ABSTRACT

DAS BHOWMIK, RAJARSHI. Reducing Downscaling & Model Uncertainties in CMIP5 Decadal Hindcasts/ Projections. (Under the direction of Dr. Sankarasubramanian Arumugam).

General Circulation Model projections are important to understand the climate change and its impact on land surface attributes. However, GCM projections exhibit model and scenario uncertainties in the near-term and the long-term predictions. Resolution of GCMs are coarser compared to the resolution of climate forcing of hydrologic models. Hence, GCM projections require post-processing. In this study, we worked with two post-processing approaches, downscaling and model uncertainty reduction. We proposed a multivariate statistical downscaling approach, based on Asynchronous Canonical Correlation Analysis (ACCA) that has the potential to preserve cross-correlation between multiple variables during downscaling. ACCA performs better than the univariate statistical downscaling approach in estimating cross-correlation between precipitation and temperature. Joint likelihood between precipitation and temperature is also improved by ACCA for months when more grid points exhibit statistical significance in the cross-correlation. In our second problem, we extended our first work to analytically investigate the limitations of univariate downscaling in estimating the cross-correlation. We assumed a linear regression model for univariate downscaling and found that univariate approach lacks any potential to estimate the cross-correlation in downscaled variables. Further, we proposed a framework, based on Fisher z-transformation to form an allowable range for the raw GCM/downscaled cross-correlation. The framework is applied for univariate downscaled products over the contiguous US and results shows that 40-50% grid points would fail to preserve the cross-correlation if univariate downscaling is applied on the GCM historical runs. Finally, we proposed an optimal weighting approach, based on model performance to reduce model uncertainty in climate change projections. Results show, spatial

bias in estimating the climate change by equal weighting remains as big as the observed climate change. We applied the optimal combination approach to reduce uncertainty in projections of changes in the mean seasonal temperature and compared the results with the equal weighting approach. Optimal weighting scheme outperformed the equal weighting scheme with lower RMSE and stronger correlation with the observed changes. In all three problems, we considered outputs from the historical and the hindcast runs from the Coupled Model Inter-comparison Project Phase-5 (CMIP5). Results are analyzed separately to understand the efficiency of the approaches under both CMIP5 experiments.

© Copyright 2016 Rajarshi Das Bhowmik

All Rights Reserved

Reducing Downscaling & Model Uncertainties in CMIP5 Decadal Hindcasts/ Projections

by
Rajarshi Das Bhowmik

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Civil Engineering

Raleigh, North Carolina

2016

APPROVED BY:

_____	_____
Dr. Sankarasubramanian Arumugam (Committee Chair)	Dr. Tushar Sinha
_____	_____
Dr. Gnanamanikam Mahinthakumar	Dr. Kenneth E. Kunkel

DEDICATION

To My Grandparents-

Late Chittaranjan Das Bhowmik

Late Sushil Kumar Dey

Late Bhupeshchandra Das Bhowmik

Late Pritikana Dey

And

Nivedita and Ava Das Bhowmik

BIOGRAPHY

Rajarshi (b. 1987) is the only child of Chinmay (Father) and Rama (Mother). Both his parents worked in government sector and currently they are enjoying life after retirement. Early days of Rajarshi's childhood were spent in a joint-family environment, with crazy siblings and close relatives under the same roof; the life, he still cherishes. Later, his parents moved from their county home to Barasat, a suburb in Kolkata (India). Rajarshi completed his schooling from Barasat Mahatma Gandhi Memorial High School. He finished his Bachelor of Engineering from BESU, Shibpur (currently, IIEST Shibpur) and joined IIT Kanpur for masters study in Civil Engineering. Rajarshi was awarded M.Tech degree in 2012, from IIT Kanpur with Hydraulics and Water Resources specialization. On a sunny day of August 8 of 2012, he moved to the city of oaks, Raleigh, to begin his PhD research.

Rajarshi remained active in student governance till his graduate study. During undergraduate study, he was one of the working committee members of a student organization, called Independent Consolidation(IC). Later at IIT Kanpur, Rajarshi actively participated in student governance and served as the Chairman, maintenance committee and the Interim President of hall of residence IV. Reading remains Rajarshi's favourite past-time since childhood and he recently gained interest in the history of the precolonial America, American civil war and civil liberties movement. Apart from these, he is also a movie-buff and music enthusiast.

Rajarshi is married to Anuradha who is a software professional and an art lover. The couple loves to spend time in cooking and chatting.

ACKNOWLEDGMENTS

'...Round the decay

Of that colossal wreck, boundless and bare

***The lone and level sands stretch far away'*-Ozymandis, P. B. Shelly**

God is gracious and never meager in bestowing me His blessings, kindness and love.

Every human being is full of possibilities and here I just fulfilled one of my journeys. It would not have been easier without my adviser Dr. Sankar Arumugam, his relentless support and guidance. Throughout the course of the study, apart from his minute observations, valuable suggestions and apt advices towards my research, he kept inspiring me to challenge new horizons and seek for the robust solution of hydro-climate problems.

I would like to express my sincere gratitude to my PhD committee members. Dr. Mahinthakumar was always eager to listen to the issues that I faced during research and advised me to pursue the appropriate course. Dr. Sinha helped me significantly to saturate in the realm of hydro-climate research and provided valuable resources whenever I needed them. Dr Kunkel's active involvement in my research and suggestions alleviated my work to a better shape.

I have grateful to the faculty members and the staffs of Department of Civil, Construction and Environmental Engineering, NCSU. I consider it a great pleasure to show my gratitude to the National Science Foundation for supporting my research. I acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and I thank the climate modeling groups (listed in the data descriptions) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for

Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

'Give them

what is desire…

***there is plenty!'*- Give, This Life, Pranabendu Dasgupta**

I was never short of friends and always valued them a lot! I am fortunate to have amazing colleagues and lab mates in my department. I am thankful to each one of them- Seung, Jason, Harminder, Amir, Dominic and Sudarshana. My family away from home includes- Rudradip, Priyam, Debraj, Mangesh, Saswata, Arunava, Michelle, ShuvaDa, SamirDa and a lovely cat-Chappals.

From a great distance of thousands of miles, AritraDa and ArnabDa never failed to keep in touch. Some friendships are made for life-Fulu, Gudul, Tanmay, Kapil and Rishida are here in my life to serve that purpose.

My best regards to my loving brothers and sisters- Bony, Tini, Oni, Rini, Tinni, Dadabhai, Rinki, Debu, Rima, Gablu, Tablu, Om and didi.

I am indebted to all my teachers for helping me become the person I am today.

And, last but not the least, I have no words to thank them-my wife, parents and family members. I love you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

It is now clearly evident that climate and its impact on hydrology is changing around the globe at different temporal and spatial scales (Gleick 1985, McCabe and Wolock 1999, Lettenmaier et al., 1992, Wilby et al., 1999, Wolock and McCabe 1999, Arora and Boer 2001, Sanakarsubramanian and Vogel 2001, Christensen et al., 2004, Clark et. al., 2016, Mondal and Majumder 2016). In an effort to capture this change, General Circulation Models (GCM) are developed that simulate past climate and project future changes (Phillips 1956, Meehl et al., 2007, Flato et al., 2013). GCM simulations for past climate are important to understand the interaction between various components- land surface, atmosphere and ocean-of earth system, impacts of carbon emissions on atmosphere, rise in the sea surface temperature and many other attributes. Further, future projections of hydro-climate are important from the perspective of planning and mitigation. GCMs are revised and updated regularly to reduce uncertainty in projections and to ensure accuracy. In the recent fifth assessment, Coupled Model Inter-comparison Project Phase 5 (CMIP5), GCMs' spatial resolution has been considerably improved (Taylor et al., 2012). In CMIP5, GCM runs that are forced with the initial ocean state, known as decadal Hindcast, are also archived, to predict climate over near-term (10-30 years).

Societal development, particularly industrialization and increased concentration of gases that absorb outgoing long-wave radiation, consequently impact both the near-term and the long-term climate and land-surface process. The planet has experienced significant changes in climatology of mean monthly hydro-climatic attributes and extremes. GCMs provide essential input variables, precipitation and temperature, to estimate runoff that are critical for watershed planning and management. For instance, effects of climate change on land-surface fluxes can also be assessed by land-surface models where GCM outputs serve as climate forcing. However, climate variables resulting from GCMs encompass various source of uncertainties and also exhibits spatial and/or temporal biases. Climate forcings, if applied for hydrologic simulations without proper post-processing, could lead to significant source of errors in land surface attributes (Sinha et al., 2014, Mazrooei et al., 2015, Seo et al., 2016). In this study, we focus my interest on two post-processing processes of climate forcing, statistical downscaling and multi-model combination that reduce uncertainty in GCM projections in climate application studies

In Chapter 2, we developed a multivariate downscaling technique that better estimates the changing cross-correlation structure across multiple variables for downscaling climate variables from coarser to finer resolution. The multivariate scheme is based on Asynchronous Canonical Correlation Analysis (ACCA). Results from the multivariate downscaling are compared with the results for univariate downscaling. In Chapter 3, we analytically investigated the limitations of univariate downscaling that leads to bias in estimating the change in cross-correlation. A framework based on Fisher z-transformation analysis is proposed to compare downscaled and observed cross-correlations. The framework could be

applied on any downscaling methods, thus being dependent on GCM and observations. We applied the framework to analyze the ability of GCMs in estimating cross-correlation across the contiguous US based on hindcast and historical runs. Lastly, in Chapter 4, we studied the spatial bias existing between the GCM projected change and observed change in temperature. Given high spatial bias exists across all models, we investigated two methods- equal and optimal weighting-for reducing model uncertainty. Hence, we proposed an optimally weighted model combination framework that can reduce uncertainty and bias in change in temperature. Along with proposing new and improved approach for post-processing of climate change projections, we also compared the performances of Hindcast and historical runs in estimating the observed changes in cross-correlation.

Finally in Chapter 5, we presents the salient findings and opportunities for further enhancing the limitations of this dissertation research.

Chapter 2. Multivariate Downscaling Approach for Preserving Cross-Correlations across Climate Variables

2.1 Introduction

Global Climate Models (GCMs) have shown potential in simulating/projecting climate variables over historical and future time periods (Hawkins and Sutton, 2008, Taylor et al., 2011). These projections are widely used to quantify future changes and associated uncertainties in hydrological responses induced by climate changes (Lettenmaier et al., 1999, Hay et al., 2002, Hanson and Dettinger, 2005, Mejia et al., 2012, Singh et al., 2014). However, the CMIP5 GCM grid resolutions are too coarse for application in hydrologic models, thereby requiring downscaling. Downscaling is basically defined as the process of interpolating coarser scale variables to a finer spatial scale (Karl et. al., 1990, Wigley et. al., 1990). Dynamic Downscaling and Statistical Downscaling are the two commonly employed types of downscaling with the former using regional climate models to downscale GCM projections (Giorgi et. al., 2001, Hay and Clark 2003, Leung et al., 2003,) the latter using statistical models to obtain projections at finer-scale (Wood et al., 2004, Wilby et al., 2004, Gangopadhyay et al., 2005, Fowler et al., 2007, Maurer and Hidalgo, 2008). The challenges in dynamic downscaling primarily arises from the difficulty in setting up the model for the spatial domain of interest, calibrating to the observed variables and due to the need of intense computational

time (Wood et al., 2004, Maurer and Hidalgo, 2008). Given the ease of model fitting and reduced computational cost, statistical downscaling has gained wide popularity over dynamic downscaling (Wilby et al., 2004, Gangopadhyay et al., 2005).

The primary assumption behind statistical downscaling is that the statistical relationship estimated between large-scale GCM projections and the local-scale data does not change with respect to time. Different types of Statistical Downscaling techniques have been employed ranging from simpler regression methods (Huth, 1999), semi-parametric methods (Stoner et al., 2012, Sankarasubramanian and Lall, 2003), non-parametric techniques (Gangopadhyay et al., 2005) to complex weather generators (Wilks and Wilby, 1999). Wood et al., (2004) provided the concept of bias correction and statistical downscaling (BCSD) based on quantile mapping. Recently, Bureau of Reclamation (2013) have developed monthly precipitation and surface temperature at 1/8° over the continental US for CMIP3 and CMIP5 projections using the BCSD concept. In this two-step process, bias correction of GCM projections is first performed based on univariate quantile mapping on the coarse GCM scale and then spatial interpolation on the bias-corrected variable is performed to obtain finer resolution (1/8°) precipitation and temperature. In essence, BCSD provides bias-corrected climate projections by relating the non-exceedance probabilities of the model data with the non-exceedance probabilities of the observed data. Asynchronous piecewise regression models, basically quantile mapping or regular regression on sorted (i.e., ascending/descending) data, also have shown their capability to downscale climate data at the local scale (Stoner et al., 2012). He et al., (2012) further extended the idea of quantile mapping for bivariate asynchronous measurements by proposing a notion of bivariate ranks and positions. However, extending the

approach beyond bivariate techniques seems to be challenging due to the prescribed ranking methods. BCSD methods have also been modified to estimate joint densities of multiple variables for preserving cross-correlation (Zhang and Georgakakos, 2011, Mehrotra and Sharma, 2016). However, both of the above approaches have not been evaluated for different seasons and over large regions. Another recently developed multivariate downscaling approach, Multivariate Adapted Constructed Analogs (MACA) produces spatially disaggregated daily time series of multiple predictor variables (Abatzoglou and Brown, 2012). MACA corrects bias in the GCM simulations, adjusts epoch for no analog situations under future climate scenarios and identifies the patterns between the GCM and the observed fields using the Constructed Analogs algorithm (Hidalgo et al., 2008). MACA was also compared with the BCSD downscaled products and had exhibited improvements in wildfire applications (Abatzoglou and Brown, 2012).

Most of the downscaling techniques are calibrated or develop downscaled products using historical simulations of 50 years. Climate projections from the Coupled Model Inter-comparison Project Phase-5 (CMIP5) substantially differ from its earlier version (CMIP3) with increased spatial resolution and also encompasses experiments that focus on decadal variability with hindcasts and future projections obtained by initializing GCMs with observed ocean states (Taylor et al., 2005). There is a growing scientific consensus that at decadal time scales (10-30 years) the choice of the scenario for greenhouse-gas emissions contributes little to the uncertainties in climate scenarios generated by different GCMs (Hawkin and Sutton, 2008, Meehl et al., 2009; Taylor et al., 2011). Studies have shown that initializing the GCMs with observed SSTs provide better decadal predictions of surface temperature, SST and other ocean

circulation features (Keenlyside et al., 2008; Smith et al., 2010). Recently, Goddard et al., (2013) compared the ability of both schemes, initialized and uninitialized, in developing decadal hindcasts and they found that initialized decadal hindcasts provide better skill in predicting observed temperature in comparison to the skill of uninitialized hindcasts in predicting observed temperature. CMIP5 hindcast runs were basically initialized with the observed ocean state at a different times with the archives having ten 10-year hindcasts (1961-1970, 1966-1975,…, 2006-2010) initialized every five years from 1960 and two 30-year hindcasts initialized in 1960 (1961-1990) and 1980 (1981-2010). However, spatial resolutions of these CMIP5 hindcasts are still coarser (around $1.4^{o}\times1.4^{o}$) than needed for many climate-application studies, thereby requiring downscaled products. To our knowledge, downscaled products for CMIP5 hindcasts are not available as of now. Given the importance of decadal hindcasts and projections for planning activity, this study also considers multivariate downscaling of hindcasts and compares the performance of downscaled products from both historical runs and hindcasts from CMIP5.

The objective of the current research is to propose and evaluate statistical downscaling techniques that preserve cross-correlation structure across variables. The proposed multivariate downscaling technique is developed by modifying the traditional Canonical Correlation Analysis (CCA) method (Hotelling 1936) for asynchronous observations. We evaluate the proposed methodology by downscaling monthly precipitation and temperature from CMIP5 over the continental US and the performance of the proposed technique is compared with the existing univariate BCSD downscaling products. The proposed downscaling technique is also applied to both hindcasts and historical simulations from CMIP5 to understand how the

technique preserves the observed cross-correlation between precipitation and temperature over the continental United States. We compare three downscaled products (BOR's BCSD, ACCA and MACA) in estimating the joint likelihood of precipitation and temperature.

The manuscript is organized as follows: Section 2.2 provides the details on the observed climate data and CMIP5 simulations and hindcasts over the continental US. Following that, we propose the asynchronous CCA methodology for preserving multivariate correlation structure in climate change projections. Then, we evaluate the performance of the proposed technique and compare with univariate BCSD products available from the Bureau of Reclamation (BOR). Finally, we summarize the findings of the study along with the discussion.

2.2 Dataset

Monthly precipitation (*pr*) and average surface (2 m) temperature (*tas*) data from the CNRM-CM5 GCM was used for downscaling. CMIP5 includes 20th century historical runs and decadal hindcast runs (Taylor et al., 2005). Experiments forced with representative concentration pathway (RCP) scenarios (Van Vuuren et al, 2011) provide future projections. Decadal runs with GCMs initialized in 2005 provide the future projections for the period, 2006-2034, under different RCPs. We chose CNRM due to its reasonable ensemble size (5 members) under historical runs and under hindcasts (9 members). CNRM-CM5 was developed jointly by CNRM-GAME (Centre National de Recherches Météorologiques—Groupe d'études de l'Atmosphère Météorologique) and CERFACS to contribute towards CMIP5 multimodel ensemble (Table 2.1). The model has improved in terms of spatial resolution, dynamic core of atmospheric resolution and radiation scheme from its earlier version (Voldoire et al., 2005).

Spatial resolution of the model is ~1.4° which is re-gridded to ~1.0° using bilinear interpolation for the purpose of further downscaling.

For the purpose of evaluating the proposed multivariate downscaling approach, we consider the period 1950-1999 from CNRM-CM5 historical run, having 5 ensemble members. The historical runs are primarily used to compare with the existing univariate approach since BOR downscaled products are primarily available for historical runs.

We have used decadal-1960 run (1961-1990) and decadal-1980 (1981-2000) run for our analysis. Decadal-1960 (Decadal-1980) was initialized in 1960 (1980) with observed SST conditions in 1960 (1980). Both experiments span for 30 years and each have 9 ensemble members. Hindcast runs are also regridded to have the same spatial resolution as historical run. We consider hindcasts for evaluating multivariate downscaling products from hindcast and historical runs.

To compare our multivariate downscaling approach, we obtain the BOR downscaled CMIP5 products. BOR has developed bias corrected 20th century historical runs ("Downscaled CMIP3 and CMIP5 Climate and CMIP5 Climate and Hydrology Projections") archived at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/. Historical runs were bias corrected at coarse spatial resolution using quantile mapping and then spatially disaggregated to finer grids (Maurer et al., 2007). Since their bias correction approach is quantile mapping, it provides us the baseline for comparing the proposed multivariate downscaling approach. The BOR database also provides observed *pr* and *tas* over the coterminous US (CONUS) for the period January 1950 to December 2010 at a spatial resolution of ~1°. The performance of our multivariate downscaling approach and the univariate approach is compared with the observed

gridded variables of *pr* and *tas*. Downscaled products of CMIP5 using MACA (Version 2) are achieved at http://maca.northwestknowledge.net/. Spatial resolution of the dataset is $1/24^0$. We upscale MACA products (monthly precipitation, monthly maximum and minimum temperature) to 1^0 using the Climate Data Operator (CDO) software. Monthly maximum and minimum temperature are averaged to obtain the monthly average temperature. Only downscaled historical runs are available in the MACA achieve.

Table 2.1: Description of GCM, CNRM-CM5, historical and hindcast runs considered for the study.

Institute	Experiment	Time Period	Emission Scenario	No. of Ensemble Members
Centre National de Recherches Me´te´orologiques and Centre Europe´en de Recherche et de Formation Avance´e (CNRM-CERFACS)	**Historical**	1950-1999	RCP 8.5	5
	Hindcast	Decadal-80 (1981-2010)	Pre-Industrial Control Run	9
		Decadal-60 (1961-1980)	Pre-Industrial Control Run	9

2.3 Motivation & Methodology

2.3.1 Motivation

To understand the need to preserve the cross-correlation structure in downscaling, we first present how the cross-correlation between observed precipitation and observed temperature has changed over the continental US between two periods 1950-1974 and 1975-1999 (Fig 2.1). Grid points exhibiting significant cross-correlation for the 1950-1974 (1975-

1999) period are shown with a '×' ('•') sign, while grids exhibiting significant cross-correlation in both periods are shown with a '▲' sign. Significance in cross-correlation are computed at 95% confidence level based on [$\pm 1.96/\sqrt{(n-3)}$]. For 25 years of data, grid points having cross-correlation greater or lesser than ± 0.42 are considered as significant.

From Figure 2.1, the Northeast region, Ohio valley and FL exhibit significant correlation in January only for the earlier period and the cross-correlation strength decreases being not significant over the later period. The change in cross-correlation for later period over the Northeast and the Ohio valley is primarily due to the increase in mean monthly temperature and decrease in mean monthly precipitation during the later period (Supplemental Information (SI) Figure – 2.7). Significant cross-correlation exists in January for both periods over Montana. Significant cross-correlations exist for both periods over CA and over Rockies during April, whereas significant cross-correlation prevails during the later (earlier) period over NC and VA (coastal Northwest) in April. Insignificant cross-correlation in the later period over coastal Northwest also appears to be due to the increase in mean monthly temperature over the later period (SI-Fig. 2.7).

In July, cross-correlation is significant during both periods over the eastern and central Sunbelt regions, coastal Northwest and northern Great Plains. Cross-correlation in July also appears to be significant over the latter (earlier) period over the upper Midwest (AZ and Northeast). From SI-Fig. 2.7, there is a marked increase in temperature for July in the later period over the eastern US. In October, significant correlation exits in both periods over Colorado. On the other hand, significant cross-correlation existing in the earlier period over the West and FL disappears in the later period in October. From SI-Fig. 2.7, it is clear that the

mean monthly temperature has decreased in the later period over the entire US except the west coast. Further, mean monthly precipitation has also increased over the Southeast US, TN valley and over upper Mid-west (SI-Fig. 2.7). Some of the coastal regions (Northwest, Northern CA and LA) also exhibit significant correlation during the later period. To summarize, number of grid points exhibiting significant cross-correlation has changed from 208,145, 323, 143 to 114, 170, 357 and 117 respectively for January, April, July and October over the later period. To summarize, the changing patterns in monthly precipitation and temperature, change the cross-correlation between precipitation and temperature over the two time periods that we considered. Thus, any downscaling method that we employ needs to accommodate the changing cross-correlation structure between the variables. Hence, we propose and evaluate multivariate downscaling methodology that preserves the correlation structure across multiple predictands and predictors.

2.3.2 Multivariate Approach

Canonical Correlation Analysis (CCA) is a regression based method. In climate studies, CCA has been applied to identify significant drivers associated with forecast skills (Barnett and Preisendorfer, 1987) and to correct systematic errors in GCM simulations (Tippett et al., 2003). In this study, we modified CCA for multivariate downscaling.

Given $\boldsymbol{X}$ being the multiple variables from the GCM runs and $\boldsymbol{Y}$ being the corresponding observed variables over $\boldsymbol{n}$ years at a given grid point, our interest is to develop

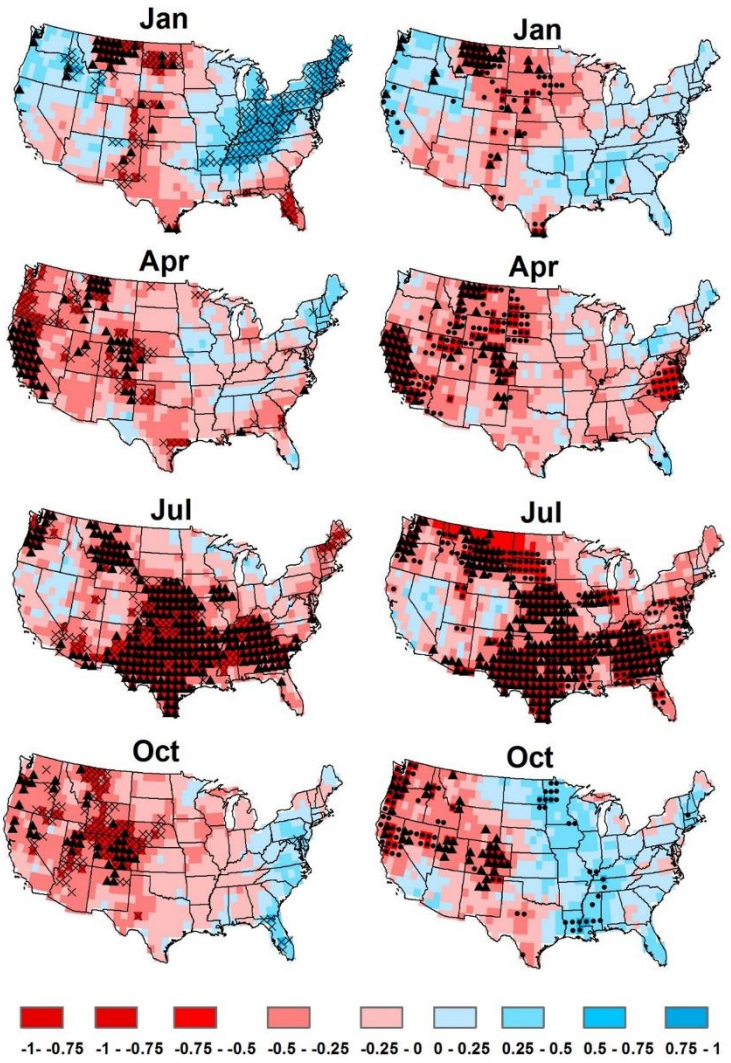

Figure 2.1: Cross-correlation between observed precipitation (Pr) and average temperature (tas) for the period 1950-1974 (left) and 1975-1999(right). Grids with significant correlation during earlier period and later period are shown with symbols [×] and [•] respectively, while grids having significant cross correlation grids over both periods are marked with [▲] symbol.

an empirical relationship between $\boldsymbol{X}$ and $\boldsymbol{Y}$, so that it could be applied for future projections (Stoner et al., 2011). For a given month, dimensions of $\boldsymbol{X}$ and $\boldsymbol{Y}$ will be $(\boldsymbol{n} \times \boldsymbol{p})$,where $\boldsymbol{p}$ is the number of variables to be downscaled. $\boldsymbol{X}$ and $\boldsymbol{Y}$ matrices are constructed for calibrating the proposed downscaling scheme. Let $\boldsymbol{Z}$ denotes the validation set of GCM projections with a

dimension of $\boldsymbol{Z}$ being ($\boldsymbol{\acute{n} \times p}$) where $\boldsymbol{\acute{n}}$ is the number of years considered for downscaling. Given the data $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ as asynchronous data with observations and GCM responses having no time-correspondence, we modify the traditional Canonical Correlation Analysis (CCA) for downscaling GCM projections.

Asynchronous Canonical Correlation Analysis (ACCA)

The steps in performing ACCA downscaling is summarized in Figure 2.2:

1. Models are developed for each grid individually. As mentioned before, dimensions of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are ($\boldsymbol{n \times p}$) and dimension of $\boldsymbol{Z}$ becomes **($\boldsymbol{\acute{n} \times p}$).**

2. Bivariate sorting of X and Y based on bivariate probability.

Given that climate change projections do not possess any time-correspondence with the observed variables, it is common to employ asynchronous regressions for downscaling climate change projections (Stoner et al., 2012). Asynchronous regression is typically developed by sorted $\boldsymbol{X}$ and sorted $\boldsymbol{Y}$ in ascending or descending order. Univariate sorting is based on the quantiles of the fitted distribution (Koenkar and Basset, 1978) or on quantile mapping (Wood et al., 2004). Multivariate sorting can be done by complex statistical methods such as bivariate ranks and positions (He et al., 2012). We propose a simpler, but yet effective, way by extending the concept of univariate sorting for multivariate sorting in which the calibration dataset $\boldsymbol{X}$ and $\boldsymbol{Y}$ were sorted based on their joint probability of occurrence. Given $\mu_{\boldsymbol{X}}$ and $\mu_{\boldsymbol{Y}}$ being the respective means of $\boldsymbol{X}$ and $\boldsymbol{Y}$ and $\boldsymbol{\Sigma_{XX}}$and $\boldsymbol{\Sigma_{YY}}$ being the co-variance matrices of $\boldsymbol{X}$ and $\boldsymbol{Y}$, we assume that $\boldsymbol{X}$ and $\boldsymbol{Y}$ follow bivariate normal distribution.

$$\boldsymbol{X} \sim \text{Bivariate Normal } (\mu_{\boldsymbol{X}}, \boldsymbol{\Sigma_{XX}}) \quad \text{Eq. 1}$$

$$\boldsymbol{Y} \sim \text{Bivariate Normal } (\mu_{\boldsymbol{Y}}\ \boldsymbol{\Sigma_{YY}}) \quad \text{Eq. 2}$$

3. The higher the joint probability of $f_X(\boldsymbol{X}_t)$ and $f_Y(\boldsymbol{Y}_t)$, the higher the ranks are that should be assigned to the multivariate vector at each time step.

Joint probabilities are transformed to log-space before comparison. Thus, both $\boldsymbol{X}$ and $\boldsymbol{Y}$ are sorted separately based on their joint probability of occurrence. These sorted $\boldsymbol{X}$ and $\boldsymbol{Y}$ correspond to each other asynchronously, hence we call the procedure as asynchronous CCA. $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ were fitted in the log space to avoid estimating negative values in the downscaled variables. Hence, we considered the temperature in Kelvin scale for developing the ACCA.

