.066	0.911	9.896
Min Temperature	*Corr*	*Mn Diff*	*St. Dev Diff*	*Corr*	*Mn Diff*	*St. Dev Diff*	*RMSE*	*Corr*	*Mn Diff*	*St. Dev Diff*	*RMSE*
BCC	0.915	1.969	0.03	0.93	-0.104	0.309	8.353	0.933	-0.155	0.075	7.572
BNU	0.921	2.222	-0.036	0.929	-0.068	0.26	8.335	0.956	-0.478	-0.044	7.353
CAN	0.92	5.476	1.07	0.927	0.006	0.33	8.319	0.96	-0.196	-0.184	7.224
CMCCCESM	0.92	1.699	-1.07	0.938	-0.375	0.32	7.957	0.967	-0.368	-0.721	6.462
FGOALS	0.921	-1.028	0.725	0.933	0.068	0.284	8.093	0.962	-0.276	-0.279	7.155
GFDL2G	0.907	2.77	-1.568	0.919	-0.125	0.081	8.547	0.953	-0.493	0.556	8.395
GFDL2M	0.922	4.216	-2.221	0.932	-0.151	0.207	8.365	0.947	-0.574	0.898	8.611
GFDLCM3	0.907	3.025	-2.631	0.917	-0.096	-0.064	8.677	0.947	-0.287	0.555	8.411
IPSLLR	0.906	1.58	-0.763	0.916	-0.369	0.582	8.596	0.954	-0.426	0.222	7.663
IPSLMR	0.92	2.385	-0.434	0.92	-0.432	0.529	8.554	0.952	-0.436	0.219	7.496
MIROC5	0.934	4.574	-0.011	0.934	-1.039	0.549	7.989	0.95	-0.388	0.392	8.168
MIROCCHEM	0.922	4.765	-0.859	0.929	-0.294	0.402	8.026	0.951	-0.364	0.064	7.381
MIROCESM	0.923	4.942	-0.908	0.93	-0.665	0.508	8.041	0.95	-0.595	0.193	7.459
MPILR	0.915	3.108	-1.163	0.926	0.215	0.081	8.309	0.947	-0.223	-0.074	7.397
MPIMR	0.914	2.933	-0.991	0.928	-0.424	0.428	8.375	0.95	-0.426	0.648	8.52
MRI	0.911	3.209	-1.412	0.923	-0.15	0.381	8.389	0.953	-0.526	0.343	7.891
NorESM	0.911	1.727	0.416	0.926	-0.432	0.322	8.349	0.95	-0.525	0.964	8.8

APPENDIX C
Demonstration of Downscaling Functions – R Script

```
##This is a test script for the main functions used to downscale##
##Not the full script!!!!##

#Load package(s) needed to perform downscaling:
library(qmap)
library(climates)
library(vegan)

#Read in historical GCM data for BCSD [NGP specific]:
Hist <- read.csv("N:\\Kirilenko_Coburn_ModelOutput\\Thesis\\Modeled\\GFDL2G\\GFDL2GnorthTmax.csv")
Hist <- Hist[,2:length(Hist[1,])]
Hist.bcsd <- as.matrix(Hist)-273.14

#Read in historical GCM data for RDA [North America]:
Hist <- read.csv("N:\\Kirilenko_Coburn_ModelOutput\\Thesis\\Modeled\\GFDL2G\\GFDL2GnaTmax.csv")
Hist <- Hist[,2:length(Hist[1,])]
Hist.rda <- as.matrix(Hist)-273.14

#Interpolate BCSD GCM data:
DIMs <- dim(Hist.bcsd)
dim(Hist.bcsd) <- c(DIMs[1],DIMs[2]/2,DIMs[2]/6)
tx <- seq(1,6,length=120)
ty <- seq(1,2,length=32)
x2 <- seq(1,6)
y2 <- seq(1,2)
Hist.bcsd <- apply(Hist.bcsd, MARGIN=c(1), FUN=interp2grid, xout=ty,
yout=tx, xin=y2, yin=x2, type=1)
Hist.bcsd <- t(Hist.bcsd)

#Import the observed data
#Remove leap years
#Remove the row names:
Data.O <- read.csv("N:\\Kirilenko_Coburn_ModelOutput\\Thesis\\Observations\\TmaxNorth.csv")
Data.O <- Data.O[,2:length(Data.O[1,])]
Data.O <- as.matrix(Data.O)
Date <- seq(as.Date("1950/1/1"), as.Date("1999/12/31"), by = c("day"))
MDdate <- (format(Date, "%m-%d"))
rownames(Data.O) <- MDdate
Data.O <- subset(Data.O, rownames(Data.O) != "02-29")
rownames(Data.O) <- NULL

#Do quantile mapping and apply function to GCM data:
q.map <- fitQmap(Data.O, Hist.bcsd, method=c("QUANT"), wet.day=F)
Validation.bcsd <- doQmap(Hist.bcsd,q.map)

#Do RDA regression and apply function to GCM data
#Real script separates data by month, but this is just a demo:
rda.mod <- rda(Data.O,Hist.rda,scale=FALSE)
Validation.rda <- predict(rda.mod,Hist.rda,type="response")

#Quick check of performance by graphing the distributions of data:
plot(density(Data.O, na.rm=T), col="black", lwd=2)
lines(density(Validation.bcsd, na.rm=T), col="green", lwd=2)
lines(density(Validation.rda, na.rm=T), col="blue", lwd=2)

##END OF SCRIPT
```

