Florida State University Libraries

Electronic Theses, Treatises and Dissertations

The Graduate School

2012

Rainfall Anomaly Prediction Using Statistical Downscaling in a Multimodel Superensemble over Tropical South America

Bradford D. Johnson

THE FLORIDA STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

RAINFALL ANOMALY PREDICTION USING STATISTICAL DOWNSCALING IN
A MULTIMODEL SUPERENSEMBLE OVER TROPICAL SOUTH AMERICA

By

BRADFORD D. JOHNSON

A Thesis submitted to the
Department of Earth, Ocean & Atmospheric Science
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Summer Semester, 2012

Bradford Johnson successfully defended this thesis on June 15th, 2012.

The members of the supervisory committee were:

Paul Ruscher
Major Professor

Guosheng Liu
Committee Member

Vasu Misra
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the thesis has been approved in accordance with university requirements.

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. T. N. Krishnamurti, for his support and input. I give a special thanks to Vinay Kumar, Ph.D. for his persistence and patience throughout this process. Without your assistance, these words may never have been put on paper. Finally, I would also like to thank my wife and my family for constant encouragement and lending an ear when I would talk endlessly about these superensemble models. They still loved me when I talked about Brazil despite my failure to plan a vacation that actually takes them there...yet. Virtue, itself, smiles at their example of patience.

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
ABSTRACT	viii
CHAPTER 1. INTRODUCTION	1
1.1. Climate Overview	1
1.2. Objectives of Thesis	2
1.3. Organization of Thesis	4
CHAPTER 2. DATA AND DOMAIN	8
CHAPTER 3. METHODOLOGY	10
3.1. Superensemble Technique	10
3.2. Synthetic Superensemble Technique	11
3.3. Member Models	13
3.4. Downscaling Methodology	14
3.5. Further SSE Refinement	15
CHAPTER 4. DESCRIPTION OF ANOMALIES	16
4.1. March – May 1984	18
4.2. March – May 1992	18
4.3. SST Influence	19
4.3.1. Case 1 – 1984	20
4.3.2. Case 2 – 1992	20
CHAPTER 5. MEMBER MODEL, ENSEMBLE MEAN, AND SSE FORECASTS	23
CHAPTER 6. SUBDOMAIN FORECASTS	33
CHAPTER 7. CONCLUDING REMARKS	41
REFERENCES	43
BIOGRAPHICAL SKETCH	48

LIST OF TABLES

Table 1. Details of the thirteen global coupled models used in this study. Acronyms are expanded in Table 2.	6
Table 2. Acronyms for model names, their respective institutional affiliation, and other terms used throughout the paper.	9
Table 3. Reference tool showing position of observed, member models, ensemble mean, and modified superensemble (SSE-B) unless specified otherwise. SSE-A is not pictured.	22
Table 4. Best and worst performing member models compared to SSE-B according to 20-year MAM mean RMS error (mm/day) in each subdomain. Values in parentheses are mean RMS errors for member models in indicated year.	37