4. A canonical correlation model is then developed using **X** and **Y**.

Canonical correlation, a parametric approach, is considered as the apex of regression based modelling (Barnett and Preisendorfer, 1987) by regressing multiple predictors and predictand in their reduced dimension. CCA defines an optimum linear combination of the prediction dataset that has the potential to explain the total variance in the predictands. CCA linearly rotates predictors $\boldsymbol{X}$ and predictands $\boldsymbol{Y}$ in such a way that their rotated components are orthogonal to each other maximizing the explained variance on rotated *Y* using rotated *X*. $\boldsymbol{U}$ is the rotated component of sorted $\boldsymbol{X}$ and $\boldsymbol{V}$ is the rotated component of sorted $\boldsymbol{Y}$. $\boldsymbol{U}$ and $\boldsymbol{V}$ can be calculated from Eq. (3) and Eq. (4) respectively.

$$\boldsymbol{U} = (\boldsymbol{X} - \boldsymbol{\mu_X})\boldsymbol{A} \qquad \text{Eq. 3}$$

$$\boldsymbol{V} = (\boldsymbol{Y} - \boldsymbol{\mu_Y})\boldsymbol{B} \qquad \text{Eq. 4}$$

Where $\boldsymbol{A}$ and $\boldsymbol{B}$ are the eigenvectors of $\boldsymbol{X}$ and $\boldsymbol{Y}$ with dimensions $\boldsymbol{p}\times\boldsymbol{p}$. Canonical correlations, $\boldsymbol{R}$, specify the correlations between $\boldsymbol{U}$ and $\boldsymbol{V}$**.** If total number of variables is *p***,** then for our case, the size of $\boldsymbol{R}$ is $\boldsymbol{p}\times\mathbf{1}$ having $\boldsymbol{p}$ canonical correlations.

5. Both eigenvectors and canonical correlations from calibration are used in validation to rotate future projections $\boldsymbol{Z}$ to obtain rotated GCM projections for developing downscaled product. Future GCM projection $\boldsymbol{Z}$ was first rotated to $\boldsymbol{U'}$ using eigenvector $\boldsymbol{A}$.

$$\boldsymbol{U'=Z\,A} \qquad \text{Eq. 5}$$

To get the rotated component of downscaled projections from the $\boldsymbol{U'}$**,** canonical correlations from calibration were used.

$$\boldsymbol{V'=U'\,R} \qquad \text{Eq. 6}$$

Finally rotated downscaled $\boldsymbol{V'}$components, are brought back to original space to get the downscaled variable using eigenvector **B** from calibration.

$$\boldsymbol{Y'=V'\,B} \qquad \text{Eq. 7}$$

6. Finally, the downscaled variables $\boldsymbol{Y'}$ are back transformed from the log space by taking the exponential. We built our model by using "cancorr" function in MATLAB R2012a.

2.3.3 Model Development & Validation (Historical and Hindcast)

Asynchronous Canonical Correlation Analysis (Asynchronous CCA) was applied on both historical runs and hindcast experiments of CNRM-CM5. The proposed framework considered the entire period January 1950 to December 1999 from raw historical GCM runs for model development. Downscaled bivariate historical GCM projections were compared with BOR's univariate downscaled products of precipitation and temperature. BOR developed the bias-corrected downscaled products based on quantile mapping by considering the entire dataset over the period 1950-1999. Hence, we compared the performance of ACCA in a similar way by developing the model using the entire dataset 1950-1999. This provides the baseline-

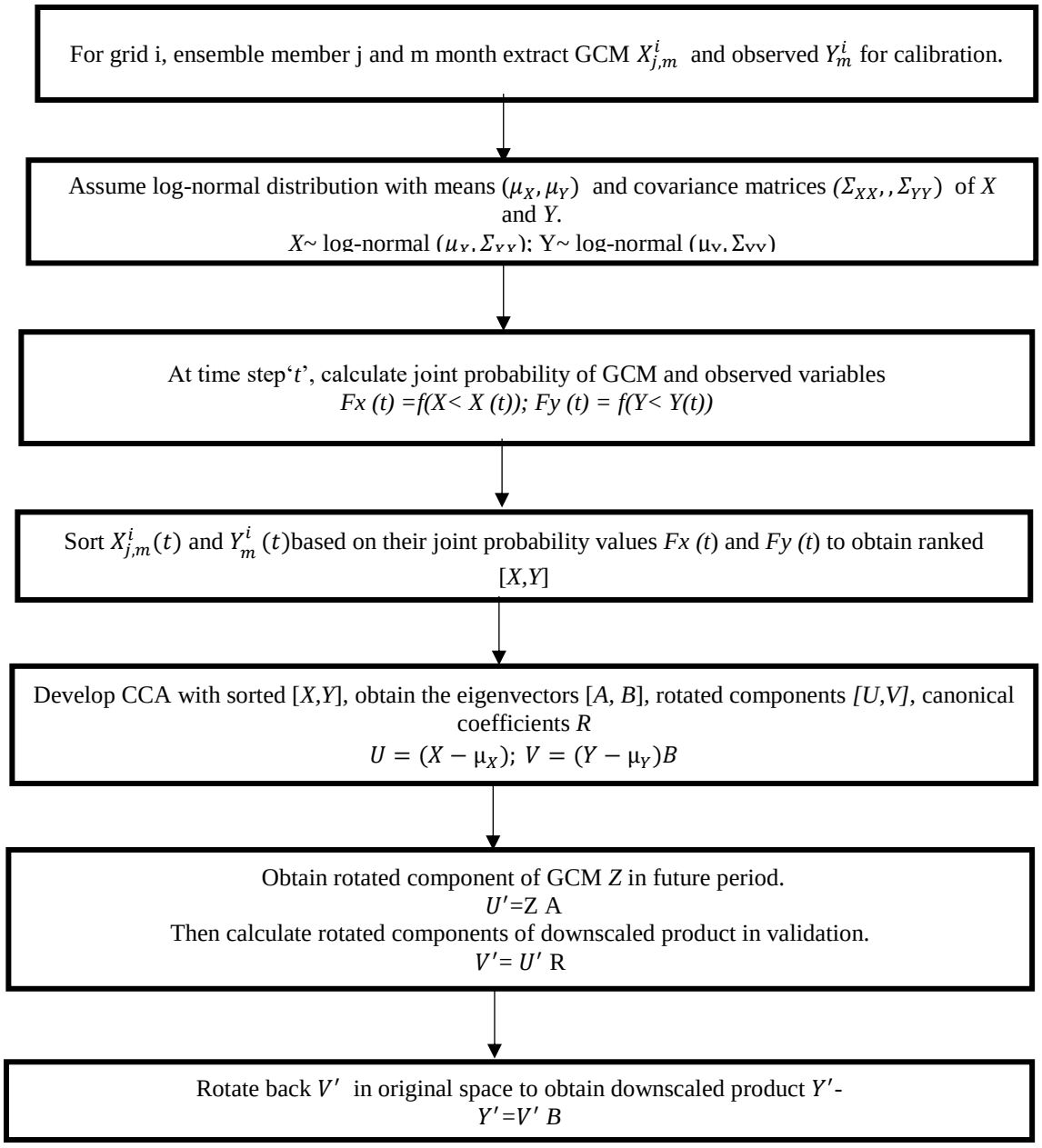

Figure 2.2: Schematic diagram illustrating the ACCA algorithm for multivariate downscaling

-comparison of the proposed technique with the existing downscaled products.

Hindcast Validation: The bivariate downscaling technique was validated for hindcast experiments by leave-five-percent-out cross-validation and by split-sample validation. Under leave-five-percent-out cross-validation, decadal-1980 experimental runs results were used for model development by leaving out the year of prediction as well as additional 5% of the remaining data (i.e., both *X* and *Y*). Then, using the model developed under leave-five-percent-out cross-validation, we obtained the downscaled *pr* and *tas* for the left out years. To compare results obtained from bivariate downscaling of hindcast runs, we developed separate univariate asynchronous regressions. Univariate asynchronous regression (ASR), suggested by O'Brien et al. (2001) and modified by Wood et al. (2004) as the quantile mapping, first sorts both the GCM variable and the observed data based on their probability of exceedance of the univariate variable and then performs the linear regression between the GCM variable and the observed data to get the regression model parameters. Model parameters thus obtained from ACCA were applied on the future GCM projections to get the downscaled projections of GCM data. The performance of the proposed regression technique, ACCA, was compared with the univariate asynchronous regression techniques (ASR) for developing downscaled variables.

2.3.4 Performance Evaluation Metric: Median Fractional Bias

It is important to note that the downscaled precipitation and temperature are developed for each ensemble member from CNRM-CM5. Downscaled ensemble members were used to calculate the fractional bias on the statistic of interest, $\boldsymbol{\lambda}$, from GCM projections. Fractional bias for statistic $\boldsymbol{\lambda}$ is defined as follows:

$$Fractional\ Bias\ for\ \lambda|^{l} = \frac{\theta'^{l} - \theta^{obs}}{\theta^{obs}} \quad \text{Eq. 8}$$

where θ (θ') represents the observed (downscaled) statistic for a given ensemble member l. We considered the bias in estimating the two observed statistics – standard deviation of *pr* and cross-correlation between *pr* and tas – for evaluating the performance of the proposed multivariate technique. We calculated the median of fractional bias using all the ensembles available under historical/hindcast run at a given grid m. We prefer the median to be as close to zero as possible to indicate the statistics of the downscaled variables preserve the observed statistic θ. We have divided medians of fractional biases into 10 categories at an interval of 0.25 for quantifying the spatial variability across the continental United States.

2.4 Results

2.4.1 Comparing the performance of BOR products with ACCA - Historical Runs

Historical runs from CNRM-CM5 consist of 5 ensemble members which result in 5 values of fractional biases for each statistic for a given grid point. The Median at each grid point was calculated over these 5 members. Figure 2.3 shows the median fractional biases in estimating the observed cross-correlation between *pr* and *tas* from the historical ensemble runs. Four months from January to October were plotted in rows while estimates from ACCA and quantile mapping estimates from BOR were plotted in two columns. Given a data length of 50 years over 1950-1999 for historical runs, cross-correlations above (below) 0.29 (-0.29) are significant at 95% confidence interval. Only those grid points for which observed cross-correlation between *pr* and *tas* are significant are marked with a black dot.

In January, ACCA maintains a low median fractional bias within (+/-) 0.25 whereas the BOR approach have median biases between -0.25 to -0.49 in the northern region. Over other regions, ACCA underestimates cross-correlation with medians within 0 to -0.24. In April, greater number of grid points over the central US exhibit significant cross-correlation compared to January. Under those regions, ACCA have median fractional bias between 0 to -0.24. Similarly, the fractional biases over western grid points are also low and negative under those techniques. BOR's univariate quantile mapping approach exhibits median fractional bias more than -1, indicating the inability to preserve the observed cross-correlation structure.

As shown in Figure 2.1, many grid points exhibit significant cross-correlation in July compared to the other three months. In July, ACCA performs better in preserving cross correlation with the median fractional biases for cross correlation between -0.24 to 0.24 (Figure 2.3). BOR's quantile mapping performs better in July in comparison to other months with median fractional bias around 0.25 to 0.49, but still worse than ACCA. In October, significant correlation between *pr* and *tas* shifts to northwest and southwest. ACCA overestimates (underestimates) observed cross-correlation between *pr* and *tas* in the Northwest (Southwest) with the median fractional biases between 0 to 0.24 (0 to -0.24).

Details of the performance of two approaches, ACCA and BOR products, in preserving the observed monthly cross-correlation between *pr* and *tas* for historical runs are summarized in Table 2.2 for grid points exhibiting significant cross-correlation. The values in the table provide fraction of total grid points that has median fractional bias within the range between -0.24 to 0.24. The relative performance of one model over other model is specified in the parenthesis. For example, in January, under ACCA, median fractional bias is within 0.24 to -

0.24 over 92% of the significant grids and ACCA better than BOR in 70% of the grids exhibiting significant cross-correlation. Considering all months, ACCA exhibits median fractional bias in cross-correlation between -0.24 to 0.24 in 90% of the grid points exhibiting significant cross-correlation. On the other hand, univariate downscaling from BOR products shows higher bias with only 0-20% of grid points having low fractional bias (in the range -0.24 to 0.24) during winter and spring months, but improves to around 55% of grid points in other months. Overall in terms of preserving cross-correlation, ACCA performs the best (% shown in parenthesis in Table 2.2) over 70% grids than univariate method. ACCA performs better than the quantile mapping products from BOR based on the number of grids (% of grid points shown in parenthesis in Table 2.2) having the lowest fractional bias in cross-correlation over all the months.

Median fractional bias in estimating the observed standard deviation for the two approaches under historical runs was summarized in Table 2.3 over twelve months. Given that the BOR's univariate ASR approach being quantile mapping, it shows negligible bias in standard deviation almost completely from the GCM runs. For this analysis, we consider all the grid points over the continental United States. During January, ACCA exhibits median fractional bias in standard deviation within - 0.24 to 0.24 over 60% of grids. Results for July are very much the same as January for ACCA. In July, ACCA shows low median fractional bias in 56% of total grids but in April and October, but the bias decreases to 39% and 22% respectively. Overall, ACCA approach is overestimating the observed standard deviation for most grid points across the country. BOR's univariate performs well, since they were developed to ensure the observed mean and standard deviation in precipitation and temperature

over all the grid points. We are not reporting the median fractional biases in mean monthly precipitation and temperature since both approaches preserve the observed mean for the historical period as regression approaches are expected to preserve the long-term mean for the calibration data set. However, the biases in estimating the mean and standard deviation are expected to increase under validation. Given that we have compared the performance of ACCA with BOR products under historical runs, we next compare the performance of ACCA with univariate downscaling, ASR, for downscaling and bias-correcting hindcasts under cross-validation and split-sample validation.

Table 2.2: Fraction of total grid points with median fractional bias in cross-correlation from CNRM-CM5 historical runs having values in the range from -0.24 to 0.25. The values in the parenthesis indicate the number of total grid points where the downscaling approaches had the lowest bias in cross-correlation. The last row provides the number of grid points having statistically significant cross-correlation.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
ACCA	0.92 (0.70)	0.90 (0.83)	0.91 (0.77)	0.88 (0.83)	0.95 (0.8)	0.98 (0.84)	0.92 (0.75)	0.91 (0.71)	0.86 (0.6)	0.79 (0.62)	0.92 (0.57)	0.94 (0.78)
BOR	0.40 (0.30)	0.14 (0.17)	0.28 (0.23)	0.12 (0.17)	0.31 (0.2)	0.55 (0.16)	0.53 (0.25)	0.53 (0.29)	0.63 (0.4)	0.53 (0.38)	0.64 (0.43)	0.23 (0.22)
Grids	100	99	54	139	190	351	386	305	192	133	97	216

Table 2.3: Fraction of total grids with median fractional bias in standard deviation of precipitation from CNRM-CM5 historical runs having values in the range from -0.24 to 0.25. The values in the parenthesis indicate the number of total grid points where the model had lowest bias in standard deviation.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
CCA	0.63 (0)	0.64 (0)	0.6 (0)	0.39 (0)	0.53 (0)	0.41 (0)	0.56 (0)	0.58 (0)	0.31 (0)	0.22 (0)	0.54 (0)	0.71 (0)
BOR	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)

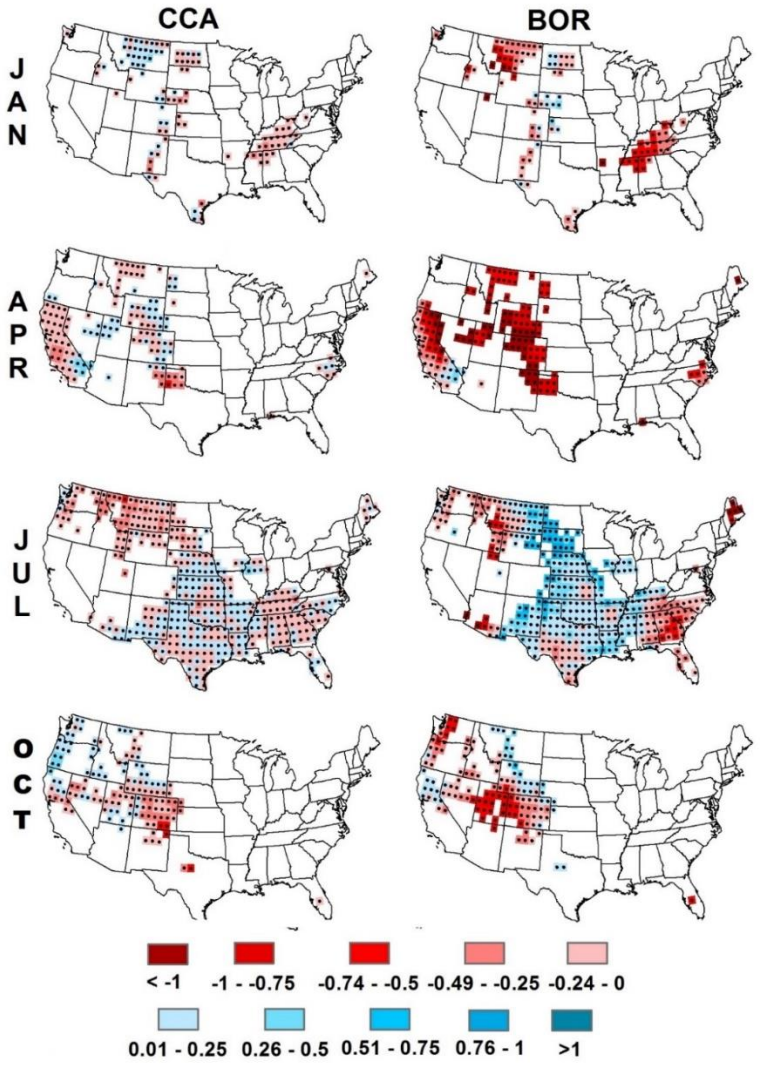

Figure 2.3: Median of fractional bias of cross correlation between monthly precipitation rate (pr) and monthly average temperature (tas) based on historical runs. Fractional biases were calculated only on those grids where observed cross correlation between pr and tas was significant for the period 1950-1999.

2.4.2 Performance comparison of ACCA and ASR for Hindcasts

Since the hindcasts typically consist of 10-year/30-year runs, we compare the performance of ACCA and ASR based on leave-five-percent-out cross-validation. Hindcast experiments from CNRM-CM5 consist of 9 ensemble members so medians were calculated over 9 members. Median fractional bias for cross-correlation between *pr* and *tas* were plotted for three approaches and for four months in Figure 2.4. Similar to historical runs, black dots were pointed at the center of grid points where observed cross-correlations between *pr* and *tas* were significant for the period 1981-2010. Significance for 30 years of monthly data is between (+/-) 0.37. Results for all the months are summarized under Table 2.4.

During January, ACCA keeps the median fractional bias for cross correlation within the low range (between -0.24 to 0.24) at most of the grid points. Univariate ASR approach underestimates cross-correlation at most of the grid points. From Table 2.4, ACCA preserves cross-correlation within the low range at 83% of significant grid points whereas univariate ASR approach performs better in 35% of the grid points. ACCA has the lowest cross-correlation in January in 79% of grid points that exhibit significant cross-correlation. In April, grid points over the western US result with underestimated cross-correlation by ACCA. Univariate ASR underestimates significantly over the Central and Eastern USA. Univariate performance improves considerably during summer with median fractional bias in low range over 76% of grid points in June and 64% of grid points in July. ACCA was able to preserve significant cross-correlations over the central United States with the median fractional bias between -0.24 to 0.24. However, ACCA underestimates the cross-correlation over the western United States. Considering all the months, ACCA performed better than the univariate ASR

over 60-70% grid points. Further, ACCA was able to limit the bias within the desired low range over 80% grid points.

With regard to preserving standard deviation, the univariate approach resulted in very small biases in estimating the standard deviation of precipitation. Results for ACCA and ASR in preserving the observed standard deviation are summarized in Table 2.5. ACCA overestimated the standard deviation in precipitation during winter months. ACCA preserves the standard deviation for at least 40% grid points with the bias being in the desired range. Comparing with ACCA performance with ASR, in general, ASR performs better in terms of capturing the standard deviation in precipitation. Overall ACCA preserves the cross-correlation better, whereas univariate preserves the standard deviation better and both schemes exhibit equal efficiency in estimating the mean. We address this trade-off by calculating joint likelihood of *pr* and *tas* from both schemes in Section 2.5.

Table 2.4: Fraction of total grid points with median fractional bias in cross-correlation from (leave-five-percent-out cross-validation) CNRM-CM5 hindcast runs having values in the range from -0.24 to 0.25. The values in the parenthesis indicate the number of total grid points where the model had lowest bias in cross-correlation. The last row provides the number of grid points having statistically significant cross-correlation.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
ACCA	0.83 (0.79)	0.83 (0.75)	0.82 (0.76)	0.87 (0.75)	0.94 (0.81)	0.96 (0.72)	0.88 (0.7)	0.88 (0.65)	0.85 (0.76)	0.79 (0.72)	0.81 (0.68)	0.87 (0.7)
ASR	0.35 (0.21)	0.63 (0.25)	0.68 (0.24)	0.13 (0.25)	0.52 (0.19)	0.76 (0.28)	0.64 (0.3)	0.74 (0.35)	0.55 (0.24)	0.57 (0.28)	0.71 (0.32)	0.75 (0.3)
Grids	138	60	146	158	222	324	378	242	214	152	150	120

Table 2.5: Fraction of total grids with median fractional bias in standard deviation from (leave-five-percent-out cross-validation) CNRM-CM5 hindcast runs having values in the range from -0.24 to 0.25. The values in the parenthesis indicate the number of total grid points where the model had lowest bias in standard deviation.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
CCA	0.53 (0.08)	0.54 (0.12)	0.49 (0.08)	0.42 (0.06)	0.49 (0.07)	0.35 (0.06)	0.47 (0.08)	0.47 (0.08)	0.45 (0.06)	0.32 (0.05)	0.51 (0.1)	0.64 (0.16)
ASR	1 (0.92)	1 (0.88)	1 (0.92)	1 (0.94)	1 (0.93)	1 (0.94)	1 (0.92)	1 (0.92)	1 (0.94)	1 (0.95)	1 (0.9)	1 (0.84)

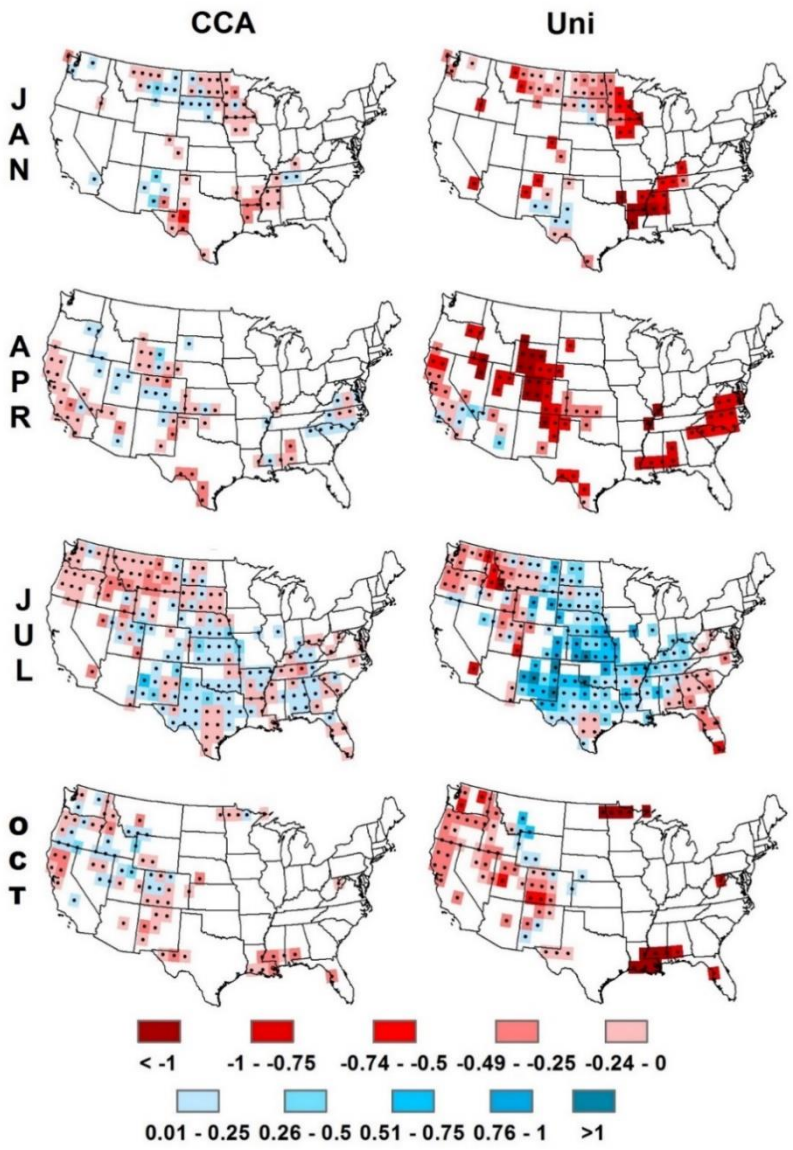

Figure 2.4: Median of fractional bias of cross correlation between monthly precipitation rate (pr) and monthly average temperature (tas) based on decadal hindcasts (leave-five-percent-out cross-validation). Fractional biases were calculated only on those grids where observed cross correlation between pr and tas was significant for the period 1981-2010.

2.4.3 Initialized and Uninitialized CMIP5 run: A Comparison

One objective of this study is to compare the performance of univariate and multivariate downscaling techniques under initialized hindcast runs and uninitialized historical runs in preserving the observed moments. Non-exceedance probabilities (cumulative distribution function, CDF) of cross-correlations from raw GCMs and downscaled GCM runs under hindcasts and historical runs are shown for four months (Figure 2.5). Observed, historical runs and hindcast runs are indicated as red, blue and green lines respectively. CDFs of nine ensemble members of hindcast runs and five ensemble members of historical runs were plotted for grid points exhibiting significant cross-correlation. Raw hindcast runs and raw historical runs underestimate the observed cross-correlation distribution in January and April (first column in Figure 2.5). But, for January and April, univariate bias correction when applied on historical runs did not show any improvements from the raw values (middle column in Figure 2.5). This could also be inferred from the bias-corrected historical estimates from ACCA showing significant improvements as opposed to the improvements resulting from ASR. However, historical estimates from ACCA overestimates the CDF of negative cross correlation compared to the observed cross-correlation CDF. Further, ACCA estimates of cross-correlation for historical runs shows lesser bias compared to the univariate downscaling of historical runs (Figure 2.5ii). However, downscaled and bias-corrected hindcast values from ACCA in January and April shifted the raw values resulting in the downscaled cross-correlation being closer to the observed cumulative distribution (last column in Figure 2.5). One potential reason for improved performance of hindcast could be due to their improved performance in predicting the observed temperature as the models are initialized with observed

SSTs (in this case 1960) prior to the period of developing projections. Studies have shown improved ability of decadal hindcasts in projecting temperature which have been attributed to both the initialized SSTs and to forcing the model with observed greenhouse gases concentration (Keenlyside et al. 2008, Smith et al. 2010, Goddard et al. 2013) The other reason for improved performance of hindcasts under ACCA is due to its ability to develop multiple regression relationships by relating multiple predictands with multiple predictors within the canonical regression framework. Thus, the CDF of hindcasts observed using multivariate downscaling better matches with the observed cross-correlation CDF.