BIBLIOGRAPHY

Abatzoglou, John T. and Timothy J. Brown. 2012. "A Comparison of Statistical Downscaling Methods Suited for Wildfire Applications." *International Journal of Climatology* 32 (5): 772-780. doi:10.1002/joc.2312.

Ahmed, Kazi Farzan, Guiling Wang, John Silander, Adam M. Wilson, Jenica M. Allen, Radley Horton, and Richard Anyah. 2013. "Statistical Downscaling and Bias Correction of Climate Model Outputs for Climate Change Impact Assessment in the US Northeast." *Global and Planetary Change* 100: 320-332.

Benestad, R. E. 2004. "Empirical-Statistical Downscaling in Climate Modeling." *Eos, Transactions American Geophysical Union* 85 (42): 417-422. doi:10.1029/2004EO420002.

Benestad, Rasmus E., and Rasmus E. Benestad. 2008. *Empirical-Statistical Downscaling*, edited by Inger Hanssen-Bauer, Deliang Chen. New Jersey: New Jersey : World Scientific Pub Co Inc.

Bindoff, N.L., P.A. Stott, K.M. AchutaRao, M.R. Allen, N. Gillett, D. Gutzler, K. Hansingo, G. Hegerl, Y. Hu, S. Jain, I.I. Mokhov, J. Overland, J. Perlwitz, R. Sebbari and X. Zhang, 2013: Detection and Attribution of Climate Change: from Global to Regional. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Bukovsky, Melissa S. 2012. "Temperature Trends in the NARCCAP Regional Climate Models." *Journal of Climate* 25 (11): 3985-3991.

Bukovsky, Melissa S. and David J. Karoly. 2011. "A Regional Modeling Study of Climate Change Impacts on Warm-Season Precipitation in the Central United States*." *Journal of Climate* 24 (7): 1985-2002.

Busuioc, Aristita, Deliang Chen, and Cecilia Hellström. 2001. "Performance of Statistical Downscaling Models in GCM Validation and Regional Climate Change Estimates: Application for Swedish Precipitation." *International Journal of Climatology* 21 (5): 557-578. doi:10.1002/joc.624.

Cannon, Alex J. and Paul H. Whitfield. 2002. "Downscaling Recent Streamflow Conditions in British Columbia, Canada using Ensemble Neural Network Models." *Journal of Hydrology* 259 (1–4): 136-151. doi:http://dx.doi.org/10.1016/S0022-1694(01)00581-9.

Cowtan, Kevin and Robert G. Way. 2014. "Coverage Bias in the HadCRUT4 Temperature Series and its Impact on Recent Temperature Trends." *Quarterly Journal of the Royal Meteorological Society* 140 (683): 1935-1944. doi:10.1002/qj.2297.

Crook, Julia A., Piers M. Forster, and Nicola Stuber. 2011. "Spatial Patterns of Modeled Climate Feedback and Contributions to Temperature Response and Polar Amplification." *Journal of Climate* 24 (14).