LIST OF FIGURES

Figure 1. MAM climatology of rainfall (mm/day) for the 1982 - 2001 period using MERRA precipitation. Downscaled rainfall (left) and coarse resolution rainfall (right) are shown. Note the similarities in anomaly positioning but the differences in resolution.	16
Figure 2. . Enlarged view of observed MERRA rainfall anomalies (mm/day). Positive anomalies are shown over Brazil during MAM of 1984 (top), while widespread negative values are evident in MAM of 1992 (bottom). Positive (negative) anomaly values are shown off the northwestern coast in the bottom (top) figure indicating enhanced (diminished) rainfall totals during documented El Niño (La Niña) events.	17
Figure 3. Kaplan Extended SST V2 anomalies (°C) for Case 1 (left column) from December 1983 (top left) to May 1984 (second panel from bottom, left). Anomalies for Case 2 from December 1991 (top right) to May 1992 (second panel from bottom, right). 6-month means from December to May for Case 1 and Case 2 are shown at bottom left and bottom right, respectively. Images are provided by the NOAA-ESRL Physical Sciences Division, Boulder, Colorado from their website at http://www.esrl.noaa.gov/psd/data/gridded/data.kaplan_sst.html	22
Figure 4. Anomaly correlation (top) and RMS error (bottom) annually from 1981 – 2001 during MAM using the downscaling technique. SSE-B outperforms each member model, the EM, and SSE-A.	24
Figure 5. Forecasts of precipitation anomalies (mm/day) in MAM 1984 over the entire domain utilizing the downscaling method. Anomaly correlations of each model are included as the top number in each panel. RMS errors are displayed as the lower number in the same manner. The anomaly correlation and RMS error of the SSE-A (not pictured) are 0.40 and 1.41, respectively.	25
Figure 6. Same as Figure 5 except using coarse resolution data. The anomaly correlation and RMS error of the SSE-A (not pictured) are -0.53 and 2.13, respectively.	26
Figure 7. Forecasts of precipitation anomalies (mm/day) in MAM 1992 over the entire domain utilizing the downscaling method. Anomaly correlations of each model are included as the top number in each panel. RMS errors are displayed as the lower number in the same manner. The anomaly correlation and RMS error of the SSE-A (not pictured) are 0.69 and 1.18, respectively.	27
Figure 8. Same as Figure 7 except using coarse resolution data. The anomaly correlation and RMS error of the SSE-A (not pictured) are -0.04 and 1.49, respectively.	28
Figure 9. Anomaly correlations in mm/day (top) and RMS error (bottom) for precipitation forecasts for Cases 1 and 2, MAM 1984 and 1992 over the entire domain. Coarse and downscaled data are included.	29
Figure 10. Scatter plot of anomaly correlations using the downscaling method during MAM from 1982-2001. Model rainfall is compared to observed rainfall in mm/day. Each model is denoted in the lower left of each panel. Mean anomaly correlations are shown in the lower right of each panel.	31

Figure 11. Graphical representation of subdomains over South America.	32
Figure 12. Average of seasonal anomaly correlations for each member model, the EM, SSE-A, and SSE-B over 20-yr period (1982 - 2001) during MAM over the entire domain and each subdomain.	34
Figure 13. Anomaly correlations (top) and RMS error (bottom) for model precipitation forecasts utilizing the downscaling method for Domain A.	35
Figure 14. Same as Figure 13 except for Domain B.	36
Figure 15. Same as Figure 13 except for Domain C.	38
Figure 16. Same as Figure 13 except for Domain D.	39
Figure 17. RMS errors (mm/day) for MAM precipitation forecasts (1984 and 1992) in the four subdomains – Domain A (upper left), Domain B (upper right), Domain C (lower left), and Domain D (lower right). While the downscaled SSE-A forecasts are generally better than the member models, the downscaled SSE-B forecasts are clearly better in every domain as evidenced by the last group in each chart.	40

ABSTRACT

Accurate forecasting of seasonal rainfall is important for many areas of commerce, including, but not limited to, agriculture, manufacturing, and power generation. There have been several studies addressing the predictability of rainfall variations over South America and the Amazon basin. The use of coupled atmospheric-ocean numerical models to predict these anomalies has had mixed results. A primary factor leading to model inaccuracy in precipitation forecasts is the coarse resolution data utilized by coupled models during the initialization phase. By using MERRA reanalysis and statistical downscaling along with the superensemble methodology, it is possible to obtain more precise forecasts of rainfall anomalies over tropical South America during austral fall. Selective inclusion (and exclusion) of superensemble member models also allows for increased accuracy. Improvement in individual member models is also possible on smaller spatial scales and in regions where substantial topographical changes were not handled well under original model initial conditions. The combination of downscaling and superensemble methodologies with other research methods presents the potential opportunity for increased accuracy not only in seasonal forecasts but on shorter temporal scales as well.

CHAPTER 1. INTRODUCTION

The South Asian monsoon has been well documented in literature. It is the most pronounced case of seasonal reversal of wind direction on the planet. Similar monsoons occur over eastern Asia, western North America, and over the Amazonian regions of South America. Much of the world's economy including, but not limited to agricultural and commercial industries, depends upon the regularity, predictability, and magnitude of the rainfall brought by these warm season events. The late arrival of the monsoon may leave crops destitute, drinking water at dangerously low levels, and energy costs at record levels due to subsequent drought and in some cases decreased generation of hydroelectric power.