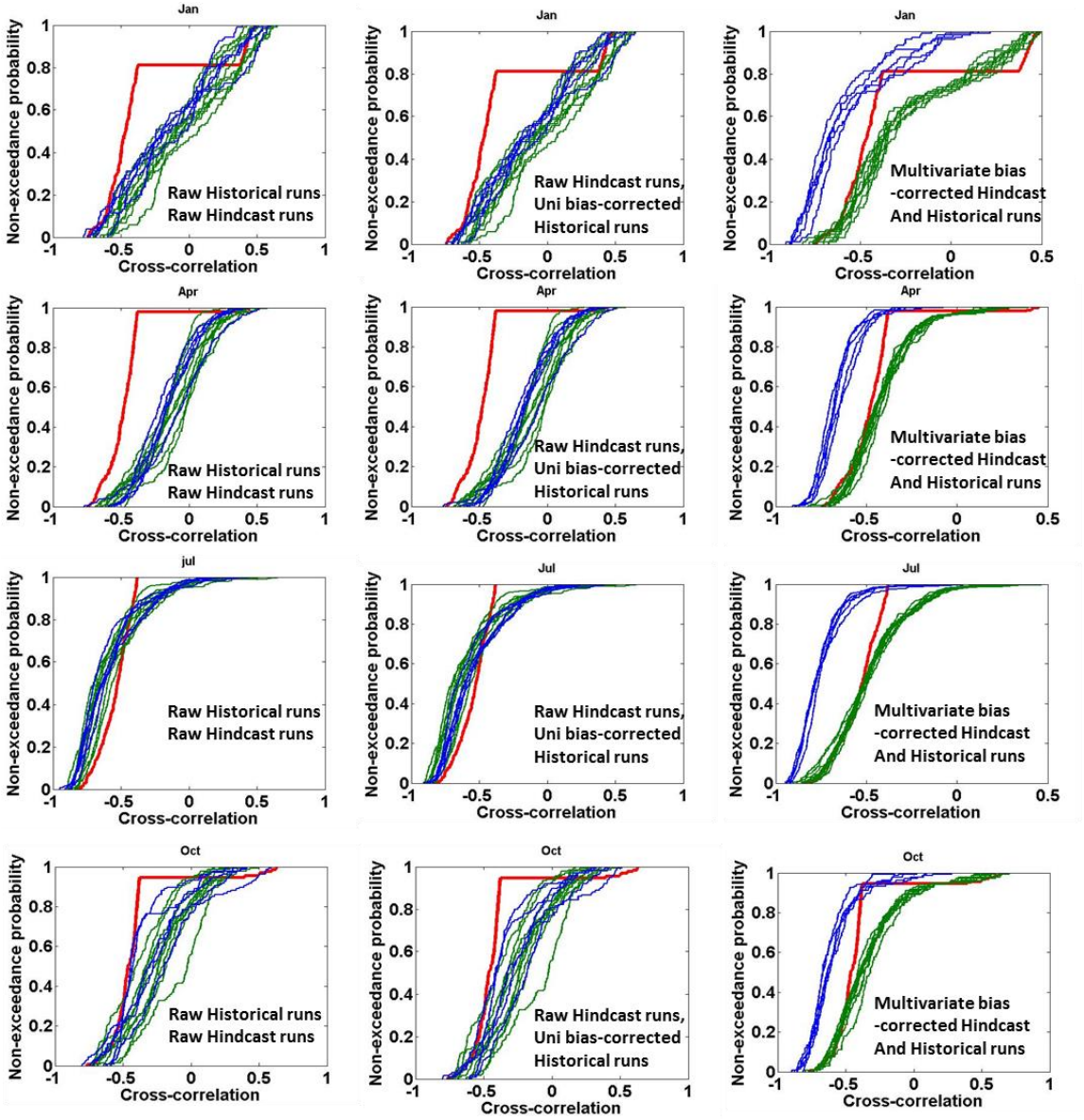

Figure 2.5: Plot of CDFs of cross-correlations for four months- January, April, July and October. Red, green and blue lines are for observed, hindcast runs (9ensemble member) and historical runs (5 ensemble members). Three columns are i) raw hindcast and raw historical runs ii) univariate downscaled historical runs and raw hindcast runs and iii) multivariate downscaled both hindcast runs and historical runs. Only statistically significant cross-correlations (period 1961-1990) were considered.

2.5 Discussion and Concluding Remarks

Currently employed univariate downscaling approaches, such as quantile mapping and ASR, have limited potential in preserving the cross-correlation between precipitation and temperature. Analyses show that cross-correlation between precipitation and temperature is changing over the continental Unites States. Cross-correlation, along with mean and standard deviation, should be considered as an important climatic attribute in spatial downscaling GCM outputs to finer resolutions. Grid points with significant cross-correlation are decreasing in magnitude resulting in insignificant cross-correlation during winter months. On the other hand, particularly summer and fall months, cross-correlation over Mid-west remained significant with limited/no change in magnitude over the Central US. Given such spatio-temporal variability in the dependency between precipitation and temperature, downscaling and bias-correction techniques need to preserve the cross-correlation between precipitation and temperature during downscaling.

Hence, we propose and evaluate a multivariate downscaling technique, Asynchronous Canonical Correlation Analysis (ACCA), for preserving the cross-correlation between precipitation and temperature. Results of multivariate downscaling were compared with univariate downscaling – quantile mapping and ASR. As a baseline, we compared the performance of ACCA with univariate downscaled historical runs from BOR. Multivariate downscaling approach, ACCA, performs better than the univariate downscaling techniques, ASR or quantile mapping, in terms of preserving the cross-correlation. Both methods showed very small bias in estimating the standard deviation of temperature, hence those results are not

presented. However, this improved estimation of cross-correlation comes with a slightly increased bias in estimating the standard deviation of precipitation.

The increased bias in estimating one of the moments could be interpreted as a tradeoff between preserving standard deviation in precipitation and preserving cross correlation between *pr* and *tas* in selecting the right downscaling approach for climate-application studies. To quantify this tradeoff, we evaluate both ACCA and ASR on their ability to estimate the joint probability of observed monthly *pr* and *tas* over the period 1981-2010 using their respective downscaled moments of precipitation and temperature by assuming bivariate lognormal distribution. Similarly, we also estimate the joint probability of estimating the observed monthly *pr* and *tas* considering the observed moments of *pr* and *tas*. For each grid point, the likelihood ratio for ACCA/ASR is calculated by dividing the joint probability of *pr* and *tas* estimated using downscaled moments of ACCA/ASR with the joint probability of *pr* and *tas*. Comparing the two likelihood ratios from ACCA and ASR, the model (i.e., ACCA/ASR) that has likelihood ratio closer to '1' is considered as the best performing approach since the joint probability estimated from that model being closer to the joint probability estimated using the observed moments of the data.

Figure 2.6(a) and Fig 2.6(c), for historical and hindcast runs respectively, provides fraction of total grids where likelihood ratio of ACCA is better than ASR likelihood ratio. The same has been calculated for grid points with statistically significant cross-correlations and results are shown in Fig 2.6(b) and Fig 2.6(d) respectively for historical and hindcast runs. Historical experiment with 5 ensemble members and hindcast experiments with 9 ensemble members are considered.

In three months, January, April and October from historical runs, shown in Fig 2.6(a), ACCA performs better over one fourth of total grid points since the likelihood ratio is close to one particularly for grid points exhibiting significant cross-correlation. During July in which almost half of the total grid points has witnessed statistically significant cross-correlation, likelihood ratio of ACCA is better than the likelihood ratio of ASR being better in more than half of the total grid points. Performance of ACCA is 5 to 10% better when only significant cross-correlation grid points are considered. During January, April and Ocober, ACCA dominates over 30% of significant cross-correlation grid points and during July, ACCA exhibits better joint likelihood ratio compared to ASR likelihood over 60% of grid points. On the other hand, ACCA approach using hindcast runs performs better than ASR over 80% of total grid points, almost for all the months and for all ensemble members, by keeping likelihood of *pr* and *tas* closer to observed likelihood. However, ACCA performance is only slightly enhanced when significant cross-correlation grid points are considered for hindcast runs. This confirms that preserving cross-correlation by ACCA results in better estimation of the joint probability of *pr* and *tas* during those months where the dependency between *pr* and *tas* is statistically significant. ACCA performance is better at estimating the joint likelihood between *pr* and *tas* when initialized runs are considered. Figure 2.6(e) shows the fraction of grid points over the contiguous US where MACA is estimating joint likelihood better than ACCA approach. Out of the four months, MACA performs better during January and October. But performance of MACA products are poorer compared to ACCA over all the months. Additional results on MACA's performance in estimating the cross-correlation is provided

under SI-Figure 2.8. Ability of MACA in estimating the cross-correlation is very much similar to that of univariate approach and ACCA outperforms MACA as well.

We also examined the skills of initialized CMIP5 runs with SSTs compared to uninitialized CMIP5 runs in terms of estimating cross-correlation between precipitation and temperature. For the raw datasets, there was little or no skill for both hindcast runs and historical runs as the CDFs depart significantly from the observed distribution. Multivariate bias correction from ACCA of hindcast runs has shown significant improvement in reducing the bias in estimation of cross-correlation. Multivariate bias correction of historical runs also shows improvements in estimating cross-correlation compared to the estimation from ASR. On the other hand, downscaled cross-correlation from ASR under historical runs did not result in any improvements in reducing the error in estimation of the observed CDF of cross-correlation.

Preserving the cross-correlation between precipitation and temperature is also important for estimating land surface states such as soil moisture and runoff using downscaled GCM projections. Simulation of flow in large basins where more grid points with significant cross-correlations exist could be sensitive particularly in predicting low flows. For instance, if the significant positive correlation between precipitation and temperature during the winter is underestimated, cross-correlations between the projected soil moisture and streamflow will also be underestimated and it is expected to impact the cross-correlations between the snow storage and runoff/snow-melt generation in the spring. Thus, preserving cross-correlations in the downscaled variables is also important for simulating hindcasts of land-surface fluxes under near-term climate change (Seo et al. 2016).

Changing patterns of cross-correlations over the continental US suggests that significant positive (in winter) and negative cross-correlation (in summer and fall) is changing. Hence, we suggest multivariate bias-corrected hindcast and historical runs for precipitation and temperature could be useful in developing better streamflow simulations. Linking this phenomenon with increases in global mean temperature leads to the hypothesis that regions with increased precipitation recycling could also benefit from applying multivariate bias correction. Following this, streamflow simulation with spatially downscaled climate variables without preserving cross-correlation could experience underestimation of positive feedback effects which could in turn lead to bias in estimation of streamflow. Since anthropogenically-forced climate change will lead to warming (Hawkins and Sutton, 2008, Hansen et al., 2010, Goddard et al., 2013), cross correlation between *pr* and *tas* will also likely to change. Thus, any downscaling procedure should preserve the cross correlation structure between large and local scale. From the streamflow perspective, increased temperature leads to increased evaporation resulting in decreased streamflow availability. Thus, in summer months during which local scale precipitation and moisture flow dominates streamflow availability, preserving cross correlation could potentially improve the low flow characteristics in streamflow simulation. Our future effort will focus on the importance of preserving cross-correlation in developing streamflow projections under near-term climate change.

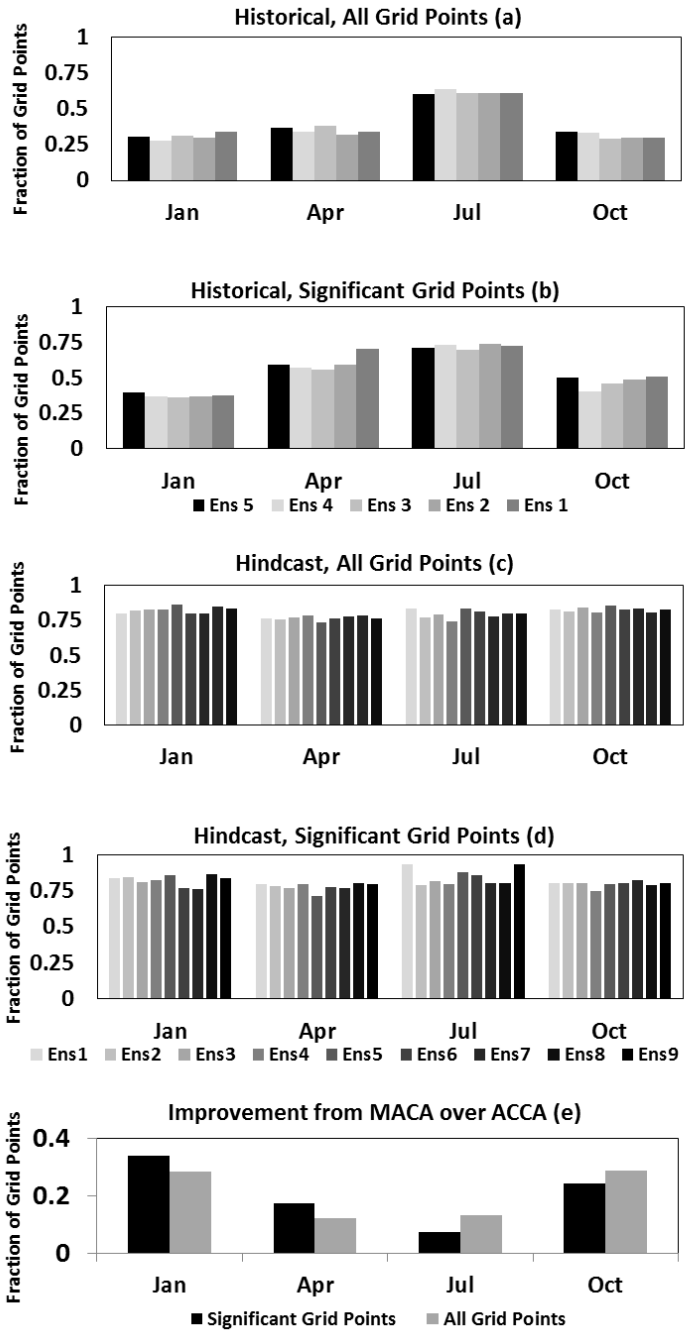

Figure 2.6: Bar chart shows fraction of total grids over which likelihood ratio of ACCA is closer to 1 compared to the likelihood ratio of ASR ((a) to (d)). Ensemble members from historical and hindcast dataset considered for the analysis. 2.6(a) and 2.6(b) are for historical runs with total grid points and significant cross-correlation grid points respectively. Hindcast runs with total grid points and significant cross-correlation grid points are shown in 2.6(c) and

2.6(d) respectively. 2.6(e) shows the fraction of grid points over which likelihood ratio of MACA is closer to 1 compared to the likelihood ratio of ACCA. The first ensemble member from the historical run is used (e).

2.6 Supplementary Information (SI)

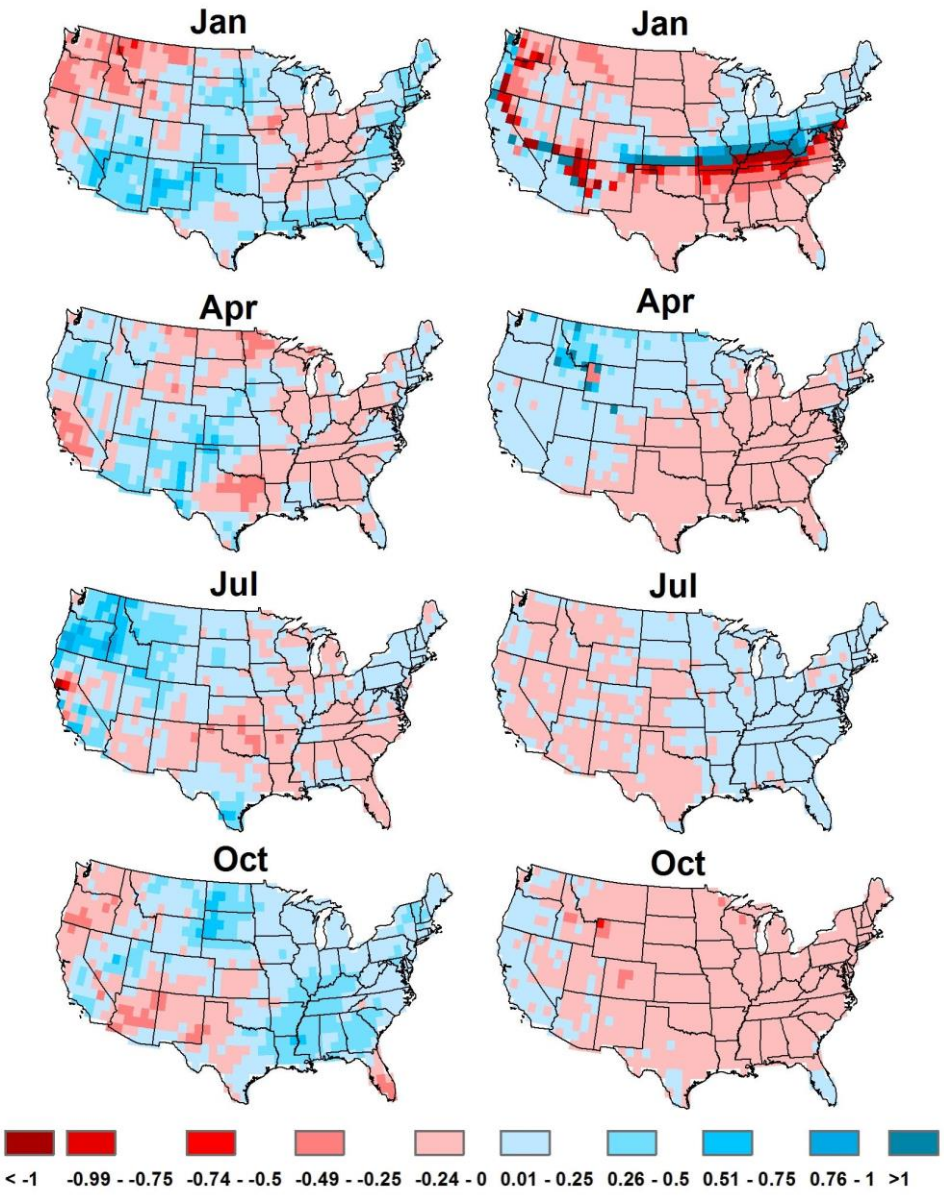

Figure 2.7: Change in the fraction of observed mean monthly precipitation (left) and observed mean monthly temperature (right) between two periods 1950-1974 and 1975-1999. To calculate the change in the fraction plotted in the left (right) column, mean monthly precipitation (temperature) for 1950-1974 is subtracted from the mean monthly precipitation (temperature) for 1975-1999 and the difference is divided by the mean monthly precipitation (temperature) of the entire period 1950-1999.

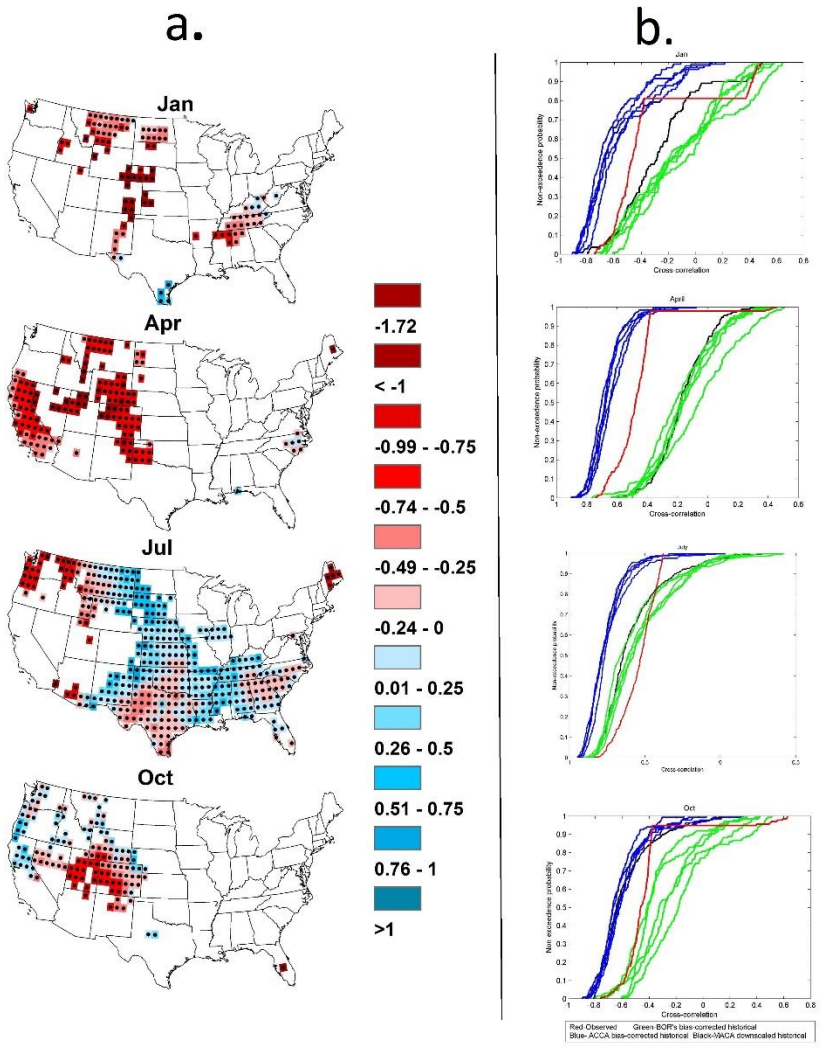

Figure 2.8: Performances of MACA downscaling technique in estimating the cross-correlation. Column a is same as Figure 2.3 but with MACA. Column b is same as Figure 2.5 but plotting MACA (Black), ACCA (Blue), BOR (Green) and observed (Red). MACA downscaled products are available for the historical runs only.

2.7 References (Chapter 1 and 2)

"Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections" archive at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/.

Abatzoglou, J. T., and T. J. Brown, 2012: A comparison of statistical downscaling methods suited for wildfire applications. Int. J. Climatol., 32: 772–780.

Arora, V. K., and G. J. Boer 2001: Effects of simulated climate change on the hydrology of major river basins, J. Geophys. Res., 106(D4), 3335–3348.

Barnett, T. P., R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Monthly Weather Review*, **115**: 1825–1850.

Christensen, N.S. and Co-authors, 2004: The Effects of Climate Change on the Hydrology and Water Resources of the Colorado River Basin. Climatic Change, 62: 337.

Clark, M. P., R. L. Wilby, E. D. Gutmann, J. A. Vano, S. Gangopadhyay, A. W. Wood,H. J. Fowler, C. Prudhomme, J. R. Arnold, L. D. Brekke, 2016: Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Current Climate Change Reports*, **2**, 2, 55.

Fisher, R. A., 1925: Statistical Methods for Research Workers. Edinburgh: Oliver & Boyd. *Text accessible at http://psychclassics.yorku.ca/Fisher/Methods/.*

Flato, G., J. and Co-authors, 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 Cambridge University Press.

Fowler, H. J., S. Blenkinsop, C. Tebaldi, 2007: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int J Climatol*., **27(12)**, 1547–1578.

Gangopadhyay, S., M. Clark, K.Werner, D. Brandon, and B. Rajagopalan, 2004: Effects of spatial and temporal aggregation on the accuracy of statistically downscaled precipitation estimates in the Upper Colorado River Basin. *Journal of Hydrometeorology*, **5(6)**, 1192-1206.

Giorgi, F. B., J. Hewitson, M. Christensen, H. Hulme, P. Von Storch, R. Whetton, L. Jones, Mearns, and C. Fu., 2001: Regional climate information — Evaluation and projections. Climate Change 2001. The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, pp. 583-638.

Gleick, P. H, 1985: Regional Hydrologic Impacts of Global Climate Changes. International Research and Development Conference; Arid Lands: Today and Tomorrow, Tucson, AZ, Office of Arid Land Studies, University of Arizona.

Goddard, L., A. Kumar, A. Solomon, D. Smith, G. Boer, P. Gonzalez, V. Kharin, W. Merryfield, C. Deser and S. J. Mason, et al., 2013: A verification framework for interannnual-to-decadal prediction experiments. *Clim. Dyn*., DOI: 10.1007/s00382-012-1481-2.

Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, **48**, RG4004.

Hanson, R. T., and M. D. Dettinger, 2005: Ground-water/surface-water responses to global climate simulations, Santa Clara-Calleguas Basin, Ventura County, California, 1950-93. *J Amer Wat Res Assoc.*, **43(3)**:517–536.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull Am Meteorol Soc.*, **90**, 1095–1107.

Hay, L. E., R. L. Wilby, and G. H. Leavesley 2000: A comparison of delta change and downscaled GCM scenarios for three mountainous basins in THE UNITED STATES1. *Journal of the American Water Resources Association*, Vol. **36**, No. 2, pp. 387-397.

Hay, L.E. and M.P. Clark, 2003: Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *Journal of Hydrology*, **282**:56-75.

He, X., Y. Yang, and J. Zhang, 2012: Bivariate Downscaling with Asynchronous Measurements. *Journal of Agricultural, Biological, and Environmental Statistics*. **17(3)**, 476–489.

Hidalgo, H., M. D. Dettinger, D. R. Cayan, 2008: Downscaling with constructed analogues – Daily precipitation and temperature fields over the United States: : California Energy Commission Report CEC-500-2007-123, 62 p.

Hotelling, H., 1936: Relations between two sets of variates. *Biometrika*, **28**:312– 377.

Huth, R., 1999: Statistical downscaling in Central Europe: evaluation of methods and potential predictors. *Clim Res.*, **13(2)**, 91–101.

Karl, T.R., W. C. Wang, M. E. Schlesinger, R. E. Knight, and D. Portman, 1990: A method of relating general circulation model simulated climate to the observed local climate. Part I. Seasonal statistics. Journal of Climate 3, 1053-79.

Keenlyside, N. S., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner, 2008: Advancing decadal-scale climate prediction in the North Atlantic sector, *Nature*, **453**, 84-88.

Koenker, R., and G. S. Bassett, 1978: Regression Quantiles. *Econometrica*, **46**, 33–50

Lettenmaier, D. P., A. W. Wood, R. N. Palmer, E. F. Wood, and E. Z. Stakhiv, 1999: Water Resources Implications of Global Warming: A U.S. Regional Perspective, *Clim. Change*, **43**, 537–579.

Lettenmaier, D. P., K. L. Brettman, L. W. Vail, S. B. Yabusaki, and M. J. Scott, 1992: Sensitivity of Pacific Northwest Water Resources to Global Warming. Northwest Environmental Journal 8 (2), 265–283.

Leung, R. L., Y. Qian, X. Bian, and A. Hunt, 2003: Hydroclimate of the Western United States based on observations and regional climate simulation of 1981–2000. Part II: Mesoscale ENSO anomalies. *J Climate*, **16**, 1912–1928.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy 2007: Fine-resolution climate projections enhance regional climate change impact studies. *Eos Trans. AGU*, **88(47***)*, **504**.

Maurer, E.P., and H.G. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences,* **12**, 551-563.

Mazrooei, A., T. Sinha, A. Sankarasubramanian, S. Kumar, and Christa Peters-Lidard 2015: Decomposition of sources of errors in seasonal streamflow forecasting over the U.S. Sunbelt, J. Geophys. Res. Atmos., 120, 11,809–11,825, doi:10.1002/2015JD023687.

McCabe, G. J. and D. M. Wolock, 1999: General-Circulation-Model Simulations of Future Snowpack in the Western United States. Journal of the American Water Resources Association, 35: 1473–1484.

Meehl, G. A., L. Goddard, J. Murphy, R. J. Stouffer, G. Boer, G Danabasoglu, et al. 2009: Decadal prediction: can it be skilful?. *Bulletin of the American Meteorological Society*. **90**, 1467–1485.

Meehl, G.A. and Co-authors, 2007: Global Climate Projections. In: 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 University Press.

Mehrotra,R. and Sharma, A. 2016: A Multivariate Quantile-Matching Bias Correction Approach with Auto-and Cross-Dependence across Multiple Time Scales: Implications for Downscaling, *Journal of Climate*, **29 (10)**,3519-3539.

Mejia, J. F., J. Huntington, B. Hatchett, D. Koracin, and R. G. Niswonger, 2012: Linking Global Climate Models to an Integrated Hydrologic Model: Using an Individual Station Downscaling Approach. Journal of Contemporary *Water Research & Education*, **147**: 17–27.

Mondal, A., P. P. Mujumdar, 2016: Detection of Change in Flood Return Levels under Global Warming. *Journal of Hydrologic Engineering*. **21**, 8.

O'Brien, T. P., D. Sornette, and R. L. McPherro, 2001: Statistical Asynchronous Regression Determining: The Relationship Between Two Quantities That Are Not Measured Simultaneously. *Journal of Geophysical Research*. **106**, 13247–13259.

Oludhe, C.,A. Sankarasubramanian, T. Sinha, N. Devineni and U. Lall, 2013: The Role of Multimodel Climate Forecasts in Improving Water and Energy Management over the Tana River Basin, Kenya. *Journal of Applied Meteorology and Climatology*, **52(11)**, 2460-2475.

Phillips, N. A. 1956: The general circulation of the atmosphere: A numerical experiment. Q.J.R. Meteorol. Soc., 82: 123–164. doi:10.1002/qj.49708235202

Sankarasubramanian, A. and R. M. Vogel, 2001: Climate Elasticity of Streamflow in the United States. Water Resour. Res. 37, 1771–1781.

Sankarasubramanian, A., and U. Lall, 2003: Flood Quantiles and Changing Climate:Seasonal forecasts and causal relations. *Water Resources Research*. **39(5),** art.no.1134.

Seo, S. B., T. Sinha, G. Mahinthakumar, A. Sankarasubramanian and M. Kumar, 2016: Identification of Dominant Source of Errors in Developing Streamflow and Groundwater Projection under Near-term Climate Change. J. Geophys. Res. Atmos., 121, doi:10.1002/2016JD025138.

Singh, H. and A. Sankarasubramanian, 2014: Systematic Uncertainty Reduction Strategies for Developing Streamflow Forecasts utilizing Multiple Climate Models and Hydrologic Models. *Water Resources Research,* **50(2)**, 1288-1307.

Smith, D. M., R. Eade, N. J. Dunstone, D. Fereday, J. M. Murphy, H. Pohlmann, A. A. Scaife, 2010: Skilful multi-year predictions of Atlantic hurricane frequency. *Nat Geosci.* doi:10. 1038/NGEO1004.

Stoner, A. M. K., K. Hayhoe, X. Yang, and D. J. Wuebbles, 2012: An asynchronous regional regression model for statistical downscaling of daily climate variables. *Int. J. Climatol*.. **33**, 2473–2494.

T. Sinha, A. Sankarasubramanian, and A. Mazrooei 2014: Decomposition of Sources of Errors in Monthly to Seasonal Streamflow Forecasts in a Rainfall-Runoff Regime, Journal of Hydrometeorology, 15, 2470–2483.

Taylor, K. E., R. J. Stoufferand, and G. A. Meehl, 2012: An overview of Cmip5 and the experiment design. *B. Am. Meteorol, Soc.*, **93**, 485–98.

Tippett, M.K., M. Barlow, and B. Lyon, 2003: Statistical correction of Central Southwest Asia winter precipitation simulations. *Internat J Climatol*., **23**, 1421–1433.

van Vuuren D. P., et al., 2011: The representative concentration pathways: an overview. *Climatic Change*, **109**: 5-31.

Voldoire, A., et al., 2011 : The CNRM-CM5.1 global climate model: description and basic evaluation. *Clim. Dyn*., DOI:10.1007/s00382-011-1259-y.