Dixon, K., K. Hayhoe, J. Lanzante, A. Stoner, and A. Radhakrishnan. 2013.*Examining the Stationarity Assumption in Statistical Downscaling of Climate Projections: Is Past Performance an Indication of Future Results?, Paper Presented at the 93rd American Meteorological Society Annual Meeting, Austin, Tex.*

Driouech, Fatima, Michel Déqué, and Emilia Sánchez-Gómez. 2010. "Weather regimes—Moroccan Precipitation Link in a Regional Climate Change Simulation." *Global and Planetary Change* 72 (1): 1-10.

Eby, Michael, Andrew J. Weaver, K. Alexander, K. Zickfeld, A. Abe-Ouchi, AA Cimatoribus, E. Crespin, SS Drijfhout, NR Edwards, and AV Eliseev. 2013. "Historical and Idealized Climate Model Experiments: An Intercomparison of Earth System Models of Intermediate Complexity." .

Edwards, Paul N. 2011. "History of Climate Modeling." *Wiley Interdisciplinary Reviews: Climate Change* 2 (1): 128-139.

Fan, Yun and Huug van den Dool. 2008. "A Global Monthly Land Surface Air Temperature Analysis for 1948?Present." *Journal of Geophysical Research: Atmospheres* 113 (D1): -D01103. doi:10.1029/2007JD008470.

Feser, Frauke, Burkhardt Rockel, Hans von Storch, Jörg Winterfeldt, and Matthias Zahn. 2011. "Regional Climate Models Add Value to Global Model Data: A Review and Selected Examples." *Bulletin of the American Meteorological Society* 92 (9).

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason and M. Rummukainen, 2013: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge.

Fyfe, John C., Nathan P. Gillett, and Francis W. Zwiers. 2013. "Overestimated Global Warming Over the Past 20 Years." *Nature Climate Change* 3 (9): 767-769.

Giorgi, Filippo. 1990. "Simulation of Regional Climate using a Limited Area Model Nested in a General Circulation Model." *Journal of Climate* 3 (9): 941-963.

Goodess, C. 2005. "STARDEX–Downscaling Climate Extremes." *UEA, Norwich*.

Gutowski Jr, William J., Steven G. Decker, Rodney A. Donavon, Zaitao Pan, Raymond W. Arritt, and Eugene S. Takle. 2003. "Temporal-Spatial Scales of Observed and Simulated Precipitation in Central US Climate." *Journal of Climate* 16 (22): 3841-3847.

Heinrich, Georg, Andreas Gobiet, and Thomas Mendlik. 2014. "Extended Regional Climate Model Projections for Europe Until the Mid-Twentyfirst Century: Combining ENSEMBLES and CMIP3." *Climate Dynamics* 42 (1-2): 521-535.

Heyen, Hauke, Eduardo Zorita, and Hans Von Storch. 1996. "Statistical Downscaling of Monthly Mean North Atlantic air-pressure to Sea Level Anomalies in the Baltic Sea." *Tellus A* 48 (2): 312-323.

Huth, Radan. 2004. "Sensitivity of Local Daily Temperature Change Estimates to the Selection of Downscaling Models and Predictors." *Journal of Climate* 17 (3): 640-652.

———. 1999. "Statistical Downscaling in Central Europe: Evaluation of Methods and Potential Predictors." *Climate Research* 13 (2): 91-101.

IPCC, 2014: Summary for policymakers. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32.

Jakob Themeßl, Matthias, Andreas Gobiet, and Armin Leuprecht. 2011. "Empirical-statistical Downscaling and Error Correction of Daily Precipitation from Regional Climate Models." *International Journal of Climatology* 31 (10): 1530-1544.

Kaper, Hans and Hans Engler. 2013. *Mathematics and Climate* Siam.

Karl, Thomas R. and Jerry M. Melillo. 2009. *Global Climate Change Impacts in the United States* Cambridge University Press.

Kendon, Elizabeth J., Nigel M. Roberts, Catherine A. Senior, and Malcolm J. Roberts. 2012. "Realism of Rainfall in a very High-Resolution Regional Climate Model." *Journal of Climate* 25 (17).

Knutti, Reto and Jan Sedlácek. 2013. "Robustness and Uncertainties in the New CMIP5 Climate Model Projections." *Nature Climate Change* 3 (4): 369-373.