1.1. Climate Overview

There are general similarities between the Asian monsoon and the South American monsoon. The Asian monsoon results from a buildup of available potential energy which converts to divergent kinetic energy then rotational kinetic energy. This subsequently leads to formation of the upper tropospheric anticyclone over Tibet (Krishnamurti and Ramanathan 1982; Wu and Zhang 1998). A similar process occurs over South America but from differing sources of potential energy. Li and Fu (2004) described this in detail. The Asian monsoon is initiated by differences in sensible heat caused by the Himalayas and the Tibetan Plateau to the north and the Indian Ocean to the south (Wu and Zhang 1998). Such a massive elevation difference is not present over the South American continent. In fact, the majority of the land mass is situated near the equator so differential heating is not the mechanism that causes the instability needed to initiate the wet season. The South American monsoon initiates from evapotranspiration processes whereby latent heat release fuels early season convection (Fu et al. 1999). Following this thinking, the effects of deforestation, the subsequent increase in albedo and lower seasonal precipitation amounts, may result from a decrease in latent heat flux as evapotranspirative processes contributions are diminished (Berbet and Costa 2003).

The South American wet season typically begins during October/November with documented cases as early as September and as late as December. The end of the wet season coincides with the austral fall and decrease in incoming solar heat flux. March is

usually the last month where above normal rainfall is observed over the domain (Wang and Fu 2002). Fu et al. (2001) showed the large sensitivity of seasonal precipitation over the eastern Amazon to the seasonal changes of SST in the Atlantic basin including the Caribbean Sea. A positive correlation exists in the southern oceans when SST leads precipitation by one to two months with respect to eastern Amazon precipitation. Larger rainfall totals over the eastern Amazon during austral fall are associated with above normal SST in the southern oceans and below normal SST in the northern oceans (Wang and Fu 2002).

There are different lines of thinking when it comes to defining the exact onset date of the South American monsoon. Kousky (1988) and Horel et al. (1989) suggest decreased outgoing longwave radiation values that approach certain thresholds over pentads can accomplish this. Marengo et al. (2001) follows increases in precipitation over several pentads to determine onset. This study also notes the progression of the monsoon throughout South America with the origin over the northwest Amazon River basin then to the southeast and finally to the mouth of the Amazon River and the Nordeste region of Brazil. Li and Fu (2004) argue the increase of convergence at the surface near and just south of the equator due to increased convection during the initiating phase cause the cross-equatorial flow to become northerly. This low-level flow is characterized by a V index, which is the average of the 925 hPa meridional winds over a domain (Wang and Fu 2002). The northerly flow causes persistent advection of moisture from the Caribbean Sea to south of the equator including western regions of the Amazon. This region coincides with lower surface pressures, increased convection, and development of the Bolivian high at 200 hPa. These factors are indicators of the existence of the South American monsoon season. The withdrawal of the monsoon is much more gradual and progresses from the southeast first then northward.

1.2. Objectives of Thesis

The purpose of this thesis is to thoroughly investigate the performance of coupled models toward precipitation climatology and anomalies over tropical South America, specifically toward the demise of the wet season. Several studies have been conducted in order to validate the ability of models to accurately predict the Asian monsoon on a seasonal timescale. There has been some success in this regard from the response of

seasonal climate to El Niño or La Niña events (Barnett et al. 1994; Meehl and Hu 2006; Philander 1990). Both dynamical and statistical methods have had reasonable levels of success for such seasonal predictions (Gadgil and Dhorde 2005; Krishnamurti et al. 1999). There have also been efforts to verify the usefulness of numerical modeling on these temporal scales over tropical South America. The ability to predict seasonal rainfall in advance has many benefits. The interannual variability of seasonal rainfall over South America has been investigated by many authors (Marengo et al. 2004; Fu and Wang 2004; Nobre et al. 2006).

Among the methods used to improve model performance for a variety of variables are dynamical and statistical downscaling. In predicting a variety of variables, including, but not limited to, precipitation, geopotential, specific humidity, and temperature, several authors have shown the coarse resolution of global models may be improved with the use of nested, higher resolution regional models over a desired domain (Chan and Misra 2011; Misra et al. 2003; Druyan et al. 2002; Nobre et al. 2001). Robertson et al. (2004) showed that dynamical downscaling might be beneficial in predicting precipitation anomalies on longer temporal scales and even down to 10-day dry spells in some areas.