Wigley, T. M. L., P. D. Jones, K. R. Briffa, and G. Smith, 1990: Obtaining sub-grid-scale information from coarse-resolution general circulation model output. J. Geophys. Res., 95(D2), 1943–1953. doi:10.1029/JD095iD02p01943.

Wilby, R. L, L. E. Hay., and G. H. Leavesley, 1999: A Comparison Downscaled and Raw GCM Output: Implications for Climate Change Scenarios in the San Juan River Basin, Colorado. J. Hydrology 225, 67–91.

Wilby, R. L., C.W. Dawson, and E.M. Barrow, 2002: SDSM—a decision support tool for the assessment of regional climate change impacts. *Environ. Model Softw*., **17 (2)**, 145–157.

Wilks, D. S., and R. L. Wilby, 1999: The weather generation game: a review of stochastic weather models. *Prog Phys Geogr*., **23(3)**, 329–357.

Wolock, D. M. and G. J. McCabe, 1999: Estimates of Runoff Using Water-Balance and Atmospheric General Circulation Models. J. Amer. Water Resour. Assoc. 35, 1341–1350.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189– 216.

Zhang, F., and A. P. Georgakakos, 2012: Joint Variavle Spatial Downscaling. *Climatic Change*, **111(3)**, 945-972.

Chapter 3. Limitations of Univariate Downscaling Techniques to Preserve Cross-correlation between Climate Variables

3.1 Introduction

Downscaling techniques are often applied on coarser resolution climate variables to translate them to finer resolution for climate change and hydrologic simulation studies (Lettenmaier et al., 1999, Hay et al., 2002, Hanson and Dettinger, 2005, Mejia et al., 2012, Singh et al., 2014). Among two broad categories of downscaling, dynamic (Giorgi et. al., 2001, Hay and Clark 2003, Leung et al., 2003) and statistical (Wood et al., 2004, Wilby et al., 2004, Gangopadhyay et al., 2005, Fowler et al., 2007, Maurer and Hidalgo, 2008), statistical downscaling (SD) is the widely used downscaling technique, because of the ease of model fitting and reduced computation cost. The most commonly employed approach is univariate downscaling (Huth, 1999, Wood et al., 2004 and Stoner et al., 2012) which basically develops an empirical relationship between large-scale climate variables and small-scale observed information. Multivariate techniques, Bivariate Ranks, Joint Variable Statistical Downscaling (JVSD), and Asynchronous Canonical Correlation Analysis (ACCA), have the potential to downscale multiple climate variables and preserve cross-correlation (Zhang and Georgakakos, 2011, He et al., 2012, Mehrotra and Sharma, 2016). However, univariate downscaling remains

a popular choice because of its simplicity (Wood et al., 2004, Maurer and Hidalgo, 2008) and its ability to estimate observed mean and standard deviation in the downscaled products.

In univariate downscaling, downscaling model is developed using observed and General Circulation Model (GCM) simulations of a single variable from the same or nearby grid points. Univariate downscaling approaches vary from simple regular linear regression (Huth, 1999), asynchronous regression (Stoner et al., 2012) to complex semi-parametric (Stoner et al., 2012, Sankarasubramanian and Lall, 2003) and non-parametric approaches (Gangopadhyay et al., 2005) and weather generators (Wilks and Wilby, 1999). Using bias correction and statistical downscaling (BCSD) philosophy (Wood et al., 2004), Bureau of Reclamation (BOR) has developed downscaled monthly precipitation (*pr*) and monthly average temperature (*tas*) dataset from Coupled Model Inter-comparison Project Phase-5 (CMIP5) and Phase-3 (CMIP3) runs over the contiguous United States (CONUS) at 1/8°. BCSD products basically preserve the observed mean and standard deviation of precipitation and temperature but exhibit significance bias in estimating observed cross-correlation between precipitation and temperature.

Over the last century, increased temperature altered the cross-correlation structure partly due to global climate change (Plummer et. al., 1999, Alexander et al., 2006, Qin et. al., 2010, IPCC 2013, Blunden and Arndt 2014, NOAA 2015). Towards this, we proposed a multivariate downscaling technique using Asynchronous Canonical Correlation Analysis (ACCA) that preserves cross-correlation structure between monthly precipitation and monthly average temperature over the CONUS (Das Bhowmik et al., 2016, under review). The study found that the number of grid points with statistically significant observed cross-correlations

is increasing (decreasing) during January and October (April and May). Our work quantified the fraction of grid points over which BCSD is able to keep the fractional bias in cross-correlation values within a limit of +/- 0.25 over the CONUS. Multivariate ACCA developed to preserve the cross-correlation between multiple variables, performed better than BCSD products and univariate approach, Asynchronous Regression (ASR).

Given this information, the first objective of the current study is to perform a systematic analytical decomposition of univariate downscaling approach. An analytical expression for estimating downscaled cross-correlation based on observed data and CMIP5 projections will quantify the limitation of univariate downscaling in preserving cross-correlation between two variables. Constraints to univariate downscaling's performance increase the importance of raw GCM projections. Secondly, the study aims to provide a framework based on statistical calculations for comparing observed and downscaled/raw GCM cross-correlation. The framework will be able to report the confidence associated in estimating cross-correlation.

. The analysis is divided into two broad parts- 1. After analytically relating two cross-correlations (univariately downscaled and the observed), we will assess the overall quality of raw GCM cross-correlation based on the framework. 2. We will investigate spatially over the CONUS to identify the regions over which cross-correlations from raw GCM projections are not within acceptable limits. The analysis would help us to understand the limitations of univariate downscale over space and time, without applying univariate downscaling to the raw dataset. Hence users can select downscaled product based on month, region and with certain confidence to improve the quality of climate forcing of hydrologic simulation.

For this paper, we assumed and applied two univariate linear regression models separately for precipitation and temperature downscaling. A framework is proposed based on Fisher Z-transformation. The framework provides bounds on the downscaled correlation to retain the null hypothesis that downscaled cross-correlation are from the same population as the observed cross-correlation. The proposed framework is independent of the type of downscaling model and can be applied to obtain the bounds for the raw GCM cross-correlation also. Combining the analytical solution from linear regression model with the bounds on the cross-correlation, we investigated the ability of three GCMs in estimating cross-correlations under univariate downscaling. Current study calculates a performance metric, fraction failure, over the CONUS for historical and hindcast runs from three GCMs. The potential of ensemble members of raw GCM projections are calculated on each grid point in terms of estimating cross-correlation. An intense spatial analysis of the performance of univariate downscaling emerges as we extend the analysis to climate region level by calculating fraction failure for climate regions.

Hindcast experiments with increased resolution was introduced in CMIP5 to focus on decadal variability. The archives have ten 10-year hindcasts (1961-1970, 1966-1975,…, 2006-2010) initialized with observed SST every five years from 1960 and two 30-year hindcasts initialized in 1960 (1961-1990) and 1980 (1981-2010). Both 20th century control runs (historical runs) and initialized hindcast runs are project on coarser grid cells thereby require downscaling prior to climate-application studies. Hindcast runs are expected to perform better than historical runs in predicting surface temperature, SST and other ocean circulation features (Keenlyside et al., 2008; Smith et al., 2010, Goddard et al., 2013). For this current

study, we evaluated the performances of both initialized and uninitialized runs under univariate downscaling.

The manuscript is organized as follows: the next section discusses about the GCM and the observed dataset, which is followed by presenting the motivation based on the fractional bias between the observed and downscaled cross-correlation between precipitation and temperature. We describe the analytical derivation of the estimated cross-correlation next and then apply Fisher Z-transformation to obtain the allowable bounds on the downscaled cross-correlation. Section 3.3.2 discusses two metrics to evaluate performance of univariate downscaling over the CONUS. Results are presented in Section 3.5 which is followed with the discussion.

3.2 Dataset

For the study we consider cross-correlation between precipitation and temperature. Variables, *pr* and *tas* values are obtained from two sources, simulations of GCMs and observed dataset. Three GCM models,CNRM-CM5, IPSL-CM5-LR and MPI-ESM-LR from 5th phase of Coupled Model Inter-comparison Project (CMIP5) are considered. These three GCMs have historical experiment runs starting from 1951, and also have hindcast runs that are initialized with observed SST in 1980 (Decadal-80). Details of the dataset are provided in Table 3.1. The analysis presented over 9 climate regions across the CONUS. Nine climate regions are-Northwest (NW), West North Central (WNC), East North Central (ENC), Central (C), Northeast (NE), Southeast (SE), South, Southwest (SW) and West(W). These nine regions

were proposed by National Centers for Environmental Information. These regions, commonly known as NCDC climate regions, are considered useful for putting current climate anomalies into historical perspective (Karl and Koss, 1984).

CNRM-GAME (Centre National de Recherches Météorologiques-Groupe d'étude de l'Atmosphère Météorologique) and CERFACS developed CNRM-CM5 to contribute towards CMIP5 multi-model ensemble. CNRM-CM5 is different from its earlier version in terms of spatial resolution, dynamic core of atmosphere and radiation schemes (Voldoire et al., 2005). The second model, IPSL-CM5-LR, is a single version of the earth system model, IPSL-CM5, developed by Institute Pierre' Simon Laplace. IPSL-CM5 includes all the previous developments, and is extended in terms of degree of complexity, components and resolutions to study responses of climate system (Dufresne et al, 2012). Max-Planck-Institut Fur Meteorologie Earth System Model (MPI-ESM) provides low resolution version called MPI-ESM-LR as our third GCM. MPI-ESM-LR, being a contribution to a wide range of CMPI5 simulations, differs from its previous versions mainly by adding subsystem models for climate processes and inclusion of carbon cycle (Giorgetta et al., 2013). Gridded observed *pr* and *tas* dataset over the CONUS is obtained from Ed Maurer's research group (Maurer et al, 2007). Univariate bias-corrected dataset of three GCMs' historical runs is provided by BOR. GCM raw historical and hindcast runs are further regridded to ~1˚using bilinear interpolation.

Near-term (10-30 years) hindcast experiments initialized (Decadal-80) with observed SST conditions of 1980 and run till 2010 were also used. Historical runs have 5 ensemble members, whereas hindcast runs consist of 9 ensemble members. Cross-correlation between

monthly precipitation and monthly average temperature is calculated for each ensemble member on each grid point. Cross correlations between *pr* and *tas* from raw GCM runs, univariate downscaled GCM data and observed dataset will be henceforth referred as raw GCM cross-correlation, downscaled cross correlation and observed cross-correlation respectively.

Cross-correlations were checked further for their statistical significance using the equation [$\pm 1.96/\sqrt{(n-3)}$] based on 95% confidence interval, where *n* is the number of years of observation. If a grid point exhibits statistically significant observed cross-correlation between *pr* and *tas*, then that grid point and the observed cross-correlation will be henceforth referred as observed significant grid point and observed significant cross-correlation respectively.

Table 3.1: Description of GCM dataset

Model Name	Experiment	Time Length	Ensemble members
CNRM-CM5	Hindcast (Decadal-80)	1981-2010	9
	Historical	1950-1999	5
MRI-ESR-LR	Hindcast (Decadal-80)	1981-2010	3
	Historical	1950-1999	1
IPSL-CM-LR	Hindcast (Decadal-80)	1981-2010	6
	Historical	1950-1999	4

3.2.1 Fractional Bias in Raw GCM Cross-correlation

The motivation to investigate limitations of univariate downscaling comes from the fact that high fractional bias exists between raw GCM and observed cross-correlation between

pr and tas over the CONUS. The analysis will give us an idea about the hurdle univariate approach has to overcome to reduce the fractional bias in estimating cross-correlation in downscaled products. Fractional bias is calculated using the equation below-

$$\text{Fractional Bias} = \frac{\theta^{GCM} - \theta^{obs}}{\theta^{obs}} \qquad \text{Eq. 1}$$

A positive value of fractional bias indicates that GCM is underestimating the observed cross-correlation and vice versa. We calculate fractional biases on grid points over the CONUS for four regions using historical and hindcast runs of CNRM-CM5 (ensemble mean values used to calculate the GCM cross-correlation). Results are presented in Figure 3.1. Additionally we marked the grid points where observed cross-correlations are statistically significant.

During January, historical and hindcast runs both experience underestimation of cross-correlation with fractional bias values more than 1 over Northeastern, Southern and Northwestern parts of the US. However, on observed significant grid points, raw GCM cross-correlation exhibits overestimation with fractional bias values more than 0.75. Most of the grid points during April experience overestimation of cross-correlation by the GCM. Observed significant grid points have fractional bias of 0.75 or more. Hindcast runs witness lesser number of significant grid points than historical runs as sample size for hindcast is smaller than historical. During July, most of the grid points undergo underestimation of cross-correlation. Almost all the Grid points over Southern US experience statistical significance but unlike the performance of the GCM during January and April, CNRM-CM5 exhibits a low fractional bias on the observed significant grid points. Underestimation of cross-correlation by historical runs can also be seen over the central and southern parts, during October. Hindcast runs, during

October exhibits mostly overestimation, with fractional bias values around 0.5. Number of observed significant grid points for these four months are 240, 290, 493 and 244 respectively for the time period 1950-1999 (historical) and 227,223, 394 and 318 respectively for the time period 1981-2010.

Overall the fractional bias values for historical and hindcast runs may differ in magnitude but follow similar patterns. Later in this paper, we will quantify the limitations of univariate downscaling by applying an allowable bound on GCM cross-correlations. In the discussion, we briefly discussed how we can combine fractional bias with the allowable bound. It would be expected from the downscaling/bias-correction scheme to reduce the fractional bias such a way the downscaled/bias-corrected cross-correlation rests within the allowable bounds created with a confidence interval of 95%.

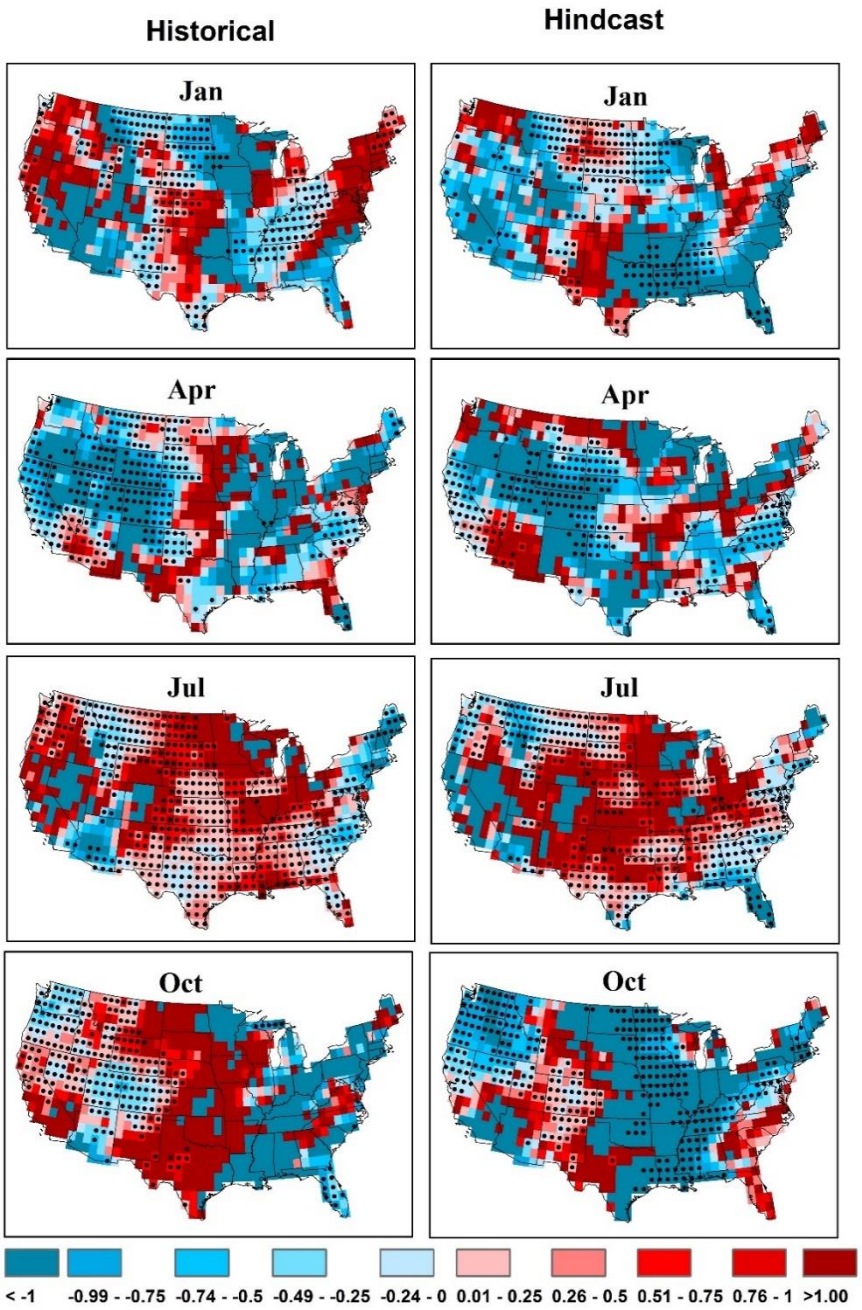

Figure 3.1: Fractional bias in estimating cross-correlation, CNRM-CM5 historical and Hindcast runs. Observed significant grid points are marked

3.3 Conceptual Framework of Univariate Downscaling

3.3.1 Basic Assumptions

Linear regression is considered as the most popular univariate downscaling approach to bring coarse-resolution climate variable to finer-resolution. Computational simplicity of the linear model and ability to preserve the mean of the observed variable during downscaling are the prime reasons behind linear regression's popularity. Linear regression considers observed and GCM projected raw values from past period for model development and estimating regression parameters (Eq. 2 and Eq. 3) which are then applied on the future GCM projections to obtain the downscaled GCM values.

We assumed a univariate linear regression model at each grid point for downscaling. Regression models are individually applied on *pr* and *tas* by relating observed as predictant and GCM estimates as predictor. Model structures are given below-

$$\hat{P}_{down} = \hat{a}.P_{GCM} + \hat{b} \quad \text{Eq. 2}$$

$$\hat{T}_{down} = \hat{c}.T_{GCM} + \hat{d} \quad \text{Eq. 3}$$

Model parameters [a,b] and [c,d] are estimated based on observed (P_{obs}, T_{obs}) and GCM projected raw precipitation and temperature (P_{GCM}, T_{GCM}) and P_{down} and T_{down} are the downscaled precipitation and temperature. Cross correlations between $\hat{P}_{down}$ and $\hat{T}_{down}$are derived analytically.

$$\frac{Corr\,(P_{down},T_{down})}{Corr(P_{obs},T_{obs})} = E.\frac{Corr(P_{GCM},T_{GCM})}{Corr\,(P_{obs},T_{obs})} \quad \text{Eq. 4}$$

where E is dependent on model slopes, standard deviations of GCM and downscaled variables. Mathematical decomposition of E shows that the value of E will always be 1 (see Supplementary Information A). So, the equation comes down to

$$\frac{Corr(P_{down},T_{down})}{Corr(P_{obs},T_{obs})} = \frac{Corr(P_{GCM},T_{GCM})}{Corr\ (P_{obs},T_{obs})} \quad \text{Eq. 5}$$

Eq. (4) is of particular importance. Irrespective of asynchronous or regular linear regression, Eq. (4) holds true as long as the linear model structure is applied for downscaling. It could be interpreted as follows: when univariate downscaling is applied individually on multiple variables using linear regression model, the downscaled product will have the same cross-correlation as the raw GCM cross correlation.

3.3.2 Fisher Z-Transformation

Since the raw GCM projections play a crucial role in estimating the cross correlation in downscaled dataset, it is important to assess quality of raw GCM cross-correlation. To verify whether the raw GCM cross-correlation is statistically different from the observed cross-correlation, we consider Fisher z-transformation. R. A. Fisher proposed z-transformation of sample correlation (Fisher 1921) when distribution of sample correlation is skewed and population correlation is not equal to zero. Z is a function of sample correlation with a sampling distribution close to normal. Fisher z-transformation is used to test the null hypothesis that two independent correlations are significantly different from one another (Fisher 1950). Statistical significances of the sample cross-correlations are not important during Fisher z-transformation analysis. The current study assumes downscaled and observed

variables follow bivariate normal distribution and applies Fisher z-transformation on them. Fisher z-transformation forced the correlations to follow normal distribution with standard deviation equal to $1/\sqrt{(n-3)}$. The test statistic,$\widehat{Z}$, to retain or reject the null hypothesis, is defined to follow the standard normal distribution. The test statistic is finally back-calculated based on observed cross-correlation, length of dataset and confidence interval to form an allowable bound for downscaled cross-correlation.

We have two cross correlations to compare- one from the downscaled/raw GCM (ρ^{Down} or ρ^{GCM}) and another one from the observed dataset (ρ^{Obs}). Fisher z-transformation applied on them to obtain Z_{obs} and Z_{GCM}[in general, $Z = 0.5\ln(\frac{1+\rho}{1-\rho})$]. Our interest is to determine whether these two cross correlations can be considered same with certain level of confidence. Hence, the test statistics $\widehat{Z}$ is introduced, $\widehat{Z}$ follows standard normal distribution with zero mean and standard deviation equal to one. $\widehat{Z}$ Can be calculated as-

$$\widehat{Z} = \frac{Z_{obs} - Z_{GCM}}{\sigma_{Z_{obs} - Z_{GCM}}} \qquad \text{Eq. 6}$$

As sampling distribution of $\widehat{Z}$ follows standard normal distribution, 95% confidence interval for $\widehat{Z}$ should be within [-1.96, +1.96]. [$\rho_{high}^{Down/GCM}, \rho_{low}^{Down/GCM}$] are the allowed values of downscaled/GCM cross-correlation given the observed cross correlation, number of years (*n*) and confidence interval. We back-calculated $\widehat{Z}$ to find out the allowable bounds of downscaled cross-correlation which must be same in value as raw GCM cross correlation (see Supplementary Information B).

$$\rho_{low}^{Down/GCM} = \frac{(C-1)+\rho^{obs}(C+1)}{(C+1)+\rho^{obs}(C-1)}] \quad \text{Eq. 7a}$$

$$\rho_{high}^{Down/GCM} = [\frac{1-C+\rho^{obs}(1+C)}{1+C+\rho^{obs}(1-C)}] \quad \text{Eq. 7b}$$

Where, C=1.96$\times \frac{2\sqrt{2}}{\sqrt{(n-3)}}$ (for 95% CI)

Analytical solution of univariate downscaling, considering a representative linear regression model, reveals that univariate techniques do not possess any additional capability to preserve observed cross correlation during downscaling. However, there could be one possibility that raw GCM cross-correlation might rest within the allowable bound over certain region and for specific months. Then we could use downscaled products for actual application ignoring limitations of univariate downscaling.

Eq. (7a) and (7b) are used to generate band of allowable raw GCM cross-correlation by varying observed cross-correlations between -1 to 1. This synthetic simulation would help us to understand the influence of observed cross-correlation, length of the dataset and confidence interval on the allowable range of downscaled cross-correlation. Two scenarios with 95% confidence interval (CI) are constructed using different length of datasets (*n*=30 and 50). Simulation results are shown in Figure 3.2. We overlaid the raw GCM cross-correlations from historical and hindcast runs on the allowable bound. Lower values of observed cross-correlation have witnessed higher bandwidth of allowable ranges whereas observed cross-correlation values close to -1 or 1 have narrower allowable bounds. This allowable downscaled cross-correlation bounds become wider to narrower as number of observation increases. A *pr* and *tas* simulation for 30 years dataset has higher scope of reaching allowable range that a

simulation of 50 years in terms of preserving cross-correlation. Allowable bounds are also wider for higher CI. To quantify the quality of raw GCM cross-correlation, we basically calculated the number of grid points that rest outside of the allowable bound.

Quality of raw GCM cross-correlations in preserving observed cross-correlations will be assessed based on Fisher z-transformation. Results reflect the potential of univariate downscaling. First, we will evaluate the performance of each ensemble members (p) in estimating cross-correlation on a particular grid point. We determine the limitations of univariate downscaling during 12 months based on a metric, called fraction failure (*f*). Fraction failure calculates the fraction of grid points over the CONUS that have raw GCM cross-correlation values beyond the allowable range Steps to calculate failure fraction and number of ensemble members failing are as follows-

1. At grid point *i (i*=1…*L),* month *m (m*=1…12*)* and for ensemble member *j (j*=1…*J)* extract the observed and GCM simulated cross correlations between precipitation and temperature, respectively $\rho_{obs}^{i,m}$ and $\rho_{GCM}^{i,m,j}$. *L* represents the total grid points (or observed significant grid points) within a region.

2. Based on $\rho_{obs}^{i,m}$ and number of years of Hindcast/simulations to calculate the allowable limits of downscaled cross correlation, $[\rho_{high}^{i,m}, \rho_{low}^{i,m}]$ using eqn. 7a and eqn. 7b. Based on that calculate the allowable ratio $[r_{high}^{i,m}, r_{low}^{i,m}]$ by dividing allowable cross correlations by $\rho_{obs}^{i,m}$. Allowable ratio doesn't vary between ensemble members.

3. As we know, for univariate downscaling with regular linear regression model ($\rho_{GCM}/\ \rho_{obs}$=ρ_{down}/ρ_{obs}), so without performing downscaling, we directly calculated r_{act}.

$$r_{act}^{i,m,j} = \frac{\rho_{GCM}^{i,m,j}}{\rho_{obs}^{i,m}} \quad \text{Eq. 8}$$

4. If $r_{act}^{i,m,j}$ is not within the bounds of $[r_{high}^{i,m}, r_{low}^{i,m}]$ we marked the grid point *i* as $I^{i,m,j}$=1, otherwise $I^{i,m,j}$=*0*, which ensures that the marked grid point is likely to fail if univariate downscaling is performed.

5. Steps 1 to 4 repeated for all grid points and all ensemble members for a month *m*.

6. Sum up the number of ensemble members to obtain ($p^{i,m}$) that have cross-correlations beyond Fisher's limit. $p^{i,m}$ represents number of ensemble members for grid point *i* and month *m* that would fail to preserve cross-correlation once univariate downscaling applied on the member's raw GCM outputs.

$$p^{i,m} = \sum_{j=1}^{J} I^{i,m,j} \quad \text{Eq. 9}$$

7. For a particular month, the fraction (f^m) of univariate downscaling are calculated using all ensemble members and all grid points.

$$\text{Failure fraction } f^m = \left[\frac{\sum_{i=1}^{L}\sum_{j=1}^{J} I^{i,m,j}}{J \times L}\right] \quad \text{Eq. 10}$$

8. Steps 1 to 7 are repeated for all grid points i=1 to L within a region and for all months m=1 to 12.

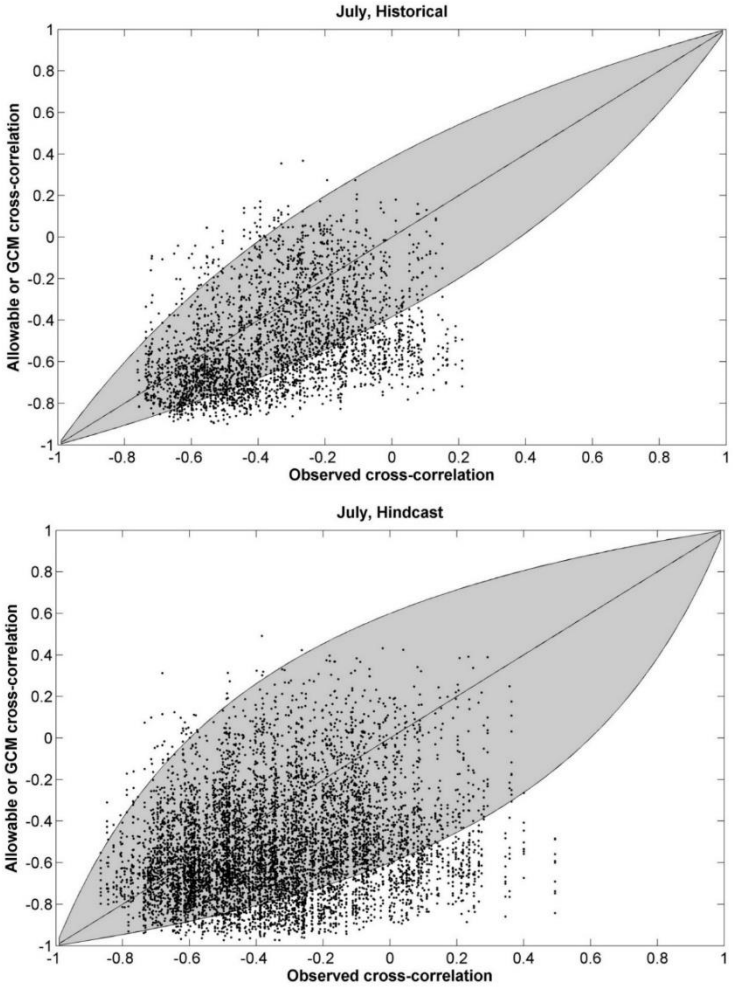

Figure 3.2: Raw GCM cross-correlations from CNRM-CM5 historical and hindcast runs (July) overlaid on allowable bounds. Allowable bounds are calculated using Fisher z-transformation.

3.4 Results

3.4.1 Quality of GCM ensemble members

Performance of CNRM-CM5 ensemble members on the observed significant grid points are evaluated in terms of estimating cross-correlation. Figure 3.3 (historical) and 3.4

(hindcast) show the observed significant grid points over the the CONUS for the time periods 1950-1999 and 1981-2010 respectively. Number of ensemble members that fail to estimate the cross-correlation within the allowable range, is indicated on each grid points with color. Gradual increase in red color from palest to the darkest represents null to all members within an ensemble.