Kosaka, Yu and Shang-Ping Xie. 2013. "Recent Global-Warming Hiatus Tied to Equatorial Pacific Surface Cooling." *Nature* 501 (7467): 403-407.

Landman, Willem A., Simon J. Mason, Peter D. Tyson, and Warren J. Tennant. 2001. "Statistical Downscaling of GCM Simulations to Streamflow." *Journal of Hydrology* 252 (1): 221-236.

Landman, Willem A. and Warren J. Tennant. 2000. "Statistical Downscaling of Monthly Forecasts." *International Journal of Climatology* 20 (13): 1521-1532.

Laprise, René, Dragana Kornic, Maja Rapaić, Leo Šeparović, Martin Leduc, Oumarou Nikiema, Alejandro Di Luca, Emilia Diaconescu, Adelina Alexandru, and Philippe Lucas-Picher. 2012. "Considerations of Domain Size and Large-Scale Driving for Nested Regional Climate Models: Impact on Internal Variability and Ability at Developing Small-Scale Details." In *Climate Change*, 181-199: Springer.

Legendre, Pierre and Loic FJ Legendre. 2012. *Numerical Ecology*. Vol. 24 Elsevier.

Lemons, Rebecca, Andrea Hewitt, Gehendra Kharel, Cherie New, Andrei Kirilenko, and Xiaodong Zhang. 2012. "Evaluation of Satellite-Derived Agro-Climate Variables in the Northern Great Plains of the United States." *Geocarto International* 27 (8): 613-626. doi:10.1080/10106049.2011.653408. http://ezproxy.library.und.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=82754424&site=ehost-live&scope=site.

Li, Gen and Shang-Ping Xie. 2014. "Tropical Biases in CMIP5 Multimodel Ensemble: The Excessive Equatorial Pacific Cold Tongue and Double ITCZ Problems*." *Journal of Climate* 27 (4): 1765-1780.

Loikith, Paul C., Duane E. Waliser, Huikyo Lee, J. David Neelin, Benjamin R. Lintner, Seth McGinnis, Linda O. Mearns, and Jinwon Kim. 2015. "Evaluation of Large-Scale Meteorological Patterns Associated with Temperature Extremes in the NARCCAP Regional Climate Model Simulations." *Climate Dynamics*: 1-18.

Luo, Lifeng, Ying Tang, Shiyuan Zhong, Xindi Bian, and Warren E. Heilman. 2013. "Will Future Climate Favor More Erratic Wildfires in the Western United States?" *Journal of Applied Meteorology and Climatology* 52 (11): 2410-2417.

Lutgens, Frederick K., and Edward J. Tarbuck. *The Atmosphere: An Introduction to Meteorology*, 10th ed. Upper Saddle River: Pearson Prentice Hall, 2007.

Martinez, Yosvany, Wei Yu, and Hai Lin. 2013. "A New statistical–dynamical Downscaling Procedure Based on EOF Analysis for Regional Time Series Generation." *Journal of Applied Meteorology and Climatology* 52 (4): 935-952.

Maurer, EP and HG Hidalgo. 2008. "Utility of Daily Vs. Monthly Large-Scale Climate Data: An Intercomparison of Two Statistical Downscaling Methods." *Hydrology and Earth System Sciences* 12 (2): 551-563.

Maurer, EP, AW Wood, JC Adam, DP Lettenmaier, and B. Nijssen. 2002. "A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States*." *Journal of Climate* 15 (22): 3237-3251.

McGuffie, K. and A. Henderson-Sellers. c2005. *Climate Modelling Primer /*. 3rd ed. ed. Chichester, West Sussex, England ; Hoboken, NJ : John Wiley,.

Mearns, LO, Steve - Sain, LR - Leung, MS - Bukovsky, Seth - McGinnis, S. - Biner, Daniel - Caya, RW - Arritt, William - Gutowski, and E. - Takle. - *Climate Change Projections of the North American Regional Climate Change Assessment Program (NARCCAP)* - Springer Netherlands.

Miller, Perry, Will Lanier, and Stu Brandt. 2001. "Using Growing Degree Days to Predict Plant Stages." *Montana State University, USA.Extension Service*.

Murphy, James. 1999. "An Evaluation of Statistical and Dynamical Techniques for Downscaling Local Climate." *Journal of Climate* 12 (8): 2256-2284.