Nobre et al. (2001) used NCEP's regional spectral model with a global forcing by the ECHAM3 AGCM at 80 km and then at 20 km. They found improvements over the AGCM at the 80 km resolution but also noted the 20 km results were less realistic. In a similar fashion, Druyan et al. (2002) forced a regional climate model at 50 km with the Goddard Institute for Space Studies GCM and developed improved forecasts.

Misra et al. (2003) utilized the Center for Ocean-Land-Atmosphere Studies global spectral model version 2.2 with a dynamical core from NCAR CCM3 forcing a regional spectral model developed by (Juang and Kanamitsu 1994). The Arakawa-Schubert deep convection scheme was replaced with the Simplified Simple Biosphere Scheme. Their results improved upon the AGCM in the Amazon River basin but degraded precipitation in the Nordeste. The higher resolutions allowed them to see the low-level jets but still underpredicted these. The authors mentioned the use of coarse data in the initialization and model biases may have contributed to these underestimations. Chan and Misra (2011) used an anomaly nesting technique whereby the model climatology is replaced with a higher resolution reanalysis to combat this problem.

In statistical downscaling, a higher resolution dataset is used as opposed to a second model by using regression and stochastic statistical methods (Wilby and Wigley 1997). Statistical downscaling is less computationally expensive than dynamical methods. Over South America, there have been few studies that utilize statistical downscaling (Valverde Ramírez et al. 2006; Solman and Nuñez 1999; Mendes and Marengo 2010).

In this thesis, I shall illustrate the performance of as many as 14 state of the art atmosphere ocean global coupled models providing 20 years of forecasts. The performance of these models individually over the domain is quite poor. A starting point in seasonal forecasts is the model climatology. This paper shows that given reliable and spatially precise rainfall observations from recently developed datasets, a model downscaling for forecasts (Krishnamurti et al. 2009; Chakraborty and Krishnamurti 2009) is possible over the domain. If one follows up next with the construction of a multimodel superensemble (Krishnamurti et al. 2006) then it is possible to obtain seasonal rainfall climatology that carries very high skill. Simple skill metrics for the evaluation of the climatology forecasts include the root mean square errors, and spatial correlations.

1.3. Organization of Thesis

The primary goals of this research project are:

- To determine skill of single member global climate models versus multiple member model ensemble versus a synthetic superensemble of these models.
- To utilize NASA's Modern-Era Retrospective Analysis dataset as a basis of downscaling data to obtain forecasts that are more accurate.
- To determine what effects selective inclusion of member models has on superensemble performance.
- To investigate model performance over subdomain regions to ascertain usefulness of methods on smaller spatial scales.
- To interpret relationships between results and findings in related studies.

A description of the data used and the domain is given in Chapter 2. The synthetic superensemble methodology, the use of coupled ocean-atmosphere models, and the downscaling technique is thoroughly explained in Chapter 3. Overall observed

climatology and the anomaly pattern of separate wet and dry season cases are explored in Chapter 4. Chapter 5 provides results of model, ensemble mean and superensemble forecasts. Simple skill metrics for the evaluation of the climatology forecasts are examined in Chapter 6. Conclusions are given in Chapter 7.

Table 1. Details of the thirteen global coupled models used in this study. Acronyms are expanded in Table 2.