Historical runs from CNRM-CM5 has 5 ensemble members. During January, 2 or more than 2 members exhibit raw GCM cross-correlations over the Ohio valley, New England and Florida that are out of the allowable bound. Observed significant grid points over Kansas, Nebraska, Texas and New Mexico generally experience 1 or no member failing to estimate the cross-correlation. However, grid points over Montana and North Dakota exhibit failure for more than the half of the members within an ensemble. During April, ensemble members are more prone to failure over the Western US (California, precisely). Almost all grid points over Colorado and Nebraska have 4 or 5 ensemble members that are expected to fail under the univariate downscaling. Whereas, on the Carolinas and Virginia, at least one member within the ensemble fails to estimate the cross-correlation. During July, almost half of the total grid points over the CONUS exhibits statistical significance. Grid points over the Dakotas and Nebraska are critical during July as more than half of the members on each grid point have cross-correlations that rest outside of the allowable bound. Observed significant grid points over the Sunbelt experience moderately better performance by CNRM-CM5 ensemble in estimating the cross-correlation. However, within the Sunbelt, small clusters of 3 to 4 observed significant grid points exist, over Colorado, Georgia and Alabama on which 50% probability remains that an ensemble member would fail under the univariate downscaling. During

October, grid points over Utah, Nevada and Florida exhibit half or more than the half of the ensemble members with GCM cross-correlations that are out of the bound.

CNRM-CM5 hindcast runs consist of 9 ensemble members and during January, typically 7 members fail to estimate the cross-correlation on observed significant grid points over Mississippi and Alabama. Grid points over the Northern US experience relatively improved performance compared to the Southern US as 1 or more than one members within an ensemble exhibit cross-correlations that are out of the allowable bounds. Almost none of the ensemble members is susceptible under the univariate downscaling for grid points over New Mexico, Central US and southern Texas. During April, grid points over Wyoming, Utah, Colorado and Nebraska witness susceptibility under the univariate downscaling over more than half of the members. Similar to these four states, eastern states, Carolinas and Virginia also exhibits similar results.

GCM's quality improves over western grid points where on each grid point, typically 4 or less than 4 members are prone to failure. Observed grid points for July are almost half of the total grid points that cover the CONUS. However, the GCM quality in estimating the cross-correlation for hindcast runs are not very severe during July. Except few grid points over Texas, North Carolina and Kansas, usually 1 to 2 members fail to estimate the cross-correlations over the CONUS. But, the ensemble of CNRM-CM5 performs weakly during October. Almost all the members within an ensemble, on grid points over Wisconsin, Illinois, Iowa and Minnesota fail to estimate the cross-correlation. Grid points over Southeast also witness similar results as their Northern counterparts. Severity of GCM in estimating the cross-correlation is further

reduced over the north-western grid points where 4 to 5 members within an ensemble are typically prone under the univariate downscaling.

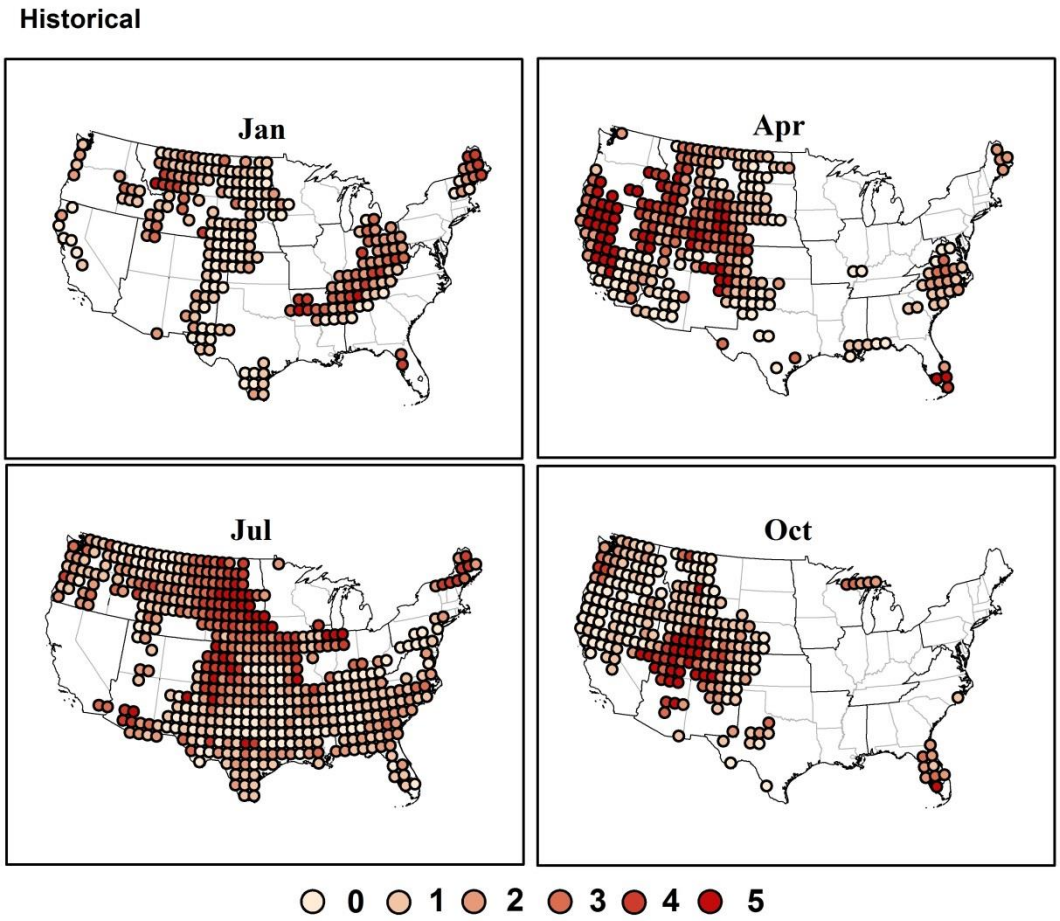

Figure 3.3: Spatial maps show number of ensemble members that have raw GCM cross-correlations (from CNRM, Historical Simulation, 1951-1999) out of allowable bound.

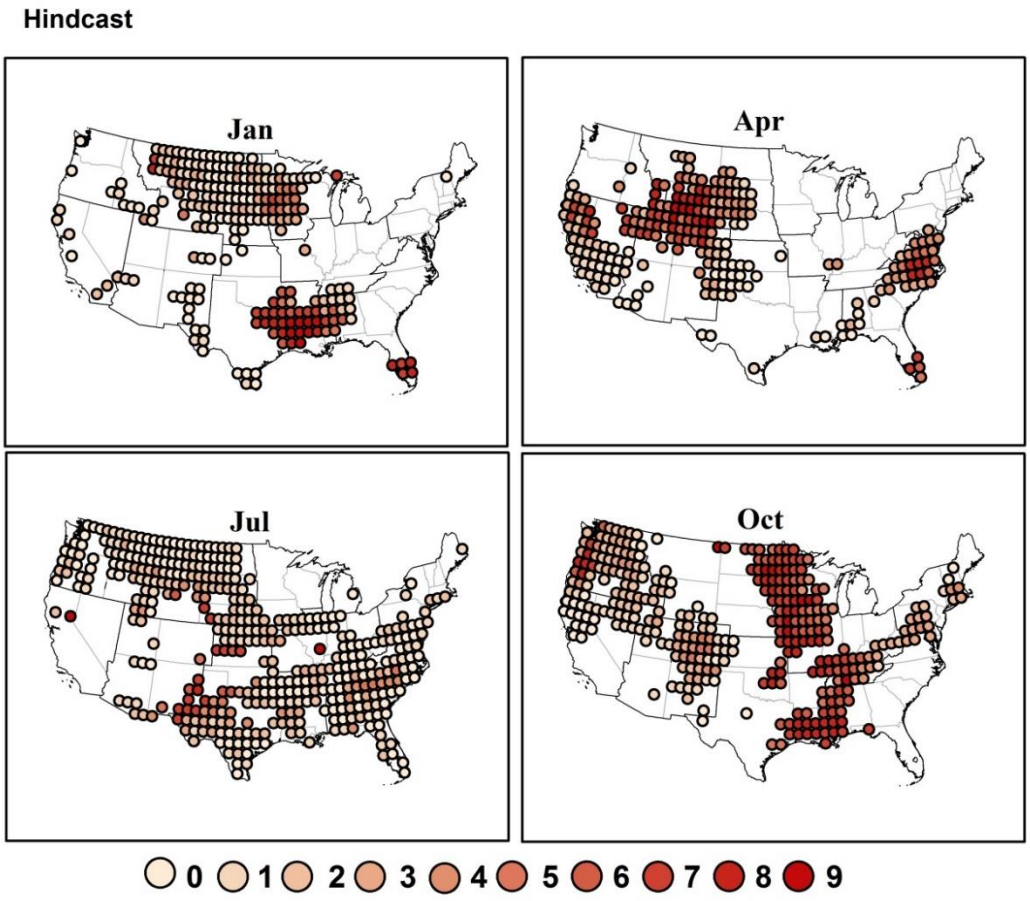

Figure 3.4: Spatial maps show number of ensemble members that have raw GCM cross-correlations (from CNRM, Hindcast Simulation, 1981-2010) out of allowable bound.

3.4.2 Fraction failure across GCMs

Fraction failure represents the fraction of total grid points within a geographic region on which GCM cross-correlations fail to rest within the allowable bound defined by Fisher z-transformation. The final value of fraction failure is calculated by taking average over ensemble members. Fraction failure values for 3 GCMs (CNRM-CM5, IPSL and MPI) are

calculated over the CONUS and results for 12 months are shown in Figure 3.5. Results for historical and hindcast runs are shown in left and right columns respectively while the rows represent individual GCMs. We considered observed significant grid points as the total grid points for this analysis. Fraction failure values from the historical runs of CNRM and MPI, around 0.5, are higher for March and December than the other months. IPSL historical runs exhibit highest value of fraction failure, 0.6 during December and higher values of fraction failure, 0.4 during March, July and August. During July, where half of the grid points are statistically significant, almost 40% of the observed significant grid points from all models would fail under the univariate downscaling. All 3 models across 12 months follow more or less similar performance in estimating the cross-correlation. However, IPSL has slightly lower fraction failure values compared to the fraction failure values of the other 2 models.

Hindcast runs from CNRM and MPI witness highest fraction failure values compared to other months, 0.4 and 0.45 respectively during April and October. January and October are most critical for IPSL hindcast runs in estimating the cross-correlation with fraction failure values of 0.4 and 0.55 respectively. Fraction failures for CNRM and MPI are both around 0.15 during July, the month with most number of observed significant grid points. Lower values of fraction failures are often witnessed during the summer months which signifies that GCM often has the ability to estimate summer negative cross-correlation between *pr* and *tas*. However, GCM performance in estimating cross-correlation does not improve across GCMs.

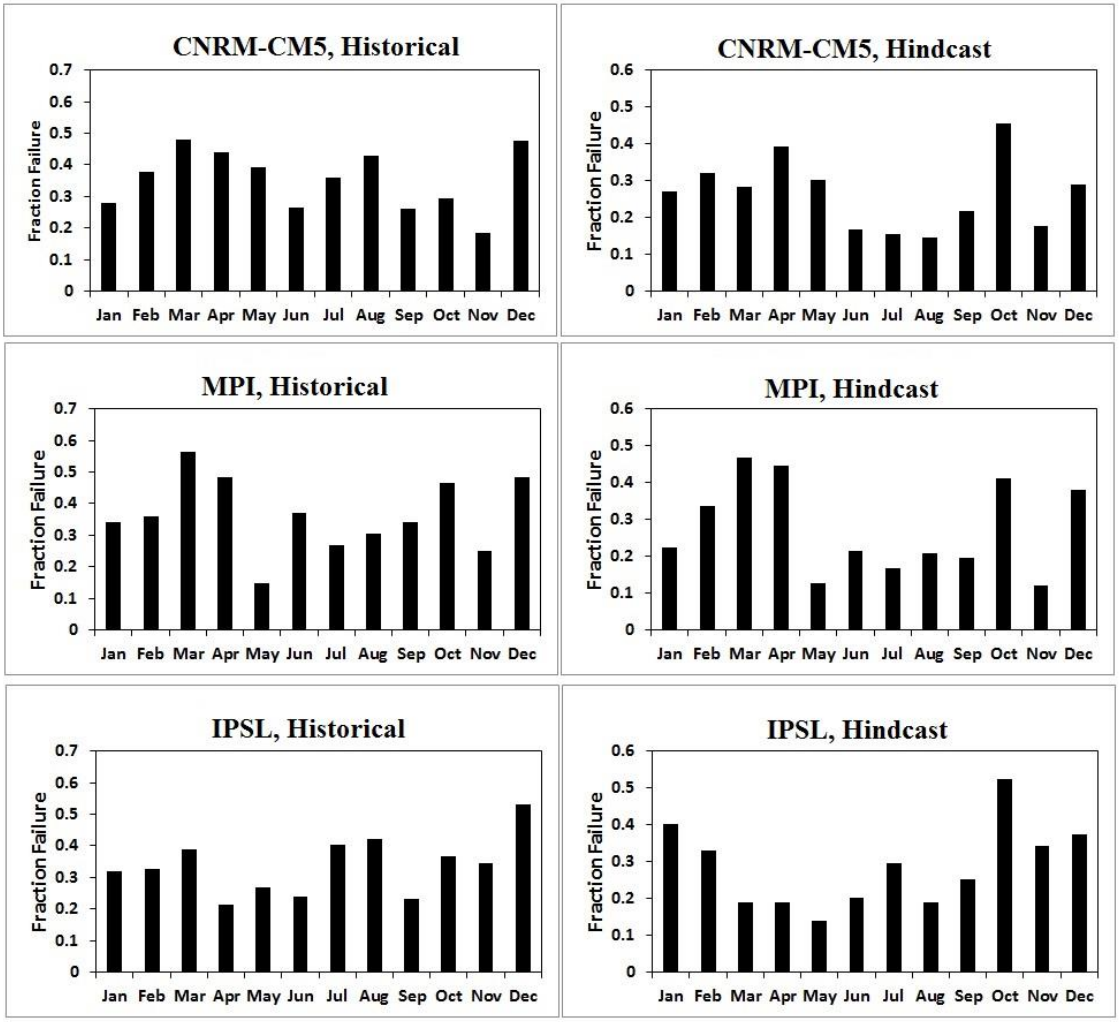

Figure 3.5: Fraction failure for three GCMs' (CNRM-CM5, MPI and IPSL) historical runs, calculated over the CONUS. Fraction failure calculated for 95% CI.

3.4.3 Fraction failure across climate regions

As it is evident that performance is not improving between GCMs, we calculated fraction failure values for only CNRM-CM5 across climate regions instead calculating over the entire CONUS. Fraction failure values across climate regions are expected to quantify the

performance of the GCM at climate region level and reveal the susceptibility or superiority of individual regions under the univariate downscaling to the user community of downscaled products. Results are shown under Table 3.2 (historical) and Table 3.3 (hindcast).

WNC, Central, South and SW regions experience higher numbers of observed significant grid points during January compared to the other the climate regions. WNC exhibits 87 observed significant grid points and one fourth of the 87 grid points would fail under the univariate downscaling. Susceptibility of grid points further worsen over the Central US where 47% of the 41 observed significant grid points fail to estimate cross-correlation. Central US experience strong cross-correlation between *pr* and *tas* which is generally positive during January and negative for July. We have shown in Chapter 3 that the observed significant grid points during January are shifting to become non-significant over the second half of the 20th century. GCM fails to capture the change which can be seen in the fractional bias plot also (Fig 3.1) where the fractional bias in estimating cross-correlation during January is very high.

During July, 35% of the 50 observed significant grid points fails over the Central US. During April, WNC, SW and West regions have 82, 68 and 59 observed significant grid points respectively. Almost half of the grid points over the Southwest and the West exhibit GCM cross correlations that are out of the allowable bound. Whereas fraction failure over the WNC has a value of 0.38, lesser than the SW and the West. During April, NW witnesses 24 observed significant grid points which is lesser compared to the number of grid points other climate regions witness but almost 80% of the grid points over the NW fails to estimate the cross-correlation. During July, WNC and South shares a high amount of observed significant grid

points over the CONUS. GCM has better performance over the South, where 26% grid points out of total 135 are failing, compared to WNC where fraction failure is 0.53 over 118 grid points. GCM cross-correlations are of moderate qualities over the SW and the NW. Failure fraction over SW and NW are 0.36 and 0.32 respectively. During October, more than half of the 67 observed grid points fails to estimate the cross-correlation. Whereas, WNC has a failure function value of 0.2, out of 53 observed significant grid points.

Cross-correlation estimation for the hindcast runs are less precise than the historical runs as number of years of observation is lesser. Hence more GCM cross-correlations rest within the allowable bound, leading to a lower fraction failure. During January, WNC experiences highest number of observed significant grid points (98) compared the other climate regions and has a fraction failure value of 0.14. Next to WNC, Southern US has 50 observed significant grid points, out of which 58% are expected to fail under the univariate downscaling. However, GCM performance in estimating cross-correlation over the WNC further weakens during April, compared to January. During April, 54% of the total 59 observed significant grid points fails over the WNC. For the Southwest also, fraction failure is 0.42 considering 46 observed significant grid points. During July, hindcast runs exhibits failure fraction of 0.13 over 106 observed significant grid points of WNC. NW, SE and South have 41, 63 and 90 observed significant grid points respectively and corresponding failure fractions are 0.66, 0.13 and 0.21. ENC and Central US experiences 46 and 45 observed significant grid points, respectively during October. GCM performance in estimating the cross-correlation during October is critical over ENC and Central climate regions with fraction failure values of 0.74

and 0.63 respectively. Northwest experiences 60 observed significant grid points but fraction failure is 0.33, a lower ratio compared to the ENC and the Central.

Table 3.2: Fraction failure for raw GCM cross-correlation across NCDC climate regions (CNRM, historical simulations, 1951-1999).

Month	NW	WNC	ENC	C	NE	SE	S	SW	W
Jan	0.3 (14)	0.22 (87)	0.23 (6)	0.47 (41)	0.46 (13)	0.29 (7)	0.26 (36)	0.18 (28)	0.09 (7)
Apr	0.78 (24)	0.38 (82)	-	0 (2)	0.4 (4)	0.33 (28)	0.16 (23)	0.44 (68)	0.57 (59)
Jul	0.31 (46)	0.53 (118)	0.58 (10)	0.35 (55)	0.41 (19)	0.24 (58)	0.26 (135)	0.36 (50)	0.6 (1)
Oct	0.17 (60)	0.2 (53)	0.48 (5)	-	-	0.46 (10)	0.18 (10)	0.56 (67)	0.13 (38)

Table 3.3: Fraction failure for raw GCM cross-correlation across NCDC climate regions (CNRM, hindcast simulations, 1981-2010).

Month	NW	WNC	ENC	C	NE	SE	S	SW	W
Jan	0.02 (9)	0.14 (98)	0.33 (31)	0.17 (4)	0 (1)	0.51 (11)	0.58 (50)	0.08 (16)	0.13 (6)
Apr	0.49 (7)	0.54 (59)	-	0.33 (3)	0.41 (3)	0.48 (34)	0.02 (21)	0.42 (46)	0.28 (50)
Jul	0.06 (41)	0.13 (106)	0.13 (6)	0.11 (49)	0.1 (14)	0.13 (63)	0.21 (90)	0.4 (23)	0.61 (2)
Oct	0.33 (60)	0.49 (27)	0.74 (46)	0.63 (45)	0.27 (15)	0.61 (4)	0.7 (41)	0.24 (50)	0.07 (30)

3.4.4 Observed cross-correlation versus GCM Performance

So far we have shown, GCM's efficiency to estimate the cross-correlation vary across climate regions and months. We also expect the efficiency to change with varying magnitude of the observed cross-correlation values. For example, on certain months, GCM might perform better in estimating the lower values of observed cross-correlation than the higher values. This assumption rely on the shape of the allowable bound. Allowable bound of GCM cross-

correlation is one of the functions of observed cross-correlation and the bound gradually decreases from wider to narrower as absolute value of the observed cross-correlation increases. The very shape of the allowable bound gives us the notion that the failure is expected to be higher when observed cross-correlation values are very high. However in reality, GCM's performance in estimating the cross-correlation can be different from what we have assumed because the number of grid points resting within an interval of observed cross-correlation values finally determines GCM's performance. In Figure 3.6, we show GCM's efficiency in estimating the cross-correlation across different intervals of the observed cross-correlation.

We plotted the fraction of grid points that fails to the estimate cross-correlation for different intervals of absolute values of the observed cross-correlations. Total number of grid points within an interval is shown in the secondary axis. The calculation is performed for 4 months by considering a single member from the CNRM-CM5 ensemble. Statistical significance of the observed cross-correlations are indicated with vertical straight lines (different for historical and hindcast runs) within each panel.

Number of grid points within an interval decreases during January and April as the observed cross-correlation increases. Historical runs during January attains highest failure when the observed cross-correlation interval is slightly greater than the statistical significance value. For hindcast runs also, if we ignore few grid points that have cross-correlations of very high value, failure is higher around the observed significant cross-correlation. During April, failure rate is higher for the lower values of the observed cross-correlation and gradually increases as the observed cross-correlation increases. This is similar to what we expected. Continuously decreasing number of total grid points within an interval leads to an almost

continuous increase in the failure rate of historical and hindcast runs. Whereas during July, most number of grid points experiences statistically significant observed cross-correlation and CNRM-CM5 fails over half of the grid points (for historical runs) for the interval of 0.3-0.4. Higher number of grid points, compared to April, are concentrated for higher values of the observed cross-correlation hence GCM performance typically decreases for higher values of the observed cross-correlation. A wider allowable bound for hindcast runs enhances GCM performance and the failure rate follows a decreasing pattern with increase in the observed cross-correlation values. October results are very similar to that of April. Number of grid points are decreasing as the observed cross-correlation increases. But fraction of grid points that fail to estimate the cross-correlation, remains constant across observed cross-correlation intervals. In summary, we infer that the failure rate in estimating the cross-correlation typically increases as the observed cross-correlation magnitude increases. Cases when more number of grid points experience higher values of the observed cross-correlations, GCM's failure rate starts decreasing for higher values of the observed cross-correlations.

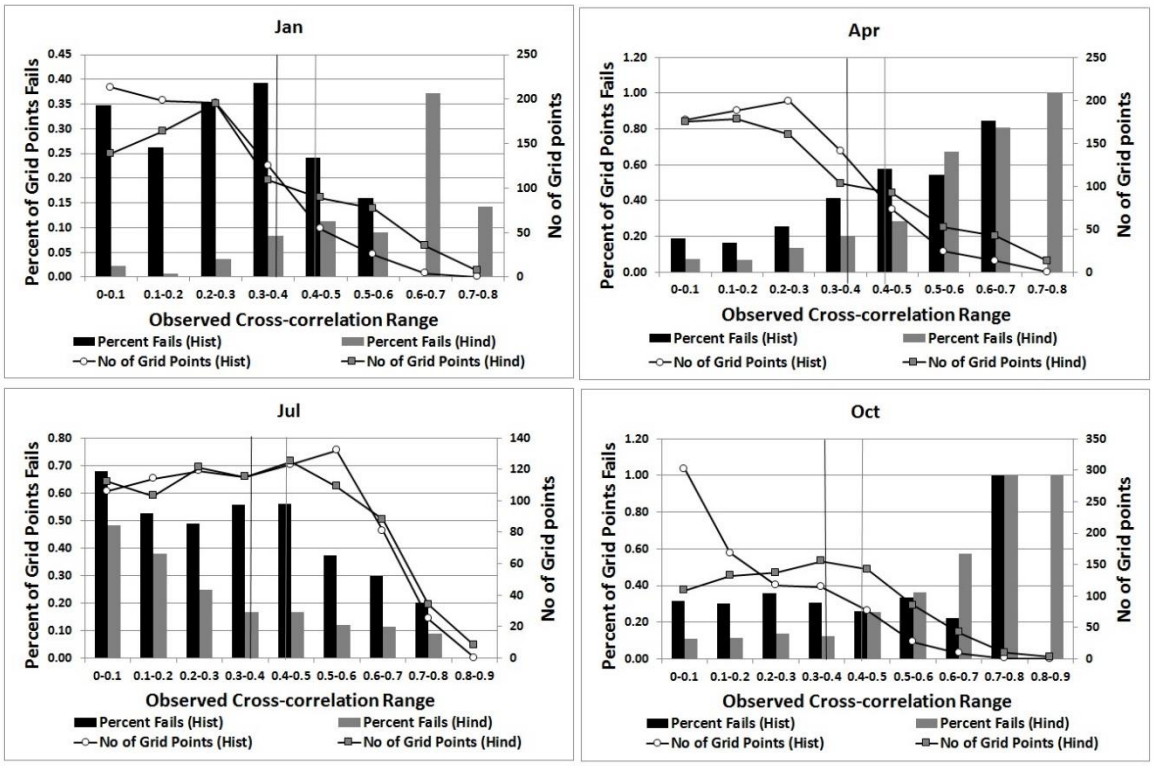

Figure 3.6: Percent of grid points fails to estimate cross-correlation for different intervals of observed cross-correlation. Number of grid point within an interval plotted in the secondary axis. Single ensemble member from CNRM-CM5 is considered for the study.

3.5 Discussion

In this study, we investigated the limitations of univariate downscaling to estimate cross-correlation between multiple variables. A univariate, regular linear regression model assumed and applied separately on climatic variables to downscale. We found that univariate downscaling does not have the potential to translate the observed cross-correlation to downscale variables. Rather, cross-correlation after univariate downscaling remains the same

as raw GCM cross-correlation. Hence raw GCM cross-correlation plays crucial role in estimating cross-correlation. Limitations of univariate downscaling might differ across different downscaling models, but the limitations in estimating cross-correlations are not expected to improve much as none of these models consider the coherence between multiple variables.

Our study shows that a high amount of fractional bias exists between observed and GCM cross-correlations between precipitation and temperature simulations from 20th Century runs (historical) and decadal hindcasts. Under this circumstances, it is important to assess the quality of raw GCM projections in terms of estimating cross-correlation. Based on Fisher z-transformation, we proposed a framework to ascertain whether downscaled/raw GCM cross-correlation and observed cross-correlation belong to the same population of cross-correlation with certain confidence interval. The allowable bound for raw GCM cross-correlation is a function of confidence interval, number of observations and observed cross-correlation. The framework we proposed is model independent, hence can be applied to check multivariate downscaled cross-correlations or cross-correlations from complex univariate methods. Fisher z-transformation assumes correlations are from bivariate normally distributed variables. Precipitation and temperature values on grid points may not follow bivariate normal distribution. Hence a prior check of multivariate normality, such as Cox-Small test, should improve the precision in estimating downscaled cross-correlation.

The framework is applied on GCM cross-correlations from historical and hindcast runs to evaluate the quality of raw GCM projections. Results show that for historical runs, around 40-50% of the observed significant grid points over contiguous US are prone to failure in terms

of cross-correlation during univariate downscaling. This percentage improves to 30-40% when the framework is applied on hindcast runs, as the CI for hindcast runs are wider. Smaller sample size for hindcast runs form a wider bound for downscaled cross-correlation. Grid points that are susceptible under univariate downscaling in terms of cross-correlation, are spread across the CONUS. At least one member within an ensemble of GCM fails to estimate cross-correlation and there exists multiple grid points where all the ensemble members fail. We extended our analysis to NCDC climate regions which reveals that failing grid points are randomly scattered over all climate regions. However, WNC and SW witness higher fraction failure compared to other climate regions. In conclusion, performance of GCMs to estimate cross-correlation is not concentrated around a particular model or climate region. A high probability always remains that downscaled products will fail to preserve the observed cross-correlation if univariate downscaling is performed. GCM failure typically enhances when observed cross-correlations are statistically significant.

As an extension to the work done so far, we are interested in combining the fractional bias in estimating cross-correlation (described in Chapter 2.1) with the allowable bound formed using Fisher z-transformation. Our previous work used the fractional bias in quantifying the improvement done by multivariate downscaling approach (ACCA) in estimation of cross-correlation. If the limitations of univariate approach persists across models and over other regions, users of downscaled products would have to rely more on the raw GCM outputs. Adjustment of the allowable bound to retain the null hypothesis made in Chapter 3.2, by making the bound wider, can improve performance of GCM. However, such adjustment has to be done by increasing the CI which eventually would lead to a higher type-I error in

estimating cross-correlation. Given an observed cross-correlation, as the fractional bias in estimating cross-correlation increases corresponding CI to retain the null hypothesis also increases. In cases when GCM witness high fractional bias in estimating cross-correlation and observed cross-correlations are also high, we advise not to increase the CI and retain the null hypothesis to cover up the fractional bias.

Simulations of various hydrologic fluxes (ex: Soil Moisture and Overland flow), depending on the type of basin and statistical significance of observed cross-correlation, rely on climate variables that exhibit preserved cross-correlation. We expect this paper will help the user community of downscaled products to select the appropriate downscaling/bias-correction approach. We applied the allowable bound on cross-correlations derived from multivariate bias-corrected (ACCA) dataset and the results shows that multivariate cross-correlations always rest within the allowable limits.