Ojima, Dennis S., Jean Steiner, Shannon McNeeley, Amber Childress, and Karen Cozzetto. 2012. "Great Plains Climate Assessment Technical Report." .

Peters, Glen P., Gregg Marland, Corinne Le Quéré, Thomas Boden, Josep G. Canadell, and Michael R. Raupach. 2012. "Rapid Growth in CO2 Emissions After the 2008-2009 Global Financial Crisis." *Nature Climate Change* 2 (1): 2-4.

Rohli, Robert V. 2008. *Climatology*, edited by Anthony J. Vega. Sudbury, Mass.: Sudbury, Mass. : Jones and Bartlett Publishers.

Rummukainen, Markku. 2010. "State-of-the-art with Regional Climate Models." *Wiley Interdisciplinary Reviews: Climate Change* 1 (1): 82-96.

Sailor, D. J., T. Hu, X. Li, and J. N. Rosen. 2000. "A Neural Network Approach to Local Downscaling of GCM Output for Assessing Wind Power Implications of Climate Change." *Renewable Energy* 19 (3): 359-378. doi:http://dx.doi.org/10.1016/S0960-1481(99)00056-7.

Schmidli, Jürg, Christoph Frei, and Pier Luigi Vidale. 2006. "Downscaling from GCM Precipitation: A Benchmark for Dynamical and Statistical Downscaling Methods." *International Journal of Climatology* 26 (5): 679-689.

Schoof, J. T. and S. C. Pryor. 2001. "Downscaling Temperature and Precipitation: A Comparison of Regression-Based Methods and Artificial Neural Networks." *International Journal of Climatology* 21 (7): 773-790. doi:10.1002/joc.655.

Sheffield, Justin, Suzana J. Camargo, Rong Fu, Qi Hu, Xianan Jiang, Nathaniel Johnson, Kristopher B. Karnauskas, Seon Tae Kim, Jim Kinter, Sanjiv Kumar, Baird Langenbrunner, Eric Maloney, Annarita Mariotti, Joyce E. Meyerson, J. David Neelin, Sumant Nigam, Zaitao Pan, Alfredo Ruiz-Barradas, Richard Seager, Yolande L. Serra, De-Zheng Sun, Chunzai Wang, Shang-Ping Xie, Jin-Yi Yu, Tao Zhang, and Ming Zhao, 2013: - *North American Climate in CMIP5 Experiments. Part II: Evaluation of Historical Simulations of Intraseasonal to Decadal Variability*. - Journal of Climate. - American Meteorological Society. doi:- 10.1175/JCLI-D-12-00593.1.

Shepard, Donald S. 1984. "Computer Mapping: The SYMAP Interpolation Algorithm." In *Spatial Statistics and Models*, 133-145: Springer.

Schmith, Torben. 2008. "Stationarity of Regression Relationships: Application to Empirical Downscaling." *Journal of Climate* 21 (17): 4529-4537.

Shukla, Jagadish, T. DelSole, M. Fennessy, J. Kinter, and D. Paolino. 2006. "Climate Model Fidelity and Projections of Climate Change." *Geophysical Research Letters* 33 (7): L07702.

Simon, Thorsten, Andreas Hense, Buda Su, Tong Jiang, Clemens Simmer, and Christian Ohlwein. 2013. "Pattern-Based Statistical Downscaling of East Asian Summer Monsoon Precipitation." *Tellus A* 65.

Solomon, Susan, Dahe Qin, Martin Manning, Melinda Marquis, Kristen Averyt, Melinda M. B. Tignor, Henry L. Miller, Zhenlin Chen, Intergovernmental Panel on Climate Change, and Intergovernmental Panel on Climate Change. 2007. *Climate Change 2007 : The Physical Science Basis : Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge ; New York: Cambridge University Press. http://www.ipcc.ch/ipccreports/ar4-wg1.htm.

Storch, Hans von, Eduardo Zorita, and Ulrich Cubasch. 1993. "Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall in Wintertime." *Journal of Climate* 6 (6): 1161-1171.

Storch, Hans von. 1999. "On the use of “inflation” in Statistical Downscaling." *Journal of Climate* 12 (12): 3505-3506.

Tatli, Hasan, H. Nüzhet Dalfes, and Ş. Sibel Menteş. 2004. "A Statistical Downscaling Method for Monthly Total Precipitation Over Turkey." *International Journal of Climatology* 24 (2): 161-180.