Name (institute) and reference	Atmospheric component			Oceanic component			Ens. member
	Model	Resolution	Initial Condition	Model	Resolution	Initial Condition	
CFS (NCEP) (Saha et al. 2006)	Global Forecast System	T62L64	CFS SST forecast	Modular Ocean Model version 3 (MOM3)	$1^\circ \times 1/3^\circ$, 40 levels	Ocean data assimilation	15
POAMA 1.5 (Australia). (Zhong et al. 2005; Wang et al. 2004)	Bureau of Meteorology Research Center (BMRC) Atmospheric Model (BAM3)	T47L17	From latest atmosphere and ocean conditions from Global Atmospheric Sampling Program	Australian Community Ocean Model 2 (ACOM2)	$2^\circ \times 0.5^\circ$ - 1.5° , 31 levels	From ocean assimilation that was based on optimum interpolation (OI) technique	10
CERFACS (France)	Action de Recherche Petite Echelle Grande Echelle (ARPEGE)	T63L31	40-yr ECMWF Reanalysis (ERA-40)	Océan Parallélisé (OPA 8.2)	$2^\circ \times 2^\circ$, 31 levels	Forced by ERA-40	9
ECMWF (Europe)	Integrated Forecast System (IFS)	T95L40	ERA-40	HOPE-E	$1.4^\circ \times 0.3^\circ$ - 1.4° , 29 levels	Forced by ERA-40	9
Frontier Research Center for Global Change (FRCGC) (SINTEX-F) (Luo et al. 2005)	ECHAM4	T106L19	NCEP-Department of Energy (DOE) reanalysis 2	OPA 8.2	$2^\circ(\text{lon}) \times 2^\circ \cos(\text{lat})$, 31 levels	SST nudging scheme	9
GFDL (Delworth et al. 2006)	AM2.1	$2.5^\circ \times 2^\circ$, 34 levels	NCEP-DOE reanalysis 2	OM3.1 (MOM4)	$1^\circ \times 1/3^\circ$, 50 levels	Ocean data assimilation	10
INGV (Italy)	ECHAM4	T42L19	Couple AMIP type	OPA 8.1	$2^\circ \times 0.5^\circ$ - 1.5° , 31 levels	Forced by ERA-40	9
LODYC (France)	IFS	T95L40	ERA-40	OPA 8.2	$2^\circ \times 2^\circ$, 31 levels	Forced by ERA-40	9
MPI (Germany)	ECHAM5	T42L19	Coupled run relaxed to observed SST	MPI Open Model Interface (MPI-OMI)	$2.5^\circ \times 0.5^\circ$ - 2.5° , 23 levels	Couple run relaxed to observed SST	9
MetFr (France)	ARPEGE	T63L31	ERA-40	OPA 8.0	182×152 GP, 31 levels	Forced by ERA-40	9

Table 1. – continued.

Name (institute) and reference	Atmospheric component			Oceanic component			Ens. member
	Model	Resolution	Initial Condition	Model	Resolution	Initial Condition	
Seoul National University (SNU) (Kug et al. 2007)	SNU	T42L21	NCEP-DOE reanalysis 2	MOM 2.2	$1^\circ \times 1/3^\circ$, 32 levels	SST nudging scheme	6
UH (Fu and Wang 2004)	ECHAM4	T31L19	NCEP-DOE reanalysis 2	UH Ocean	$2^\circ \times 1^\circ$, 2 levels	Thermocline-depth nudging	10
UKMO (United Kingdom)	HadAM3	$2.5^\circ \times$ 3.75° , 19 levels	ERA-40	GloSea OGCM Third Hadley Center Coupled Ocean- Atmosphere GCM (HadCM3) based		Forced by ERA-40	9

CHAPTER 2. DATA AND DOMAIN

Thirteen coupled models were used in this study to investigate the rainfall forecasts over tropical South America. A brief explanation of resolution, physics and components of the models has been given in the Table 1. These are discussed in more detail later. Forecasts from each model cover different temporal scales. Most of the models were integrated from 1 May to 30 September for the boreal summer season, and integrated from 1 November to 31 March for the boreal winter. The acronyms for the model names are provided in Table 2. Entire datasets for the model's forecast were obtained from the CliPAS project (Wang et al. 2008). A high resolution (0.5° latitude by 0.667° longitude) monthly observed global rainfall is obtained using the Modern Era Retrospective-analysis for Research and Applications (MERRA) dataset (Rienecker et al. 2011). The temporal resolution of this dataset is from 1979 to present day. This study will focus on the 20 year period from 1982-2001. The domain encompasses the entire Amazon basin and the adjacent regions from 30°S to 10°N latitude and 30°W to 80°W longitude.

Table 2. Acronyms for model names, their respective institutional affiliation, and other terms used throughout the paper.