3.6 Supplementary Information (SI)

A. Analytical Analysis of Univariate Approach

Linear model is fitted between $[P^{GCM}, T^{GCM}]$ and $[P^{obs}, T^{Obs}]$,. Model parameters are a, b, c and d. ξ^p and ξ^t are the model errors.

$$P^{Obs} = aP^{GCM} + b + \xi^p \quad \text{Eq. 1}$$

$$T^{Obs} = cT^{GCM} + d + \xi^t \quad \text{Eq. 2}$$

Model parameters can be expressed as-

$$a = \frac{cov(P^{GCM}, P^{Obs})}{var(P^{GCM})}, c = \frac{cov(T^{GCM}, T^{Obs})}{var(T^{GCM})},$$

Downscaling products are estimated using model parameters-

$$P^{Down} = aP^{GCM} + b \quad \text{Eq. 3}$$

$$T^{Down} = cT^{GCM} + d \quad \text{Eq. 4}$$

Hence,

$$cov(P^{Down}, T^{Down}) = cov\,(aP^{GCM} + b, cT^{GCM} + d) = a.c.cov(P^{GCM}, T^{GCM})$$

Or, $\sigma_{PDown}.\sigma_{TDown}.cov(P^{Down}, T^{Down}) = a.c.\sigma_{PGCM}.\sigma_{TGCM}.cov(P^{GCM}, T^{GCM})$

Or, $\frac{cov(P^{Down}, T^{Down})}{cov(P^{GCM}, T^{GCM})} = a.c.\frac{\sigma_{PGCM}.\sigma_{TGCM}}{\sigma_{PDown}.\sigma_{TDown}} = E$

Or, $E = a.c.\frac{\sigma_{PGCM}.\sigma_{TGCM}}{\sigma_{PDown}.\sigma_{TDown}} = a.c.\frac{\sigma_{PGCM}.\sigma_{TGCM}}{a.\sigma_{PGCM}.c.\sigma_{TGCM}}$ [as, $\text{var}(P^{Down}) = a^2 var(P^{GCM})$ and $\text{var}(T^{Down}) = C^2 var(T^{GCM})$ from model formulation.]

So, E=1

B. Fisher z-transformation

ρ^{Obs} *and* ρ^{Down} are the observed and downscaled cross-correlations

Null Hypothesis: $\rho^{Obs} = \rho^{Down}$

Z-Transformation:

$$\rho^{Obs} = 0.5.\ln[\tfrac{1+\rho^{Obs}}{1-\rho^{Obs}}] \quad \text{Eq. 1}$$

$$\rho^{Down} = 0.5.\ln[\tfrac{1+\rho^{Down}}{1-\rho^{Down}}] \quad \text{Eq. 2}$$

Test Statistic: $\tilde{z} = \dfrac{Z^{Obs} - Z^{Down}}{\sigma_{ZObs-ZDown}}$ Eq. 3

$\tilde{z}$ follows standard normal distribution, $\tilde{z} \sim N(0,1)$ and $\sigma_{Z^{Obs}-Z^{Down}} = \sqrt{\frac{2}{n-3}}$. 'n' is the number of observations.

Bounds for test statistic $\tilde{z}$ are [-1.96, 1.96] and [-2.58, 2.58] to retain the null hypothesis with 95% or 99% CI respectively.

Let us assume, the allowable range for downscaled cross-correlation is $[\rho_{high}^{Down}, \rho_{low}^{Down}]$.

Considering the lower confidence interval for $\tilde{z}$, we back-calculated the highest value of allowable downscaled cross-correlation

$$\tilde{z}_{low} = -CI$$

$$or, \frac{1}{2.\sigma_{Z^{Obs}-Z^{Down}}} \ln\left[\frac{(1+\rho^{Obs})(1-\rho_{high}^{Down})}{(1+\rho_{hign}^{Down})(1-\rho^{Obs})}\right] = -CI$$

$$or, \frac{(1+\rho^{Obs})(1-\rho_{high}^{Down})}{(1+\rho_{high}^{Down})(1-\rho^{Obs})} = \exp(-CI \times 2 \times \sigma_{Z^{Obs}-Z^{Down}}) = C \text{ [Where, } C = \exp(-CI \times \tfrac{2\times\sqrt{2}}{\sqrt{n-3}}]$$

$$or, \frac{(1-\rho_{high}^{Down})}{(1+\rho_{high}^{Down})} = C \times \frac{(1-\rho^{Obs})}{(1+\rho^{Obs})}$$

$$or, \frac{1}{2} + \frac{\rho_{high}^{DOwn}}{2} = \frac{1+\rho^{Obe}}{(1+C)+\rho^{Obe}(1-C)}$$

$$or, \rho_{high}^{Down} = \frac{1-C+\rho^{Obs}(1+C)}{1+C+\rho^{Obs}(1-C)}$$

Similarly, back-calculating for positive bound of CI gives us ρ_{low}^{Down}. Which is

$$\rho_{low}^{Down} = \frac{C - 1 + \rho^{Obs}(C + 1)}{1 + C + \rho^{Obs}(C - 1)}$$

Table 3.4: Synthetic simulation result for Confidence Interval (CI) values of downscaled GCM cross-correlation given observed cross-correlation and three fractional bias (fb) values- 0.05, 0.10 and 0.20.

	Historical			Hindcast		
Observed Corr	**fb=0.20**	**fb=0.10**	**fb=0.05**	**fb0.20**	**fb0.10**	**fb0.05**
(+/-)0.1	0.08	0.03	0.02	0.06	0.03	0.01
(+/-)0.2	0.16	0.08	0.04	0.12	0.06	0.03
(+/-)0.3	0.26	0.13	0.06	0.2	0.1	0.05
(+/-)0.4	0.37	0.18	0.09	0.28	0.14	0.07
(+/-)0.5	0.51	0.26	0.13	0.4	0.2	0.1
(+/-)0.6	0.7	0.37	0.18	0.57	0.29	0.14
(+/-)0.7	0.91	0.54	0.27	0.81	0.43	0.21
(+/-)0.8	1	0.82	0.45	1	0.69	0.35

3.7 References

"Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections" archive at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/.

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res., **111**, D05109. doi:10.1029/2005JD006290.

Bhaskaran, B., and J. Mitchell, 1998: Simulated changes in southeast Asian monsoon precipitation resulting from anthropogenic emissions. Int. J. Climatol., **15**, 1455–1462.

Blunden, J., and D. S. Arndt, eds. 2014: State of the climate in 2013. B. Am. Meteorol, Soc., **95(7)**:S1-S257.

Brunetti, M., M. Maugeri, F. Monti and T. Nanni, 2004: Changes in daily precipitation frequency and distribution in Italy over the last 120 years. J Geophys Res, **109**:D05102, doi:10.1029/2003JD004296

Dufresne, J., Foujols, M., Denvil, S. et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5 Clim Dyn (2013) 40: 2123. doi:10.1007/s00382-012-1636-1

Fisher, R. A (1950), Statistical Methods for Research Workers, 11th ed. London: Oliver & Boyd

Fisher, R. A., 1925: Statistical Methods for Research Workers. Edinburgh: Oliver & Boyd. Text accessible at http://psychclassics.yorku.ca/Fisher/Methods/.

Fowler, H. J., S. Blenkinsop, C. Tebaldi, 2007: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int J Climatol., **27(12)**, 1547–1578.

Gangopadhyay, S., M. Clark, K.Werner, D. Brandon, and B. Rajagopalan, 2004: Effects of spatial and temporal aggregation on the accuracy of statistically downscaled precipitation estimates in the Upper Colorado River Basin. Journal of Hydrometeorology, **5(6)**, 1192-1206.

Giorgetta, M. A., et al. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, J. Adv. Model. Earth Syst., 5, doi:10.1002/jame.20038,2013

Goswami, B. N., V. Venugopal, D. Sengupta, M. S. Madhusoodanan and P. K. Xavier 2006: Increasing trend of extreme rain events over India in a warming environment. Science **314**:1442–1445.

He, X., Y. Yang, and J. Zhang, 2012: Bivariate Downscaling with Asynchronous Measurements. Journal of Agricultural, Biological, and Environmental Statistics. **17(3)**, 476–489.

Huth, R., 1999: Statistical downscaling in Central Europe: evaluation of methods and potential predictors. Clim Res., **13(2)**, 91–101.

IPCC (Intergovernmental Panel on Climate Change), 2013: Climate change 2013: The physicalscience basis. Working Group I contribution to the IPCC Fifth Assessment Report. Cambridge University Press. www.ipcc.ch/report/ar5/wg1.

Karl T. R. and Koss W. J., 1984: Regional and National Monthly, Seasonal and Annual Temperature Weighted by Area, 1895-1983. Historical Climatology Series 4-3, National Climatic Data Center, Asheville, NC, 38 pp.

Karl, T.R., D. R. Easterling, 1999: Climate extremes: Selected review and future research directions. Clim Change **42**:309–325.

Karl, T.R., W. Knight Richard and P. Neil, 1995: Trends in high-frequency climate variability in the twentieth century. Nature **377**:217–220.

Krishnamurthy, C. K. B., U. Lall and H. H. Kwon, 2009: Changing frequency and intensity of rainfall extremes over India from 1951 to 2003. Journal of Climate, **22 (18)**, 4737-4746.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy 2007: Fine-resolution climate projections enhance regional climate change impact studies. Eos Trans. AGU, **88(47**), **504**.

Maurer, E.P., and H.G. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences,* **12**, 551-563.

Mehrotra,R. and Sharma, A. 2016: A Multivariate Quantile-Matching Bias Correction Approach with Auto-and Cross-Dependence across Multiple Time Scales: Implications for Downscaling, Journal of Climate, **29 (10)**,3519-3539.

NOAA (National Oceanic and Atmospheric Administration), 2015: National Centers for Environmental Information. https://www.ncei.noaa.gov/.

Plummer, N., M. Salinger, N. N. James, S. Ramasamy, K.J. Hennessy, R. M. Leighton, T. Blair, C. M. Page and J.M. Lough, 1999: Changes in Climate Extremes Over the Australian Region and New Zealand During the Twentieth Century. Clim Change **42(1):**183–202.

Qin, N., X. Chen, G. Fu, J. Zhai and X. Xue, 2010: Precipitation and temperature trends for the Southwest China: 1960–2007. Geoenvironmental Disaster, DOI: 10.1002/hyp.7792.

Sankarasubramanian, A., and U. Lall, 2003: Flood Quantiles and Changing Climate:Seasonal forecasts and causal relations. Water Resources Research. **39(5),** art.no.1134.

Stoner, A. M. K., K. Hayhoe, X. Yang, and D. J. Wuebbles, 2012: An asynchronous regional regression model for statistical downscaling of daily climate variables. Int. J. Climatol.. **33**, 2473–2494.

Suppiah, R. and K. J. Hennessy, 1998: Trends in total rainfall, heavy rain events and number of dry days in Australia, 1910–1990. Int J Climatol **18(10):**1141–1164.

Taylor, K. E., R. J. Stoufferand, and G. A. Meehl, 2012: An overview of Cmip5 and the experiment design. B. Am. Meteorol, Soc., **93**, 485–98.

Voldoire, A., et al., 2011 : The CNRM-CM5.1 global climate model: description and basic evaluation. Clim. Dyn., DOI:10.1007/s00382-011-1259-y.

Wilby, R. L., C.W. Dawson, and E.M. Barrow, 2002: SDSM—a decision support tool for the assessment of regional climate change impacts. Environ. Model Softw., **17 (2)**, 145–157.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change, **62**, 189– 216.

Zhang, F., and A. P. Georgakakos, 2012: Joint Variavle Spatial Downscaling. Climatic Change, **111(3)**, 945-972.

Chapter 4. Reducing Model Uncertainty in Climate Change Projection using Multi-model Combination

4.1 Introduction

Climate is significantly changing particularly in terms of temperature over the last few decades (Plummer et. al., 1999, Alexander et al., 2006, Qin et. al., 2010, IPCC 2013, Blunden and Arndt 2014, NOAA 2015). The challenge for the Global Circulation Models (GCM) and other earth system models is to project the changes in different statistical properties of precipitation, temperature and other hydro-climate variables. Skillful predictions of future changes provide valuable information to the policy-makers and scientific agencies. GCMs are important as they are based primarily on physical processes, cover larger areas and provide long-term prediction under several types of experiments consisting of different emission scenarios, initialization schemes etc. (Phillips 1956, Meehl et al., 2007, Flato et al., 2013). One GCM differentiates itself from other models by model formation, nature of inputs and can have errors from various sources and uncertainties in climate change prediction. Major simulation errors may arise either from model parameterization or from inadequate-understanding of physical processes (Knutti 2008). Rather than considering one model to simulate climate, many studies have recommended considering multiple models for future projections to reduce the uncertainty (Palmar et al., 2005, Yun et al., 2003, Hagedorn et al., 2005).

One way to reduce the uncertainty in a single model is to pool ensembles from all the models which is known as equal weighting. Equal weighting has its own merit and demerits. Strong correlated errors between models could diminish the expected improvements (Weigel et al. 2010). A recent study (Sanderson et al., 2014) addressed the large degree of interdependency across models and proposes a method for combing model results into single or multivariate distributions. From a number of available models, it is difficult to identify the least skillful model as model strengths and weaknesses are functions of variable of interest, location, season etc. (Hagedorn et al., 2005). Performance based weighting schemes have shown improved average prediction skill in seasonal forecasting (Rajagopalan et al., 2002, Robertson et al., 2003, Colman & Davey 2003, Devineni & Sankarasubramanian 2010 a,b). Statistical estimation of weights conditioned on the dominant prediction conditions (Devineni & Sankarasubramanian 2010) could maximize the skill of multi-model forecasts. Interdependence of model has also been investigated to forecast sea surface temperature and showed improved projection (Khan et al. 2014). Instead of simple static weighting, multi-model combination could be performed by dynamic weighting (Chowdhury & Sharma 2009, Devineni & Sankarasubramanian 2010) where model weights are allowed to change with respect to time.

In this current work, we are interested in reducing uncertainty in GCM projections of change in mean seasonal temperature over the contiguous United States. The variable exhibits substantial spatial variability. Climate change projections from GCM simulation have a high amount of spatial bias compared to observed change in mean seasonal temperature. Model errors in projecting the change have correlated errors with non-zero mean, hence assigning

equal weight to the models may not serve the purpose to reduce the mean square error in combined projections.

Model uncertainty is the dominant source of the near-term uncertainty with a limited role of scenario uncertainty (Hawkins & Sutton 2009). To reduce model uncertainty and model bias under the near-term climate change, we propose an optimal multi-model combination framework. Optimal multi-model combination approaches have shown significant improvements in seasonal forecasting and need to be evaluated in the context of the near-term climate change perspective. Coupled Model Inter-Comparison Project (CMIP5) historical runs, spanning over last half of the 20^{th} century, provides a long term change in mean seasonal temperature. Optimal model combination framework is applied on historical simulations to calculate the model weights which can be applied later to future projections under different emission scenarios to calculate the future changes. CMIP5 also provides decadal hindcast output as a part of core of the near-term experiments (Taylor, 2005). Hindcast runs are initialized with observed Sea-Surface Temperatures (SST) and allowed to run for following 10 or 30 years. We applied optimal and equal weighting multi-model combination framework individually on historical and hindcast runs to evaluate the effectiveness of optimal framework over equal weighting in terms of reducing uncertainty in climate change projections.

Performance of raw GCM projections is first evaluated to capture the observed change in mean of seasonal temperature over 9 NCDC (National Climate Data Center) climate regions over the contiguous US. Then, we propose an optimal weighting framework for multi-model combination to reduce uncertainty in model projections that relies on model diagnostics and performance. Results of the optimal combination framework are compared with results of the

equal weighting scheme. In Section 4.2, we discuss the datasets followed by a model error analysis plan in Section 4.3. Results are presented under section 4.4 and finally, discussions are in Section 4.5.

4.2 Data

We obtained time series of raw monthly average temperature simulations from 10 GCMs that contribute to the 20th Century experiment (also referred as historical runs) of Coupled Model Inter-comparison Project Phase-5 (CMIP 5). We focus our interest over the time period of January, 1950 to December, 1999, 50 years data. Details, like full name of the GCMs, parent organization and ensemble size are provided in Table 4.1. Similar datasets are obtained for hindcast runs of CMIP5. Hindcast experiments which are initialized with observed SST for the beginning year, are the core experiment set up for the near-term projections. Hindcast runs are named for the initialization year (ex: decadal-60, decadal-80) and available as 10 or 30 years of simulation. We considered two hindcast runs, decadal-60 and decadal 80, to construct a continuous time series for the period 1962-1999. Among 10 models that contribute to the 20th Century experiment, only 3 models, bcc-csm 1-1, ccsm-4 and cmcc-cm, have decadal-60 and decadal-80 runs both. Detailed information regarding the Hindcast runs are provided in Table 4.1. Whenever it is needed, raw monthly average temperature dataset is regridded to ($1^0 \times 1^0$) using bilinear interpolation to maintain a uniform resolution across various GCMs.

We extracted raw monthly average temperature from historical and hindcast runs over 9 climate regions of the contiguous United States, commonly known as NCDC (National

Climatic Data Center) climate regions (Northwest, West North Central, East North Central, Central, Northeast, Southeast, South, Southwest and West) that are used to put current climate anomalies into historical perspective (Karl and Koss, 1984). Climate model simulations are compared with observed changes. The gridded monthly average temperature dataset over the contiguous US is obtained from Prof Ed Maurer's website (Maurer et al, 2007; http://www.engr.scu.edu/~emaurer/gridded_obs/index_gridded_obs.html) and regridded using bilinear interpolation to 1°.

Table 4.1: Description of Models. For Hindcast runs, number of ensemble members is shown only for those models that have both decadal-60 and decadal-80 runs.

Modelling Group	Climate Model Id	Historical Experiments	Hindcast Experiments	
			Decadal 60	Decadal 80
		No. of Ensemble Members	No. of Ensemble Members	No. of Ensemble Members
Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology, Australia	ACCESS1-0	**1**	**-**	*****
	ACCESS1-3	**1**	**-**	*****
Beijing Climate Center, China Meteorological Administration	BCC-CSM1-1	**1**	**4**	**4**
	BCC-CSM1-1-M	**1**	**-**	*****
Canadian Centre for Climate Modelling and Analysis	CanESM2	**5**	**-**	*****
National Center for Atmospheric Research	CCSM4	**5**	**-**	*****
Community Earth System Model Contributors	CESM1-BGC	**1**	**-**	*****
	CESM1-CAM5	**3**	**-**	*****
Centro Euro-Mediterraneo per I Cambiamenti Climatici	CMCC-CM	**1**	**3**	**3**
Centre National de Recherches Météorologiques/ Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	CNRM-CM5	**1**	**10**	**10**

4.3 Model Error Analysis Plan

The analysis consists of three parts. First we define our variable of interest to measure the extent to which climate has changed in terms of mean seasonal temperature over the contiguous US for last 50 years. Then, we discuss the spatial bias in estimating the mean seasonal temperature using the equal weighting model combination and provide insights on how spatial bias of model predicted climate change could affect the multi-model combination scheme. Following that, we propose an optimal multi-model combination framework that combines climate change projections from multiple models.

4.3.1 Variable of Interest

Our variable of interest is the change in the long term mean seasonal temperature (in ^{0}C) defined as $\Delta\mu$. From GCM and observed time series of monthly average temperature, we obtained seasonal temperature by taking arithmetic average of monthly values over a season. In the current study, we divided a year into four seasons- January-February-March (JFM), April-May-June (AMJ), July-August-September (JAS) and October-November-December (OND). As these signals are statistical properties of a climate response to external forcing, we obtain them from individual ensemble of predictions. We execute the following steps below to extract $\Delta\mu$ from both hindcast and historical experiments.

1. For grid point *i* over a climate region *c*, given model *m* and given ensemble member *k*, obtain the monthly average temperature for the entire period and arithmetically average them to seasonal temperatures for four seasons.

2. For a particular season, divide the data into two time slices, T1 and T2. Calculate the mean for both periods separately, $\mu_{m,c}^{i,k,T1}$ and $\mu_{m,c}^{i,k,T2}$, and subtract them (earlier is subtracted from the later) to obtain climate change, named $\Delta\mu_{m,c}^{i,k}$.

$$\Delta\mu_{m,c}^{i,k} = \mu_{m,c}^{i,k,T2} - \mu_{m,c}^{i,k,T1} \quad \text{Eq. 1}$$

3. Repeat step 2 for all members ($k=1 \ldots K$) in an ensemble, then average them over ensemble members to obtain climate change signal $\Delta\mu_{m,c}^{i}$ from model m at grid i.

$$\Delta\mu_{m,c}^{i} = \frac{1}{K}\sum_{k=1}^{K} \Delta\mu_{m,c}^{i,k} \quad \text{Eq. 2}$$

4. Repeat steps 1 to 3 for all grids ($i=1 \ldots I$), all models (m=1…10), at each climate region (c=1…9) and for other seasons. Observed change in mean seasonal temperature is denoted as $\Delta\mu_{obs,c}^{i}$ and can be obtained as-

$$\Delta\mu_{obs,c}^{i} = \mu_{obs,c}^{i,T2} - \mu_{obs,c}^{i,T1} \quad \text{Eq. 3}$$

We discarded the use of individual ensemble members as they carry equal probability of occurrence. $\Delta\mu_{obs,c}^{i}$ or $\Delta\mu_{m,c}^{i}$ are spatially distributed and do not have any temporal dimension. For historical runs, T1 and T2 are 1951-1974 and 1975-1999 respectively. During hindcast runs, two periods are defined as 1962-1980 and 1981-1999.

$\Delta\mu_{obs,c}^{i}$ is calculated between two periods 1950-1974 and 1975-1999, over all the climate regions and for four seasons. Observed changes values are plotted in Figure 4.1. During

JFM, West North Central (WNC) and East North Central (ENC) climate regions experienced a higher rise in mean seasonal temperature, up to 2℃, for the later period compared to other climate regions. Apart from ENC and WNC, Northwestern (NW) and Western (W) US experienced a rise in mean seasonal temperature during JFM of up to 0.75-1℃. Some grid points over Central (C) and Southern (S) US have a decrease in mean seasonal temperature up to 1℃. A few grid points over the Southeast (Florida) also experienced an increase in winter season temperature.

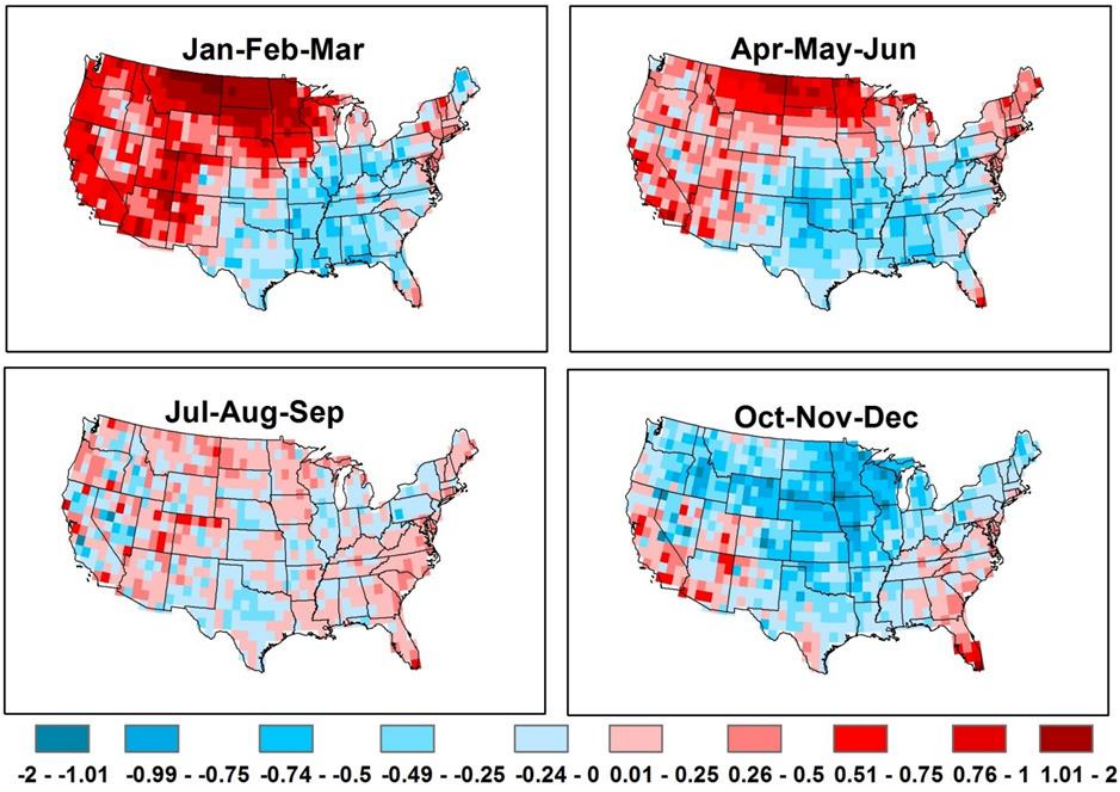

Figure 4.1: Observed changes in mean of seasonal temperature between two periods 1951-1974 and 1975-1999

The amount of increase in mean seasonal temperature was lesser during AMJ compared to JFM. Most grid points over ENC and WNC experienced increase up to 1° C. More grid points, compared to JFM, over the Central and Southern US experienced decrease in mean seasonal temperature during the later period. Over the Northeast (NE), most of the grid points exhibits increase in mean seasonal temperature. Increase in mean seasonal temperature is lower during JAS compared to AMJ or JFM over the Northern US. Northern US experienced increase up to 0.5° C. Grid points covering the Southeastern US experienced decrease ($\pm$ 0.25) in mean seasonal temperature during JAS. Overall the changes during JAS over the contiguous US are lower in magnitude and not very spatially coherent. Finally during OND, decreases in mean seasonal temperature are prominent over the Central, East North Central and West North Central US with decreases as large as 2℃. Whereas, Southeast (SE) and Western US experienced increases in mean seasonal temperature between 1℃ to 2℃. The changes are prominent and should be simulated GCM model projections. Results shown in Figure 4.1 are in line with the earlier studies that had addressed changing patterns of trends and frequencies of temperature over the contiguous US. The multimodel combination approach should be framed in such a way that captures the observed magnitude and trends in mean seasonal temperature.

The accurate simulation of observed changes by GCMs provide confidence in their reliability to capture future changes. Hence, the following analysis helps to understand GCM behavior and any particular GCM's superiority over the others to simulate the observed change. Each panel in Figure 4.2 shows change in mean seasonal temperature projected by 10 GCMs and their corresponding observed change over a climate region as boxplots for the

season JFM. At each panel, we have 10 boxplots (in blue) created from 10 GCMs and the last boxplot (in green) represents the observed changes. Quartiles of each boxplot represents spatial variation of GCM response. During JFM, ENC and WNC are two crucial regions, we can locate the observed medians close to 1°C. While, other than cmcc-cm, none of the 10 GCMs has the median of climate change close to that of observed. Over WNC, the spatial variation exhibited by the GCMs (except cmcc-cm) is lower compare to the observed. On the other hand, over the South, Southeast and Central, the median of observed change is negative, i.e., mean of seasonal temperature is decreasing during JFM over these three climate regions. Most of the GCMs not only fails to estimate the spatial variation of climate change projections over the three climate regions but also project a positive change. In brief, Figure 4.2 reveals the necessity of an effective multi-model combination approach to reduce uncertainty in multiple model climate change projections.

4.3.2 Bias Analysis

GCMs have spatial bias in errors which can be as high in magnitude as the observed spatial mean of climate change projections. Earlier studies neglected the bias in error and only considered the variance of errors (Weigel et al., 2010). This scenario completely changes when bias in the error is present and it has potential to lead to failure of the equal weighting scheme. Equal weighting could not remove the bias in errors because the bias remains even with limits taken over the average of infinite number of models. Hence we argue that any multi-model uncertainty reduction scheme should explicitly consider reduction in bias.