Team, R. Core. 2000.*R Language Definition*.

Themeßl, Jakob. Matthias, Andreas Gobiet, and Armin Leuprecht. 2011. "Empirical-statistical Downscaling and Error Correction of Daily Precipitation from Regional Climate Models." *International Journal of Climatology* 31 (10): 1530-1544.

Trzaska, Sylwia and Emilie Schnarr. 2014. "A Review of Downscaling Methods for Climate Change Projections." *Center for International Earth Science Network*. Accessed March 2015. http://www.ciesin.org/publications.html

Tyler, D. E., 1982: On the optimality of the simultaneous redundancy transformations. Psychometrika, 47, 77–86.

U.S. Environmental Protection Agency. 2014. Climate change indicators in the United States, 2014. Third edition. EPA 430-R-14-004. www.epa.gov/climatechange/indicators.

Van Vuuren, Detlef P., Jae Edmonds, Mikiko Kainuma, Keywan Riahi, Allison Thomson, Kathy Hibbard, George C. Hurtt, Tom Kram, Volker Krey, and Jean-Francois Lamarque. 2011. "The Representative Concentration Pathways: An Overview." *Climatic Change* 109 (1-2): 5-31.

Vasiliades, L., A. Loukas, and G. Patsonas. 2009. "Evaluation of a Statistical Downscaling Procedure for the Estimation of Climate Change Impacts on Droughts." *Natural Hazards and Earth System Science* 9 (3): 879-894.

Vrac, M., P. Marbaix, D. Paillard, and P. Naveau. 2007. "Non-Linear Statistical Downscaling of Present and LGM Precipitation and Temperatures Over Europe." *Climate of the Past* 3 (4): 669-682.

Vrac, M., ML Stein, K. Hayhoe, and X-Z Liang. 2007. "A General Method for Validating Statistical Downscaling Methods Under Future Climate Change." *Geophysical Research Letters* 34 (18).

Wang, Yuqing, L. Ruby Leung, John L. McGREGOR, Dong-Kyou Lee, Wei-Chyung Wang, Yihui Ding, and Fujio Kimura. 2004. "Regional Climate Modeling: Progress, Challenges, and Prospects." *気象集誌* 82 (6): 1599-1628.

Wang, Fang, Song Yang, Wayne Higgins, Qiaoping Li, and Zhiyan Zuo. 2013. "Long-Term Changes in Total and Extreme Precipitation Over China and the United States and their Links to Oceanic+Atmospheric Features." *International Journal of Climatology*: n/a-n/a. doi:10.1002/joc.3685.

WASA, 1998. “Changing waves and storms in the Northeast Atlantic?” *Bull. Amer. Met. Soc*., 79, 741–760

Weart, Spencer R. 2008. *The Discovery of Global Warming: Revised and Expanded Edition* Harvard University Press.

Wilby, RL, SP Charles, E. Zorita, B. Timbal, P. Whetton, and LO Mearns. 2004. "Guidelines for use of Climate Scenarios Developed from Statistical Downscaling Methods." .

Wilby, Robert L., Lauren E. Hay, William J. Gutowski, Raymond W. Arritt, Eugene S. Takle, Zaitao Pan, George H. Leavesley, and Martyn P. Clark. 2000. "Hydrological Responses to Dynamically and Statistically Downscaled Climate Model Output." *Geophysical Research Letters* 27 (8): 1199-1202.

Wilby, Robert L., TML Wigley, D. Conway, PD Jones, BC Hewitson, J. Main, and DS Wilks. 1998. "Statistical Downscaling of General Circulation Model Output: A Comparison of Methods." *Water Resources Research* 34 (11): 2995-3008.

Wilby, Robert L. 1997. "Non-Stationarity in Daily Precipitation Series: Implications for GCM Down-Scaling using Atmospheric Circulation Indices." *International Journal of Climatology* 17 (4): 439-454.

Wilks, Daniel S. 2011. *Statistical Methods in the Atmospheric Sciences*. Vol. 100 Academic press.

Wood, Andrew W., Lai R. Leung, V. Sridhar, and DP Lettenmaier. 2004. "Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs." *Climatic Change* 62 (1-3): 189-216.