MERRA	Modern-Era Retrospective Analysis for Research and Applications (NASA), United States
BMRC	Bureau of Meteorology Research Center, Australia (also POAMAI)
CERF	European Center for Research and Advanced Training in Scientific Computation, France (CERFACS)
ECMWF	The European Center for Medium-Range Weather Forecasting, United Kingdom
GFDL	Geophysical Fluid Dynamics Lab, United States
INGV	Istituto Nazionale de Geofisica e Vulcanologia, Italy
LODY	Laboratoire d’Océanographie Dynamique et de Climatologie, France
MAXP	Max-Planck Institut für Meteorologie, Germany
METF	Centre National de Recherches Météorologiques, Météo-France, France
NCEP	National Centers for Environmental Prediction, United States
SINT	Scale Interaction Experiment-FRCGC
SNU	Seoul National University, South Korea
UH	University of Hawaii, United States
UKMO	Met Office, United Kingdom
EM	13 member model ensemble mean
SSE	Synthetic superensemble
SSE-A	Synthetic superensemble of all 13 member models
SSE-B	Modified synthetic superensemble of 11 member models
MAM	March-April-May
RMSE	Root mean square error
AGCM	Atmospheric Global Circulation Model

CHAPTER 3. METHODOLOGY

3.1. Superensemble Technique

Krishnamurti et al. (1999, 2000a) developed a superensemble technique that utilizes a set of multimodel forecasts to derive a single consensus forecast that differs from the ensemble mean of the models. This method was inspired by the research of Lorenz (1963) who suggested that “very-long-range” forecasts may be developed when using two or more numerical solutions. The dataset utilized by Krishnamurti et al. is separated into two different time segments, the training phase and the forecast phase. The individual model performance is derived using multiple linear regression against the observations in the domain during the training phase. Each model in the ensemble is assigned a weight based upon the results of the regression. These weights are then used in the forecast phase of the superensemble. The full field variables are not used in determining the weights assigned to each model. The multiple regression technique utilizes the temporal anomalies of the variables instead. The individual model weights are implemented in the forecast construction in the following manner:

$$S = \bar{O} + \sum_{i=1}^N a_i (F_i - \bar{F}_i), \quad (1)$$

where S is the constructed forecast; O is the observed climatology; a_i is the weight for the i th member in the ensemble; and F and $\bar{F}_i$ are the forecasts and forecast climatological values, respectively, for the i th model’s forecast. N designates the number of member models the summation is taken over. The climatology O and a_i originate from different sources. Chakraborty and Krishnamurti (2006) calculate S for 52 different seasons. The anomalies calculated in this study are for different years so the coefficients do not need to add up to zero. If the superensemble performs poorly over the course of a year during the forecast period, this can result in differences between S and $\bar{O}$.

The model weights a_i are determined through minimization of the error term G , which is expressed as

$$G = \sum_{t=1}^{N_{train}} (S'_t - O'_t)^2. \quad (2)$$

N_{train} is the number of time samples in the training phase of the superensemble. O'_t and S'_t are the observed and superensemble field anomalies respectively at any training time t . This process is performed over the entire domain at each grid point and vertical level in the aforementioned dataset during the forecast phase. Each model is given a weight at each grid point in all models in three-dimensional space. The member model with the strongest performance at one grid point possesses a higher weight whereas a poor performing model may even have a negative weight. Using this methodology, you achieve weights that do not vary greatly for a model grid or any particular variable from year to year. This is indicative of the stability of the method and its ability to perform well on the seasonal time scale.

This forecast method is different from the traditional bias removed ensemble mean because it does not use $1/N$ as its weights. Individually assigning weights to the model and variable grid points has proven to perform better than letting $a_i = 1/N$ for $I = 1, 2, 3, \dots, N$ (Stefanova and Krishnamurti 2002). In a multimodel bias-removed ensemble,

$$E = \frac{1}{N} \sum_{i=1}^N (F_i - \overline{F_i}) + \overline{O}, \quad (3)$$

the superensemble scales the individual model forecast by its relative performance during the training period in a manner mathematically equivalent to weighting it.

3.2. Synthetic Superensemble Technique

The synthetic superensemble methodology is a variant of the superensemble methodology which is shown to provide better forecast skill than that of the conventional superensemble (Yun et al. 2005). The synthetic superensemble forecast is obtained when synthetic datasets are used in the superensemble algorithm in place of the member model forecast datasets. One set of synthetic data is created corresponding to each model by a linear regression with the observed (analysis) field in the EOF space. This technique is described below.