To calculate the spatial bias in GCM models projections over the contiguous United States, we consider $\Delta\mu^{i}_{obs,c}$ and $\Delta\mu^{i}_{m,c}$ to be the observed and GCM projected climate change on grid point *i*, model *m* and over the region *c* for a particular season. First, spatial bias is calculated for individual models by taking spatial average of the climate change projections, named as $e_{m,c}$.

$$e_{m,c} = \frac{1}{I}\left[\sum_{i=1}^{I} \Delta\mu^{i}_{obs,c} - \sum_{i=1}^{I} \Delta\mu^{i}_{m,c}\right] \qquad \text{Eq. 4}$$

Spatial bias that remains after equal weighting, is defined as ξ_c . ξ_c for a particular season, is also considered as the average error, can be calculated using Eqn 5.

$$\xi_c = \frac{1}{I}\sum_{i=1}^{I}\left(\Delta\mu^{i}_{obs,c} - \frac{1}{M}\sum_{m=1}^{M}\Delta\mu^{i}_{m,c}\right) \qquad \text{Eq. 5}$$

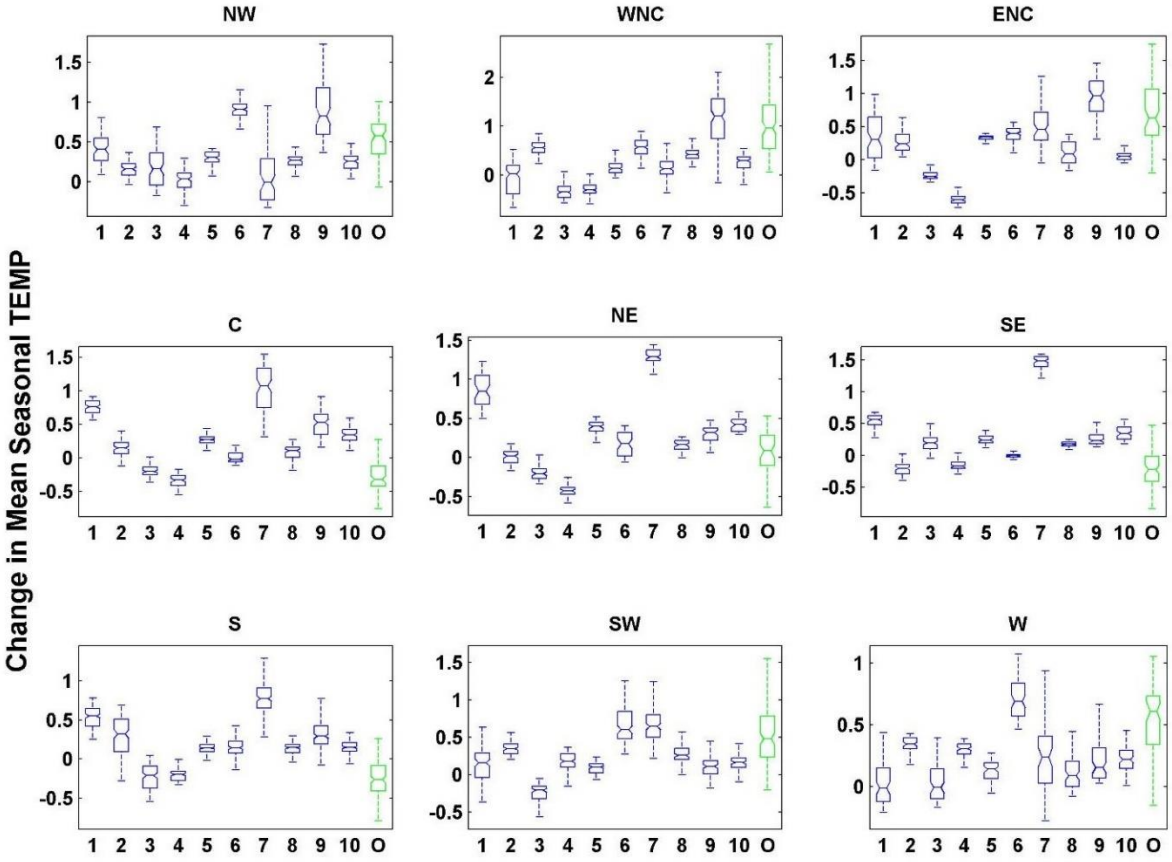

Figure 4.2: GCM performances over NCDC regions to capture observed (green) changes in mean seasonal temperature for Jan-Feb-Mar. X-axis represents 10 GCM models and the observed (Green); Y-axis represents change in mean of seasonal temperature.

The term $\frac{1}{M}\sum_{m=1}^{M}\Delta\mu_{m,c}^{i}$ in Eqn 5 represents multimodel combined projection from equal weighting scheme at grid point *i* over climate region *c*. We defined a metric as bias-ratio (f_c) that calculates the ratio between average error from equally combined model projections and spatial average of observed changes. The ratio gives us the idea about the proportion of error from multimodel combination compared to the observed change, hence equal weighting scheme's ability to eliminate the spatial bias. The metric,f_c, can be calculated using Eqn 6.

$$f_c = \frac{\xi_c}{\frac{1}{I}\sum_{i=1}^{I} \Delta\mu_{obs,c}^{i}} \quad \text{Eq. 6}$$

Bias-ratio, f_c is calculated over the climate regions and for 4 seasons, results from historical runs are shown in Table 4.2. As we already mentioned, ENC and WNC are the two climate regions that experienced higher magnitude of climate change during JFM. Over these two climate regions bias to original change ratio is around 0.8. Over Northwest and West, amount of spatial bias from equal weighting are 0.53 and 0.64 of the observed climate change respectively. During AMJ when observed changes are lower compared to JFM, bias-ratio decreases. Equal weighting improves performance with bias-ratio equal to 0.60 and 0.59 over WNC and ENC respectively. Equal weighting performs better over Southwestern US, with bias-ratio close to zero which signifies correlated error in GCM climate change projections are negligible over SW during AMJ. During JAS, the bias-ratio exhibits negative values for 8 out of 9 NCDC climate regions. Climate models are overestimating the change in mean of seasonal temperature. Finally, during OND the spatial bias after equal weighting remains even higher than the observed changes. Over Southeast and Central, where mean seasonal temperature is decreasing, spatial bias in projections are 1.5 times bigger than the observed climate change. Over Western US, the bias-ratio is 2.58 which denotes a very high amount of bias that could not be eliminated by equal weighting alone, unless a proper bias correction is applied along with multimodel combination of climate change projections. We used the term- $e_{m,c}$, calculated using Eqn 4 to eliminate the spatial bias of individual model projections.

Table 4.2: Bias-ratio (historical runs) after equal weighting is applied for multi-model combination

Season	NW	WNC	ENC	C	NE	SE	S	SW	W
JFM	0.53	0.82	0.80	0.82	7.85	0.81	0.79	0.57	0.64
AMJ	-0.23	0.60	0.59	0.61	-1.01	0.60	0.53	-0.02	-0.42
JAS	-0.75	-1.58	-1.55	-1.75	5.73	-1.95	-1.12	-0.75	-0.83
OND	1.82	1.97	1.46	1.50	1.78	1.52	2.26	1.92	2.58

4.3.3 Optimal model combination

1. We propose here an optimal weighting scheme, to reduce the uncertainty in the estimation of $\Delta\mu^{i}_{obs,c}$ by optimally combining all the models within a climate region. Seasonal climate change response, $\Delta\mu^{i}_{m,c}$ & $\Delta\mu^{i}_{obs,c}$, for a particular grid point is extracted by steps discussed earlier in 4.3.1. $\Delta\mu^{i}_{m,c}$ is the set of multiple model predictions [$\Delta\mu^{i}_{1,c}$ ... $\Delta\mu^{i}_{M,c}$] and $\Delta\mu^{i}_{obs,c}$ is the observed change in mean seasonal temperature for a particular season, over climate region *c*. First the spatial bias is eliminated from model projections by adding $e_{m,c}$ which is calculated using Eqn 4, to model projections. Bias corrected model projections are denoted as $\Delta\mu'^{i}_{m,c}$.

$$\Delta\mu'^{i}_{m,c} = \Delta\mu^{i}_{m,c} + e_{m,c} \qquad \text{Eq. 7}$$

2. Model error ($E^{i}_{m,c}$) at each grid point is calculated over all the grid points within a climate region and all models participating

$$E^{i}_{m,c} = \Delta\mu'^{i}_{m,c} - \Delta\mu^{i}_{obs,c} \qquad \text{Eq. 8}$$

Model error matrix is constructed using model error values from all models and over all grid points within a climate region-

$$E_c = \begin{pmatrix} e_{1,c}^1 & \cdots & e_{1,c}^I \\ \vdots & \ddots & \vdots \\ e_{M,c}^1 & \cdots & e_{M,c}^I \end{pmatrix} \quad \text{Eq. 9}$$

Dimension of E_c is ($M \times I$), where I is the total number of grid points within climate region c and M is the total number of models.

3. For M=10 (for Historical runs) models calculate the variance covariance matrix $\Sigma_{EE,c}$where diagonals represent error variances exhibited by models over I grid points. Since we have 10 models, dimension of $\Sigma_{EE,c}$ is (10×10). Variance covariance matrix $\Sigma_{EE,c}$ is used further to calculate weights for model combination.

$$\Sigma_{EE,c} = cov(E_c) \quad \text{Eq. 10}$$

4. Over a climate region c, we assume weights associated for multi-model combination are $\boldsymbol{w_c} = [w_{1,c} \ldots w_{M,c}]$. Multi-model prediction combining weights and model responses are given as $\Delta\mu^i_{multi,c}$.

$$\Delta\mu^i_{multi,c} = \sum_{m=1}^{M} w_{m,c}\, \Delta\mu'^i_{m,c} \quad \text{Eq. 11}$$

5. To obtain a lower mean square error, our objective is to minimize the determinant of error co-variance matrix. Application of $\Sigma_{EE,c}$ as the estimation of error variance co-variance matrix is sensitive because if pair wise correlation between models are high, optimization could be difficult (Cleman & Winkler 1986), hence a large dataset is preferable to estimate the error variance co-variance matrix. Unbiased combination prediction is achieved by conditioning the sum of all weights to unity. Vector of model weights over a climate region is defined as $\boldsymbol{w_c} = [w_{1,c} \ldots w_{M,c}]$, and the vector $\boldsymbol{w_c^i}$ is constant over grid points within a region.

Objective function: min $(\boldsymbol{w_c'}\Sigma_{EE,c}\boldsymbol{w_c})$ s.t., sum$(w_{M,c})$=1 Eq. 12

Optimization was performed by Genetic Algorithm function with default features in MATLAB 2012a. Steps 1 to 3 repeated and optimization performed for other grids points. A step by step process for data extraction, bias correction and optimal weighting framework is provided under Figure 4.3.

Comparison between two approaches, equal and optimal, is performed based on two statistical parameters- root mean square error (RMSE) and correlation. RMSE is calculated using the following equation-

$$RMSE_c = \sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(\Delta\mu^i_{multi,c} - \Delta\mu^i_{obs,c}\right)^2} \quad \text{Eq. 13}$$

Correlation is also calculated between multimodel combined projections and observed changes.

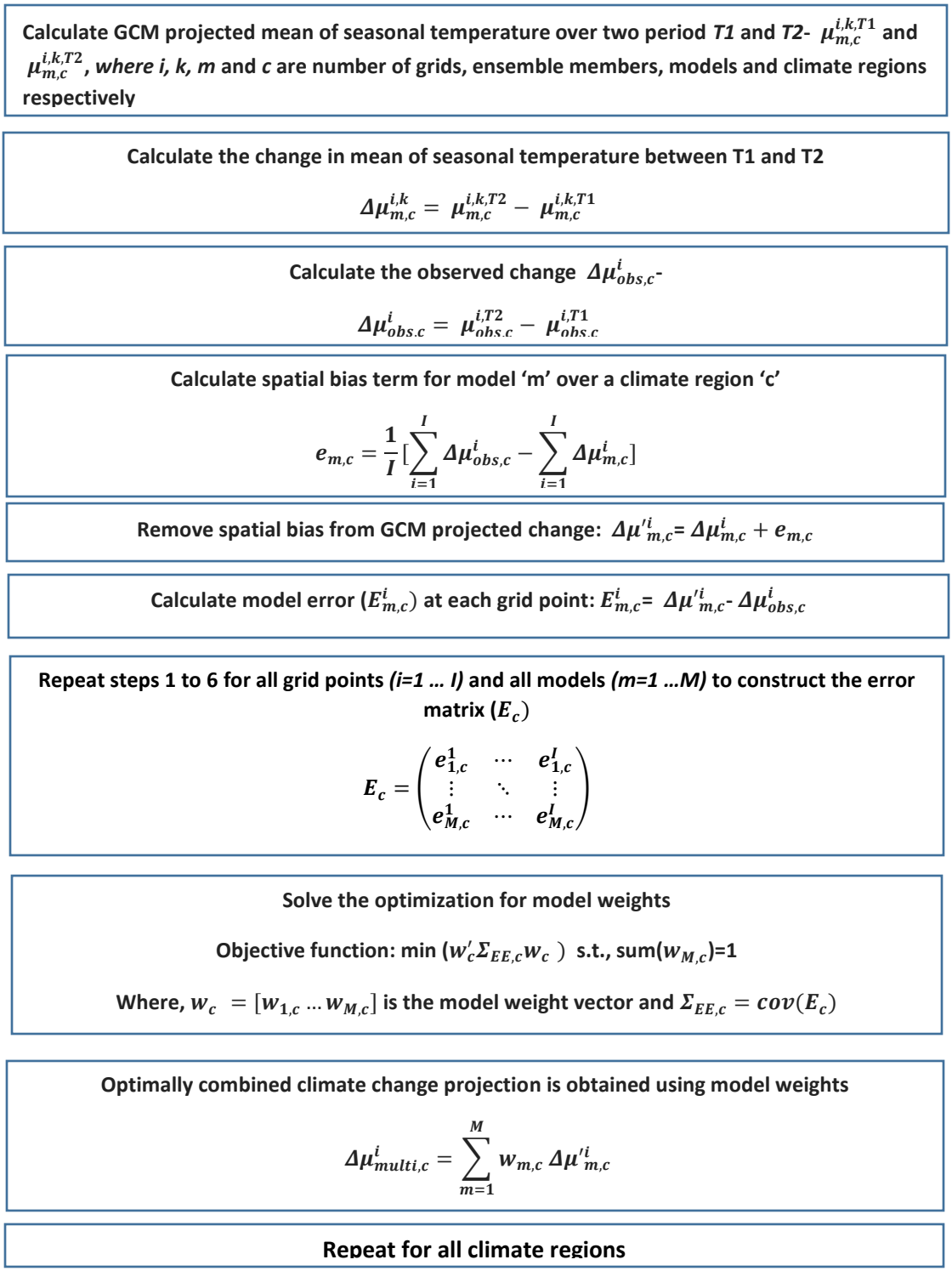

Figure 4.3: Schematic diagram showing steps of Optimal Weighting Framework for multi-model Combination

4.4 Results

Two model combination techniques, optimal and equal, are applied on model predicted changes of mean seasonal temperature from the historical and the hindcast runs separately. Model combination frameworks are individually set up for each NCDC climate regions and for each season. Multimodel combined change, given a season and a climate region, is analyzed based on the observed changes. Root mean square error (RMSE) and correlation between the multimodel combined changes and the observed changes are calculated. Model weights are obtained for a given region and season. RMSE is considered as a measure of accuracy, providing the standard deviation in difference between observed and multi-model combined changes. An ideal multi-model combined change would expect RMSE values close to zero. RMSE shares the same unit as change in mean seasonal temperature. On the other hand, correlation between the observed and the reveals whether multimodel combined changes are able to detect the observed changes. Multi-model combined changes could have a lower RMSE, as close to zero but it should also capture the future changes. Models weights from optimal approach, unlike equal weighting, do not depend on the number of models. Rather weights from the optimal combination portray the ability of one or more than one models to capture the observed changes. Hence, we examine the model weights generated over a region to understand how much the optimal weighting differs from the equal weighting.

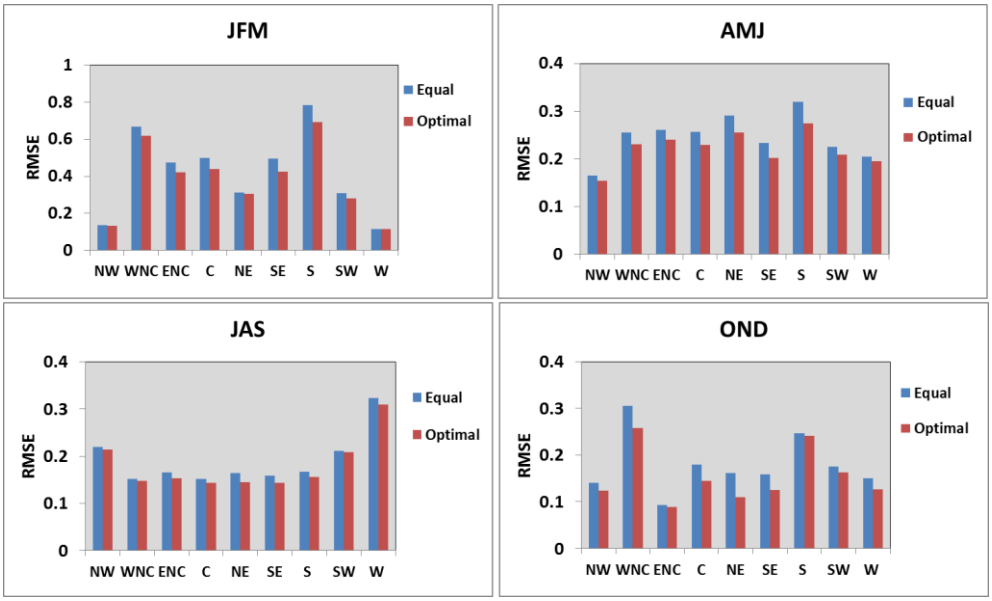

Figure 4.4: Performance of Optimal and Equal weighting schemes for Historical runs

Historical runs:

Figure 4.4 shows RMSE values of the optimal and the equal weighting schemes obtained for different seasons and for 9 NCDC climate regions. Duirng JFM, the Northwest, West North Central and East North Central have experienced higher values of observed changes in mean seasonal temperature, up to 2.0℃, compared to other climate regions. Both approaches perform better over the Northwest with RMSE values around 0.13, while, over WNC the RMSE for the equal weighting is 0.68, whereas optimal weighting performances better with a RMSE value of 0.61. The difference in RMSE between the optimal and the equal model combination over WNC and ENC remains same, optimal model combination exhibits RMSE around 0.42. During JFM, the Southern regions experience relatively weak performance from both model combination schemes. RMSE for the optimal and the equal weighting schemes are 0.69 and 0.78 respectively. Optimal and equal weighting schemes

perform best during JFM with similar efficiency in terms of RMSE over the Western US. RMSEs for both schemes over the Western US are 0.11. From Figure 4.1, the Northeast and the Southwest experienced 1-2^{o} C increase in the observed changes in mean seasonal temperatures values. RMSE values for optimal weighting over NE and SW, 0.30 and 0.27 respectively, are better than RMSE values from the equal weighting for which RMSE values are 0.31 and 0.30 respectively for NE and SW. Next to JFM, AMJ experienced the second largest temperature increase after JFM. During AMJ, ENC and WNC experienced changes up to 1-1.5℃ increase in mean seasonal temperature. Part of SE also experienced temperature increase whereas, the Central US have decrease in mean seasonal temperature. Optimal model combination performs slightly better compared to the equal weighting model combination with RMSE being 0.23 and 0.24 respectively for WNC and ENC. RMSE values for the equal model weighting schemes are 0.26 for both regions, WNC and ENC. Over SE, the equal and the optimal weighting schemes exhibit RMSEs as 0.23 and 0.20 respectively. In the Central US, optimal model combination continues performing better than the equal weighing with RMSE of 0.23, while equal has a RMSE of 0.25. Similar to JFM, during AMJ both schemes perform poorly over the Southern US. RMSE values for the Southern US are 0.31 and 0.27 for the equal and the optimal model combination respectively, while, over NW, both approaches perform better compared to other regions. RMSE values for NW are 0.16 (equal) and 0.15 (optimal).

Observed changes in mean seasonal temperatures for JAS are neither as high as JFM or AMJ nor the grid points with higher values are clustered around the northern states. Change values are randomly scatted and varies between -0.25- 0.25℃. RMSE values for JAS are also lower than JFM or AMJ. From Figure 4.4, optimal weighting scheme performs moderately

steady for climate regions WNC and SW with RMSE values around 0.14. Though, equal weighting scheme performance also follows a steady pattern over WNC and SW regions, whereas RMSEs for equal weighting scheme are slightly higher compared to optimal model combination and have values around 0.15. Western US has little or no change in observed mean seasonal temperature and RMSEs from model combination approaches end up being higher compared to the other regions. RMSE values over the Western US for the equal and the optimal model combination are 0.32 and 0.30 respectively.

During OND, Southeast, West and Southwest are the climate regions where mean seasonal temperature are rising by up to 2°C for the period 1975-1999. On the other hand, mean seasonal temperature is decreasing for the period 1975-1999 by 1-2°C over the West North Central and the Central US. Figure 4.4 shows, over SE, W and SW RMSE values resulted from optimal model combination approaches are 0.12 (SE and W, both) and 0.16 (SW). Optimal model combination performs better than equal model combination which has RMSE values as 0.15, 0.14 and 0.17 respectively, for SE, W and SW. RMSE for both approaches are higher for WNC compare to the other regions in this season. RMSE values over WNC are 0.30 (equal) and 0.25 (optimal) whereas for the Central US, RMSE values are 0.18 and 0.14 respectively for the equal and the optimal model combination approaches.

Correlation between the observed and the model combined projection for the change in mean seasonal temperature are calculated for four seasons and over 9 NCDC climate regions. Results of correlations for the historical runs are plotted in Figure 4.5. Out of three climate regions, NW, WNC and ENC, where changes in mean seasonal temperature experienced a rise up to

2℃- over WNC equal weighting approach fails to detect the observed changes. Over WNC, correlation for optimal weighting has a lower value of 0.39, compared to ENC and NW but the combined projection simulates an increase in temperature. Optimal performance improves for ENC and NW with correlation values being 0.59 and 0.72 respectively. Over the NE and SE regions which also experienced rise in temperature with an order between 1-2℃, both optimal and equal approaches are able to detect the observed change. Correlation values for optimal weighting are 0.39 and 0.61 respectively for NE and SW. On the other hand, correlation for the equal weighting on these two regions are weaker than the optimal with correlations values being 0.18 and 0.30 respectively for NE and SW. Over the South, RMSE is high and also both approaches fail to detect the observed change. Similar to JFM, during AMJ correlation between model combined projections using equal weighting and observed changes are negative over WNC and the Southern US; additionally equal weighting fails to detect the trend over the west. Optimal weighting has relatively higher and positive correlation with the observed, compared to equal weighting over climate regions, except the Western US. Observed changes are of higher values over ENC and WNC and the optimal weighting projects positive correlations with correlation values around 0.30 over the two regions. Optimal weighting scheme exhibits a strong performance over NE where the correlation with the observed is 0.70. During JAS, over four climate regions, ENC, C, SE and South, both approaches fail to detect the observed change. However the negative correlation for the optimal combination scheme is not as strong as the correlation for equal weighting scheme. Optimal weighting performs strongly over NE with the correlations close to 0.80. Observed changes during JAS have lower magnitude and have moderate values of temperature changes over the climate regions. During OND, out of

three climate regions that have experienced significant increase in mean seasonal temperature, equal weighting scheme exhibits negative correlation for all three regions. On the other hand, optimal weighting projects all positive correlations for SE, W and SW with values 0.49, 0.68 and 0.24 respectively. Only the Southern US experience negative correlation between the observed and the multimodel combined projection by optimal weighting. Between two states, ENC and WNC that have experienced decrease in the mean seasonal temperature, equal weighting fails to detect the trend in WNC. Optimal combination scheme have correlation values as 0.25 and 0.34 respectively for WNC and ENC while equal weighting exhibits correlation values as -0.67 and 0.16 respectively for WNC and ENC.

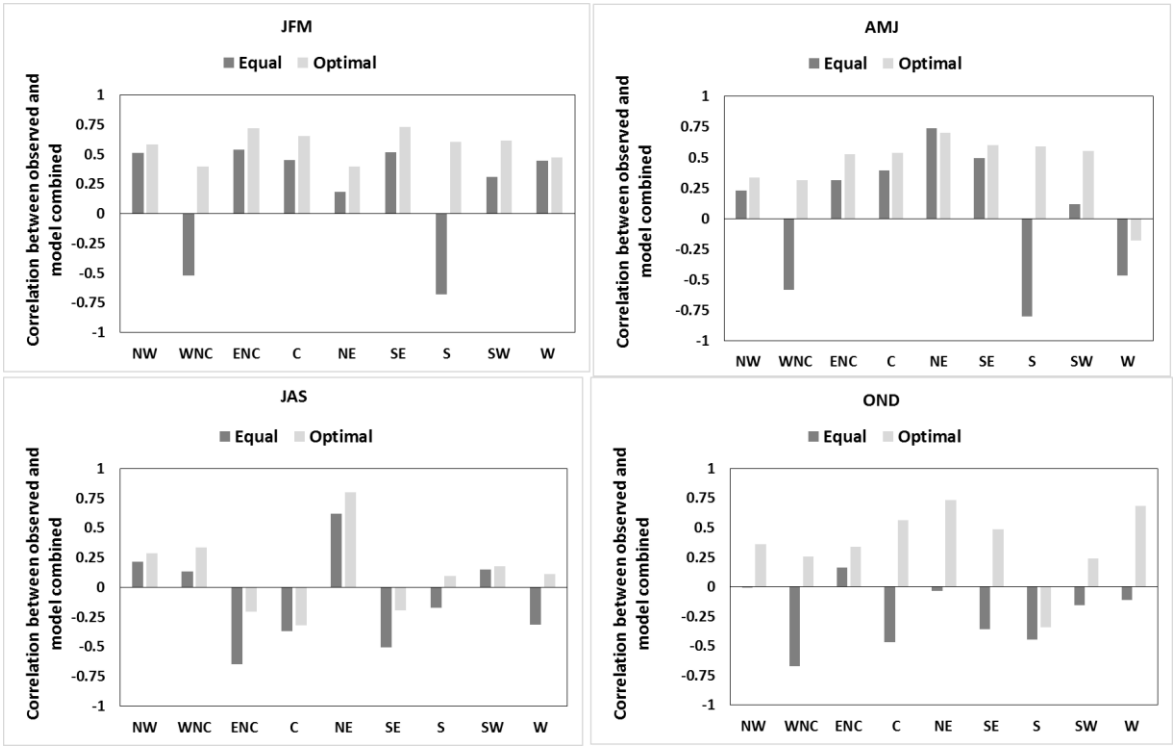

Figure 4.5: Correlations between observed change and multi-model combined projections of climate change for historical runs.

Over each climate region, we calculated model weights for 10 GCMs that would help to reduce the uncertainty by combining GCM model projections of change in mean seasonal temperature. For optimal weighting approach model weights can vary from 0 to 1 however model weights are fixed for equal weighting scheme. For historical runs, a model weight of 0.1 is imposed on each GCM during the equal weighting. We would check the model weights to understand whether model weights generated from optimal weighting framework are different in magnitude from model weights provided by the equal weighting scheme. For a particular climate region, given a season, model weights representing 10 models are calculated and plotted in Figure 4.6. In Figure 4.6, each panel represent a season, climate regions are plotted in x-axis and model weights over regions are plotted in y-axis.

For NW and JFM season, the highest weight value equal of 0.15 is assigned to access 1.0 whereas two models, cesm1-bgc and cmcc-cm, have model weights equal to 0.12. For WNC, access 1-3 and cmcc-cm have model weights of values 0.19 and 0.17 respectively. Apparently access 1-3 and cmcc-cm perform better in general over all regions, including NE and SE. During AMJ, access 1-3 and bcc-csm 1-1 performs better over WNC with model weights equal to 0.19 and 0.16 respectively. Over ENC also, bcc-csm 1-1 has a higher model weight of 0.16 while other models have weights around 0.1, whereas, during AMJ access 1.0 has little or no significance in terms of optimal model combination and model weights are close to zero. However, this scenario changes during JAS when model access 1.0 has dominance over the other models over regions like ENC, SE and South with model weight values as 0.18, 0.35 and 0.19 respectively. Western US has higher increase of mean seasonal temperature change during JAS. Models cmcc-cm and ccsm4 perform slightly better over the Western US

with model weight values around 0.19 and 0.14 respectively. During OND, model weights for model access 1-3 are 0.14 (SW) and 0.19(W). Over SW and W, cnrm-cm5 has model weights close to 0.16 and 0.14 respectively and cmcc-cm has model weights close to 0.19 and 0.18 respectively. Changes over ENC are well captured by GCM bcc-csm 1-1 and cesm-bgc with models weights as 0.14 and 0.17. Best performance among all model and all season, is achieved by access-1.0 over WNC with model weight equal to 0.82.

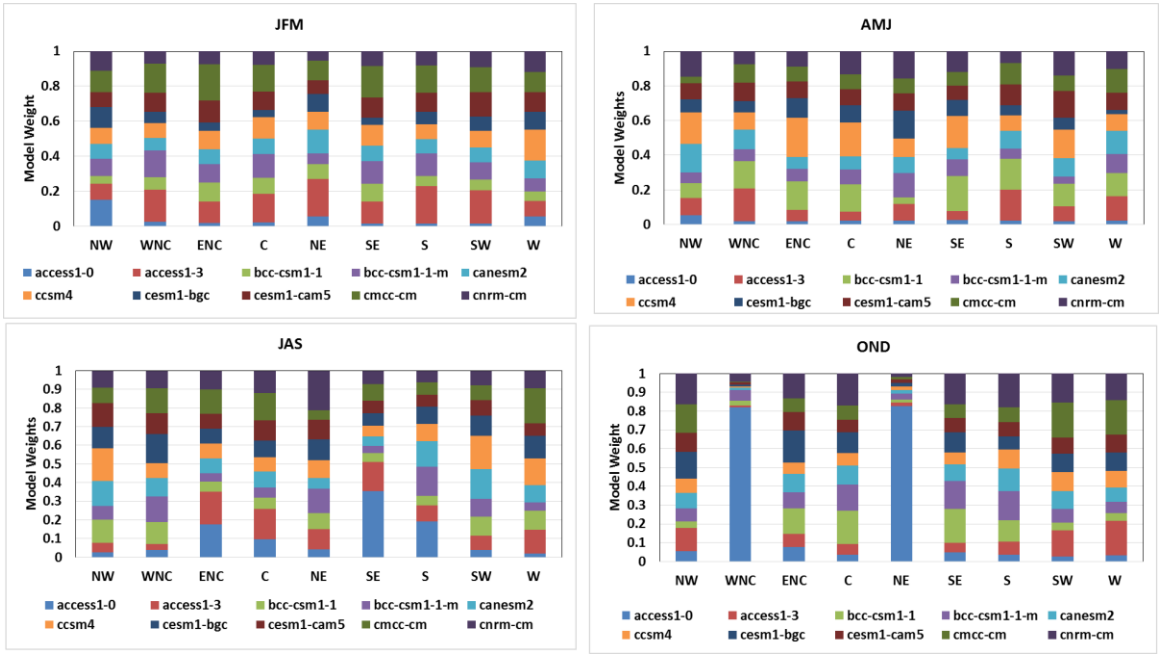

Figure 4.6: Model weights from optimal combination approach for historical runs

Hindcast Runs

Hindcast runs are divided into two time slices, 1962-1980 and 1981-1999, different from the historical runs. During JFM, East North Central and West North Central regions mainly witnessed increase in mean seasonal temperature, up to 2°C, Central and Southwestern

US also experiencde increase in seasonal temperature up to 1° C. Optimal and the equal weighting schemes are applied on climate change projections from the hindcast runs, RMSE values are calculated and plotted in Figure 4.7. While evaluating RMSE and correlations for the hindcast runs, we assumed that the observed changes experienced by the climate region during the period 1961-1999 would be similar to observed changes calculated for the time period 1950-1999.

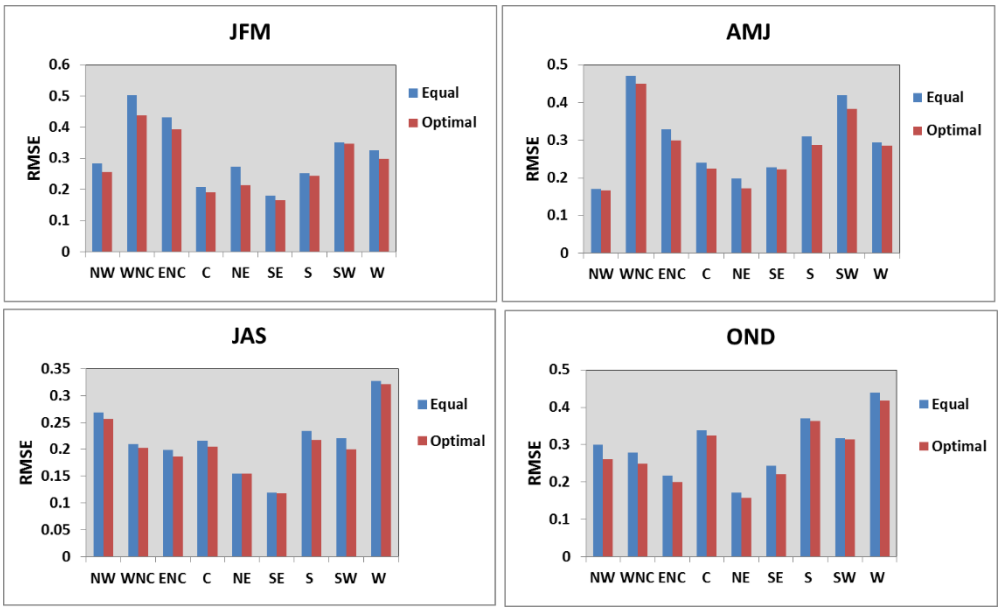

Figure 4.7: Performance of Optimal and Equal weighting schemes for Hindcast runs

RMSE values resulted from the optimal and the equal weighting approaches are higher over ENC and WNC compared to other climate regions. RMSE values for the optimal weighting schemes are 0.43 and 0.39 respectively for WNC and ENC, whereas equal weighting scheme performs weaker than the optimal with RMSE values of 0.50 and 0.43 respectively for WNC and ENC. Over the Central US, both approaches perform equally well with RMSE values of 0.20 and 0.19 respectively for equal and optimal model combination framework.

Over Southwestern US, RMSE values are 0.35 and 0.34 for equal and optimal weighting respectively. Next to JFM, AMJ experienced increase in mean seasonal temperature for the later period of 1981-1999 over the Western US by an amount of 1.2° C and decrease over the Central and Southern US by an amount of 0.8°C. RMSE values from the optimal and the equal weighting approaches are relatively lower over the Western US and do not differ much from each other. RMSE value of the optimal weighting scheme is 0.28, whereas RMSE of equal weighting model combination is 0.29. Optimal performs better than the equal weighting approach over all climate regions during AMJ. RMSE values for the Central US are 0.24 (equal) and 0.22 (optimal). Over the Southern US, RMSE values are 0.30 and 0.28, respectively for equal and optimal weighting model combination approaches. During JAS, West and Southwestern US experienced average 0.6° C rise in mean seasonal temperature and ENC and the Central US experienced average 0.4°C decrease. However, for the Western US, both approaches perform similarly with RMSE values of 0.32, optimal weighting performs slightly better over the Southwest with RMSE value of 0.20. Over the Southwestern US, RMSE value for equal weighting is 0.22. Over the ENC and the Central US, RMSE values for optimal model combinations are better than the RMSE for equal weighting approaches with values 0.18 and 0.20 respectively for the ENC and the Central US. RMSE values for equal weighting approach are 0.19 and 0.21 respectively for the ENC and the Central US. During OND, five climate regions, Central, East North Central, South, West North central and Northwest experienced average 1°C decrease in mean seasonal temperature for the period 1981-1999. Among these five regions, both models perform with similar ability for the East North Central. RMSE values over the ENC are 0.21 and 0.19 respectively for the equal and optimal model

weighting schemes. Both approaches have relatively higher RMSEs over the South with values 0.37 and 0.36 respectively for the equal and optimal model combination schemes. Over the WNC, difference in RMSEs between the equal and the optimal are more than 0.03 with optimal being better in reducing uncertainty in the climate change projections. Overall, optimal model weighting scheme for hindcast runs performs better than equal weighting schemes over 9 climate regions and for all the seasons. Variance in error after model combination is higher when the observed change is higher in magnitude and mean of seasonal temperature is increasing, for example the winter months.

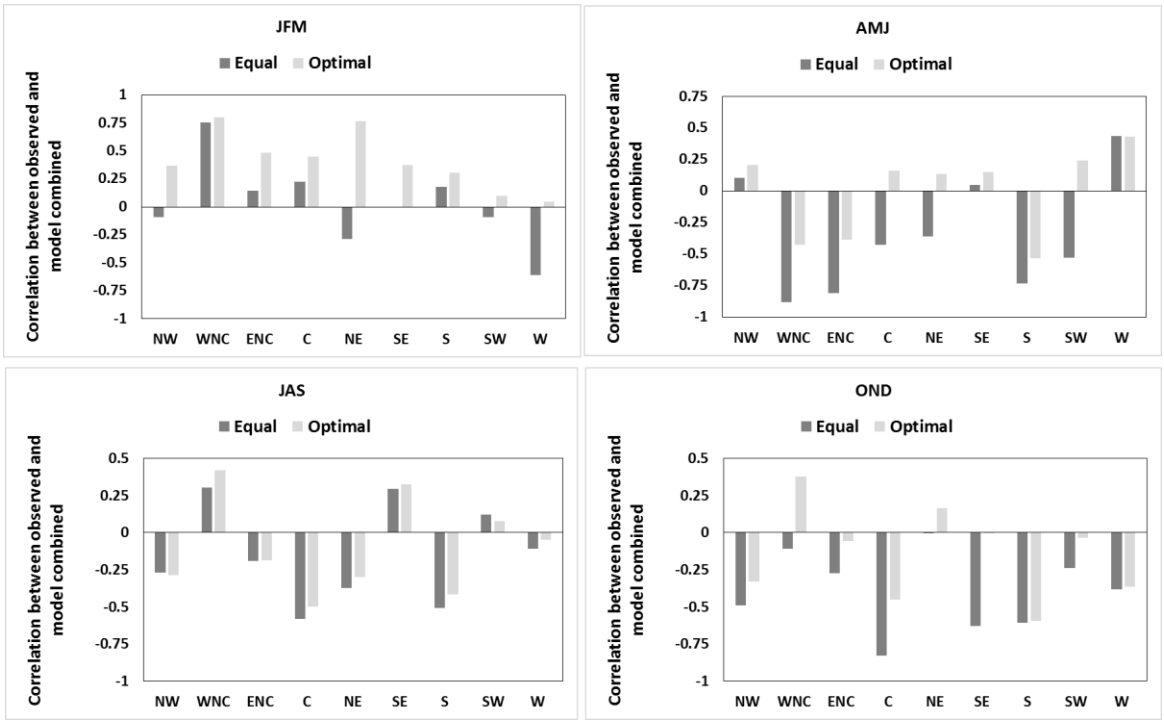

Figure 4.8 Correlations between observed change and multi-model combined projections of climate change for Hindcast runs.

Figure 4.8 shows the correlation between multimodel combined change from hindcast runs and observed changes over 9 climate regions and 4 seasons. During JFM, equally

combined climate change projections for the NE, SW and Western US could not capture the observed changes. Over the WNC, both approaches show strong correlations with the observed changes with values 0.75 and 0.79 respectively for the equal and the optimal combination approaches. Optimal weighting exhibits stronger correlation over the ENC with value 0.48 compare to correlation value of 0.18 resulted from the equal weighting. Optimally combined change projections exhibits positive correlations with the observed change though almost has no correlation over the Western US. Equal weighting scheme fails to detect the the observed change over 6 out of 9 climate regions during AMJ. Optimal weighting also fails to detect the trend in observed changes but over 3 regions during AMJ. Negative correlation values for the optimal weighting are relatively of lesser magnitude compared to the respective negative correlation value from the equal weighting. Western US is important from the observed change perspective and both approaches perform with similar ability to detect the observed with correlation values of 0.43 for both. Over the Central US, optimal weighting maintains a positive correlation of value 0.16 but fails over the Southern US with correlation value equal to 0.53. Overall, correlations between the observed and the model combined changes are weak during the AMJ season. During JAS, over the Southwest and West, two major regions where temperature has increased, optimal and the equal weighting approaches have very low correlations with the observed and over the Western US both approaches fail to detect the observed change. However, over the SW equal weighting performs slightly better than the optimal model combination with correlation values 0.12 and 0.07 respectively for the equal and the optimal. Over the ENC and the Central US, observed mean seasonal temperature is decreasing for JAS. Both regions have experienced failure to detect the observed change within

by both model combination approaches. Hindcast runs during JAS have scattered grid points with either increasing or decreasing mean seasonal temperature. Hence, superiority of any model combination approach remains inconclusive. Model combination schemes also fails to detect the observed change over five regions for the OND season-C, ENC, S, WNC and NW. However, optimal weighting scheme is able to detect the observed change over the Southeast but the equal weighting scheme fails to detect.

Similar to Figure 4.6, model weights are calculated for 3 GCMs from the hindcast runs and barplots are plotted in Figure 4.9 for 4 seasons. Three models from the hindcast runs are bcc-csm-1-1, ccsm-4 and cmcc-cm are shown under x-axis. Model weights from the equal weighting scheme are constant over each of the three models with value equal to 1/3.

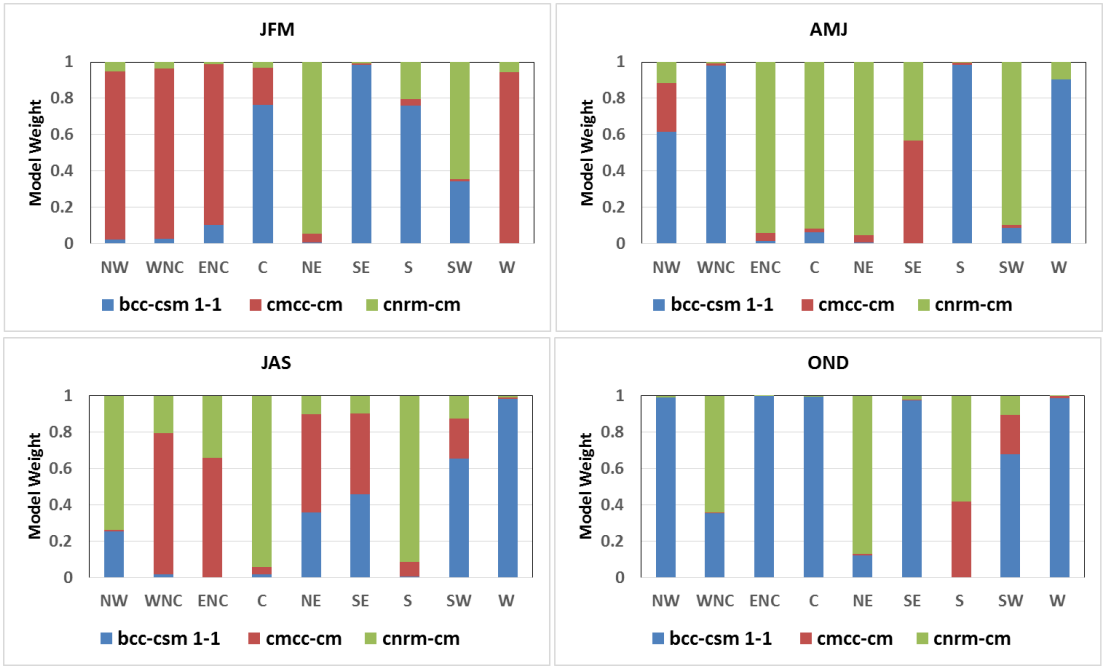

Figure 4.9: Model weights from optimal combination approach for Hindcast runs

During JFM, ccsm-4 has model weights of 0.94 and 0.88 respectively over ENC and WNC. In case of Central US, bcc-csm-1-1 is alone sufficient to reduce the uncertainty in the climate projection with model weight around 0.76. Another important region during JFM is the Southwest where bcc-csm-1-1 and ccsm-4 are dominant with model weight values of 0.34 and 0.64 respectively to address the model uncertainty. During AMJ, Western US experienced higher increase in temperature compared to other climate regions and bcc-csm-1-1 claims the major share of model weights with model weight value as 0.9. Over the Central and the Southern US- other two important climate regions, bcc-csm-1-1 and cmcc-cm have shown their complete dominance on each region respectively. Model bcc-csm-1-1 has the model weights of value 0.98 over the Southern US whereas the same metric over the Central US is close to 0.92 for cmcc-cm. During JAS, over the Western US, bcc-csm-1-1 has model weight close to 1. While, cmcc-cm performs better over the Central and the Southern US with model weight values more than 0.9. During OND, over the Central, ENC and Northwestern US model weights of bcc-csm 1-1 are close to 1. Model importance over the Southern US are shared between ccsm-4 and cmcc-cm with weights close to 0.4 and 0.6 respectively, whereas over the WNC the share is between bcc-csm-1-1 and cmcc-cm with weights from the optimal weighting being 0.35 and 0.65.

To summarize the performance of the optimal and the equal weighting multimodel combination approach, we calculated the fraction of total grid points over a climate region on which squared error from the optional weighting is lesser than the squared error from the equal weighting. The analysis is a one to one comparison between two approaches in terms of

reducing uncertainty in the multimodel projection. Results are shown under Table 4.3 (historical) and Table 4.4 (hindcast).

For the historical run, during JFM over the WNC and the ENC climate regions where the observed change is higher compared to the other regions, optimal framework performs better than the equal over more than 80% grid points. NW, NE and SW where observed changes are moderate, equal performs slightly better with the metric values as 0.55, 0.63 and 0.50 respectively. During AMJ, 55% grid points over the WNC and 78% grid points over the ENC exhibits superior performance by the optimal scheme. During JAS, where the observed changes have lower values, equal weighting and the optimal weighting perform with the similar ability. When the observed change experienced decrease in temperature over the SE, S and SW, optimal weighting performs better over 60% grid points.

Table 4.3: Percent of grid points where optimal weighting is performing better than equal weighting for Historical runs

Season	NW	WNC	ENC	C	NE	SE	S	SW	W
JFM	0.55	0.80	0.83	0.56	0.63	0.57	0.97	0.50	0.64
AMJ	0.65	0.55	0.78	0.57	0.76	0.64	0.83	0.65	0.66
JAS	0.62	0.55	0.75	0.78	0.69	0.57	0.43	0.58	0.59
OND	0.79	0.69	0.58	0.67	0.65	0.67	0.61	0.62	0.76

Table 4.4: Percent of grid points where optimal weighting is performing better than equal weighting for Hindcast runs

Season	NW	WNC	ENC	C	NE	SE	S	SW	WNC
JFM	0.61	0.77	0.74	0.50	0.76	0.57	0.52	0.59	0.63
AMJ	0.54	0.60	0.89	0.71	0.71	0.61	0.72	0.76	0.54
JAS	0.62	0.50	0.49	0.57	0.57	0.46	0.64	0.55	0.49
OND	0.75	0.70	0.51	0.57	0.65	0.56	0.65	0.57	0.67

For the hindcast runs, during JFM optimal model combination has lower RMSE compared to the equal weighting approach on 77% and 74% grid points respectively for the WNC and the ENC. Lowest percentage for optimal weighting is over the Central US where equal weighting performs better on 50% of total grid points. During AMJ, over ENC optimal combination performs better over 89% of total grid points. Overall, over 5 climate regions optimal weighting's superiority ranges above 70% of total grid points. During JAS, optimal and equal approach performs with similar ability as the observed changes are of lower magnitude, hence GCMs exhibits a lower spatial bias. Optimal combination is also able to address the decrease in temperature during OND where it performs better than the equal weighting on more than 70% grid points over NW and WNC. Except few climate regions, during JAS for the hindcast and the historical runs both, optimal model combination has dominance over the equal weighting.

4.5 Discussion

In this current study, we proposed an optimal model combination approach to reduce the uncertainty in climate change projection. We focused our attention on the change in mean seasonal temperature. Among four seasons within a year, JFM has experienced increase in mean seasonal temperature during the later period of 1950-1999 over half of the total grid points covering the contiguous US. Northern climate regions (NW, ENC and WNC) during JFM and AMJ experienced increase in mean seasonal temperature up to 2°C. On the other hand, mean seasonal temperature has decreased over the southern climate regions during OND

by 1-2°C. GCMs fail to simulate the spatial mean and variance of the observed change and neither does a bias-corrected GCM time series have the ability to identify the climate change. Hence, multimodel combination is needed to quantify the future changes based on individual model's performance during the hindcast period. Equal weighting is a common model combination approach, and has shown effectiveness over the seasonal forecasting, weather forecast etc. Climate change projection from GCM possess high spatial bias. Spatial bias contributes to model to model corrected error. Equal weighting could not remove the correlated errors, hence spatial bias remains within the multimodel projections. Our study shows that during JFM the amount of spatial bias present in multimodel projection could be equal or higher than the observed change. We proposed an optimal weighting framework, combined with bias correction to reduce the squared error. Optimal weighting framework is applied on the historical and the hindcast runs, over 9 NCDC climate regions of the contiguous US. RMSE and correlations, resulted from the optimal weighting are compared with RMSE and correlation from the equal weighting. For climate regions over which changes are higher, optimal weighting performs better than the equal to keep a lower RMSE. Over many climate regions, equal weighting fails to detect the observed change. Optimal weighting detects the observed change by exhibiting a moderately strong correlation with the observed.

The current study investigates model weights generated by the optimal Weighting, as the weights do not depend on the number of models, unlike equal weighting. For the historical runs, model weights from the optimal weighting are different from that of equal weighting. Usually, 2-3 GCMs have dominance over model combination but not any particular model is dominant over all seasons and all climate regions. During the hindcast runs also, model weights

from the optimal weighting differs from the weights from the equal weighting. Usually a single GCM, out of 3, is sufficient to simulate the observed change. In conclusion, an optimal weighting scheme effectively combines multiple GCM models and reduces RMSE in multimodel projection of climate change better than equal weighting. There is a need to apply the model weights on future projection to quantify the change in mean seasonal temperature for upcoming decades. The analysis in spatial bias of GCM projections reveal that equal weighting fails to eliminate the correlated error between models. A future study to obtain a stable estimate of spatial bias, as a function of time or time independent, is important to project a more appropriate future change. Our study concludes that no particular model is important for the optimal combination for all seasons and over all regions. Hence, ensemble of multiple models performs better within an optimal weighting framework.

4.6 References

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res., **111**, D05109. doi:10.1029/2005JD006290.

Blunden, J., and D. S. Arndt, eds. 2014: State of the climate in 2013. B. Am. Meteorol, Soc., **95(7)**:S1-S257.

Chowdhury, S., and A. Sharma, 2009a: Long-range Niño-3.4 predictions using pairwise dynamic combinations of multiple models. J. Climate, **22**, 793–805, doi:10.1175/2008JCLI2210.1.

Colman, A. W., and M. K. Davey, 2003: Statistical prediction of global sea-surface temperature anomalies. Int. J. Climatol., **23**,1677–1697.

Devineni, N., and A. Sankarasubramanian, 2010a: Improved categorical winter precipitation forecasts through multimodel combinations of coupled GCMs. Geophys. Res. Lett., **37**, L24704, doi:10.1029/2010GL044989.

Flato, G., J. and Co-authors, 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 Cambridge University Press.

Hagedorn, R., F.J. Doblas-Reyes and T. N.Palmer 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. Tellus A. **57**, 219–233. doi:10.1111/j.1600-0870.2005.00103.x.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. Bull Am Meteorol Soc., **90**, 1095–1107.

IPCC (Intergovernmental Panel on Climate Change), 2013: Climate change 2013: The physicalscience basis. Working Group I contribution to the IPCC Fifth Assessment Report. Cambridge University Press. www.ipcc.ch/report/ar5/wg1.

Karl, T. R., and W. J. Koss, 1984: Regional and National Monthly, Seasonal and Annual Temperature Weighted by Area, 1895-1983. Historical Climatology Series 4-3, National Climatic Data Center, Asheville, NC, 38 pp.

Khan, M Z.H., R. Mehotra, A. Sharma and A. Sankarasubramanian, 2014: Global Sea Surface Temperature Forecasts Using an Improved Multi-model Approach, Journal of Climate, **27(10)**, 3505-3515.

Knutti, R., 2008: Should we believe model predictions of future climate change? Philos. Trans. Roy. Soc., **366A**, 4647–4664.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy 2007: Fine-resolution climate projections enhance regional climate change impact studies. Eos Trans. AGU, **88(47)**, 504.

Meehl, G.A. and Co-authors, 2007: Global Climate Projections. In: 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 University Press.

NOAA (National Oceanic and Atmospheric Administration), 2015: National Centers for Environmental Information. https://www.ncei.noaa.gov/.

Palmer, T.N., F.J Doblas-Reyes, R. Hagedorna and A. Weisheimer, 2005: Probabilistic prediction of climate using multi-model ensembles: from basics to applications. Phil. Trans. R. Soc. B. **360**, 1991–1998. doi:10.1098/rstb.2005.1750.

Phillips, N. A. 1956: The general circulation of the atmosphere: A numerical experiment. Q.J.R. Meteorol. Soc., 82: 123–164. doi:10.1002/qj.49708235202

Plummer, N., M. Salinger, N. N. James, S. Ramasamy, K.J. Hennessy, R. M. Leighton, T. Blair, C. M. Page and J.M. Lough, 1999: Changes in Climate Extremes Over the Australian Region and New Zealand During the Twentieth Century. Clim Change **42(1):**183–202.

Qin, N., X. Chen, G. Fu, J. Zhai and X. Xue, 2010: Precipitation and temperature trends for the Southwest China: 1960–2007. Geoenvironmental Disaster, DOI: 10.1002/hyp.7792.

Rajagopalan, B., U. Lall, and S. E. Zebiak, 2002: Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles. Mon. Wea. Rev., **130**, 1792–18

Robertson, A. W., U. Lall, S. E. Zebiak, and L. Goddard, 2004: Improved combination of multiple atmospheric GCM ensembles for seasonal prediction. Monthly Weather Review. 132 2732-2744. DOI: http://dx.doi.org/10.1175/MWR2818.1.

Taylor, K. E., R. J. Stoufferand, and G. A. Meehl, 2012: An overview of Cmip5 and the experiment design. B. Am. Meteorol, Soc., **93**, 485–98.

Tebaldi, C. and R. Knutti, 2007: The use of the multi-model ensemble in probabilistic climate projections. Philosophical Transactions of the Royal Society A, **365**, 2053-2075, doi:10.1098/rsta.2007.2076.

Sanderson, B. M., R. Knutti, and P. Caldwell, 2015: Addressing interdependency in a multimodel ensemble by interpolation of model properties. Journal of Climate, **28**, 5150-5170, doi:10.1175/JCLI-D-14-00361.1.

Weigel, A., R. Knutti, M. Liniger and C. Appenzeller, 2010: Risks of model weighting in multi-model climate projections. Journal of Climate, **23**, 4175-4191, doi:10.1175/2010JCLI3594.1,

Yun, W.T., L. Stefanova and T. N. Krishnamurti 2003: Improvement of the multimodel supersensemble technique for seasonal forecasts. J. Clim. **16**, 3834–3840.

Chapter 5. Summary and Future Problems

Bias Correction and Statistical Downscaling (BCSD) of precipitation and temperature are commonly required to bring the large scale variables available from GCMs to a finer grid-scale for ingesting them into models related to climate-application studies. Most of the currently employed procedures on BCSD primarily consider a univariate approach by developing a statistical relationship between large-scale precipitation/temperature with the local-scale precipitation/temperature ignoring the interdependency between the two variables. In this study, a semi-parametric multivariate approach, Asynchronous Canonical Correlation Analysis (ACCA), is proposed and applied to Global Climate Model simulations from the Coupled Model Inter-comparison Project-phase 5 (CMIP5) to downscale monthly precipitation and temperature over the continental US. This multivariate downscaling framework was first applied to the CNRM-CM5 GCM for the period 1950-1999 and compared with the bias corrected dataset using quantile mapping from Bureau of Reclamation (BOR). Further, the methods were applied to CNRM-CM5 hindcasts and compared with univariate asynchronous regression. A metric, fractional bias, was defined, and the distribution of fractional bias from ensembles was considered for comparing with the univariate approach. Asynchronous CCA performs better than the univariate approach in preserving the cross-correlation at grids where observed cross correlations are significant, while reducing fractional biases in mean and standard deviation. Results also show that preservation of cross-correlation

even with slightly increased bias in cross-correlation results in increased likelihood of estimating the joint probability of precipitation and temperature for months exhibiting significant cross-correlation.

Univariate downscaling has the ability to preserve observed mean and variance of climate forcing. But multivariate downscaling performs better to preserve the cross-correlation between multiple variables. In the second problem, the intrinsic property of univariate downscaling technique to estimate observed cross-correlation is analyzed by assuming and applying a linear regression model separately on two climate variables. Analytical solution suggests that univariate technique does not have the ability to estimate observed cross-correlation in downscaled variables. Constraint on univariate approach's' performance increases the role of raw GCM dataset. The current study proposes a framework, based on Fisher z-transformation to evaluate the quality of raw GCM/downscaled cross-correlation based on the observed cross-correlation. The framework is applied on cross-correlations between precipitation and temperatures projections from 3 General Circulation Models of CMIP5 multi-model ensemble. Results are analyzed over contiguous US and 9 NCDC climate regions. Results show that 40-50% (historical) and 30-40% (hindcast) grid points have inability to estimate cross-correlation between precipitation and temperature when univariate downscaling is applied. Equal probability exists, irrespective of the model, ensemble size or climate region, that univariate downscaling would be unable to estimate the observed cross-correlation of downscaled variables within the allowable band.

In first two problems, the current study deals with the uncertainty arising from downscaling approaches, while model uncertainty in climate change projections is addressed

in the third and last problem. GCM projections have high spatial bias in estimating the change in mean seasonal temperature over contiguous US and 9 NCDC climate regions. Observed mean of seasonal temperature for the latter half of 1950-1999 has increased over the Northern climate regions for January-February-March, whereas, Southern climate regions experienced a decrease between 1-2 °C in mean seasonal temperature for October-November-December. Equal weighting of GCM models can reduce the uncertainty in model projections but high spatial bias reduces the performance of equal weighting. Hence, to reduce uncertainty and eliminate spatial bias in climate change projections, the current study proposed an optimally weighted model combination approach. RMSE and correlation between observed change and multi-model combined change are calculated. Results show that optimal weighting scheme performs better to maintain a lower RMSE than equal weighting scheme. Slope in observed change is also maintained when optimal weighting is performed for multi-model combination.

Future course of the current work leads to several problems on the dependency between cross-correlation and hydrologic fluxes and model combination approach to reduce uncertainty in extreme climate prediction. Multivariate bias-corrected hindcast and historical runs with accurately estimated cross-correlation could improve the estimation of hydrologic fluxes. In my future work, I would like to investigate the sensitivity between cross-correlation of climate forcing and simulated hydrologic variables. We expect that multivariate bias-correction of climate variables could estimate the future changes in cross-correlation between precipitation and temperature. Hence, I would like to test the hypothesis that regions where increased precipitation recycling had occurred during the past, experience improved estimation of hydrologic fluxes from multivariate bias-corrected climate forcing. On the other hand, Fisher

z-transformation provided an allowable range to estimate downscaled cross-correlation based on the observed. I would like to apply the Fisher z-transformation framework on downscaled products from multivariate or other complex univariate approaches. The analysis would assess the extent to which multivariate or complex univariate approaches are able to estimate the observed cross-correlation. On multi model combination side, the future course is based on the reduction of model uncertainty in extreme climate predictions for the future period and stable estimate of spatial bias in climate change projections. The current study on downscaling and model uncertainty reductions provides the basis for future works. I expect to improve the quality of climate forcing hence reducing the error from downscaling during hydrologic simulations. On the other hand, my efforts to reduce model uncertainty in climate change projection lead to improve future changes exhibited by the GCMs. For long-term climate change projections, scenario uncertainties starts playing dominant role compared to model uncertainty. A future research encompassing the reduction of model and scenario uncertainty would improve the long-term climate change projections. I expect, the study I performed, would help to look at the future with a clearer and optimistic vision.