The observational data can be written as a linear combination of the principal component (PC) and the spatial EOFs:

$$O(x, t) = \sum_{k=1}^M P_k(t) \Phi_k(x), \quad (4)$$

where $P_k(t)$ is the PC time series, $\Phi_k(x)$ denotes the EOF component of the k th mode, and M is the number of modes selected. In the present study, M was chosen such that the EOFs explain 99% of the variance in the actual dataset. The results are not dependent on this number as long as the variance explained by the EOFs is high (95%).

Similar to the observations, the model datasets are also expanded in terms of PC time series and EOFs:

$$F_i(x, t) = \sum_{k=1}^M F_{i,k}(t) \phi_{i,k}(x), \quad (5)$$

where I denotes the i th-member model in the ensemble, $F_{i,k}(t)$ and $\phi_{i,k}(x)$ are the PC time series and spatial EOF, respectively, for the k th mode, and M is the number of modes selected.

Now, a consistent pattern of the forecast and observation data is created by multiple linear regression. The k th mode of the PC time series of the observations is written as a linear combination of the k th mode of the PC time series of different models with an error term:

$$P_k(t) = \sum_{i=1}^N \alpha_{i,k} F_{i,k}(t) + \varepsilon_{i,k}, \quad (6)$$

where I denotes the i th member in the ensemble, $\alpha_{i,k}$ is the weight of the i th model for mode k , and $\varepsilon_{i,k}$ is the error term. The α s are estimated using a multiple linear regression, which minimizes the error variance $E(\varepsilon^2)$. After the weights are computed, the regression improved PC time series of the k th mode for the i th model is defined as

$$F_{i,k}^{\text{reg}}(t) = \alpha_{i,k} F_{i,k}(t), \quad (7)$$

and the synthetic data for the i th-member model is defined as

$$F_i^{\text{reg}}(x, t) = \sum_{k=1}^M F_{i,k}^{\text{reg}}(t) \Phi_k(x), \quad (8)$$

where $\Phi_k(x)$ represents the EOF components of the observations. Once the synthetic data $[F_{i,k}^{\text{reg}}(x, t)]$ are generated in this manner for every model of the ensemble, these are then passed on to the input of the conventional superensemble algorithm to construct what is called a synthetic superensemble forecast. The synthetic superensemble algorithm

is schematically described in Fig. 1. It was found that the synthetic superensemble performs better than the conventional superensemble (Yun et al. 2005). In this paper, the phrases *superensemble* and *synthetic superensemble* will be used interchangeably to refer to the synthetic superensemble forecasts.

3.3. Member Models

As many as 13 state-of-the-art atmosphere–ocean models were used for the construction of the superensemble. The datasets used were acquired from our lab’s personal contacts with the data producers. In Table 1, some details on model resolution, physical parameterizations, description of the ocean modeling, years of model runs and relevant references are provided. This also includes the number of ensemble forecasts provided for each model run. Table 2 provides a guide of frequently used acronyms. The ensemble mean forecasts from a single model’s several runs are also included in this study. These model forecasts are cast at a common horizontal resolution of 2.5° latitude by 2.5° longitude for the construction of model ensembles. The models carry a diversity of physical and oceanic parameterizations. This provides a robust ensemble for the seasonal climate experiments.

A coupled version of the Community Climate Model, version 3 (CCM3), from the National Center for Atmospheric Research (NCAR) was also included in this suite of member models. The atmospheric component of this model was run at the T63 resolution and with 18 vertical levels. The oceanic component was adapted from Hamburg Ocean Primitive Equation (HOPE) global model (Wolff et al. 1997), which has a resolution of 5° in longitude and 0.5° – 5.0° in latitude with higher resolution over the equatorial latitudes. The oceanic initial conditions were obtained from a comprehensive coupled assimilation where the initialized SSTs were relaxed to observed SST (Krishnamurti et al. 2002).

The Predictive Ocean Atmosphere Model for Australia-1 (POAMA1) is also included in this multimodel suite. The atmospheric component of the POAMA1 is run at an R47 resolution (rhomboidal truncation at 47 waves) and it has 17 vertical levels. A detailed description of this model can be found in (Wang et al. 2004).

Two independent versions of ECHAM4 initialized from NCEP-DOE reanalysis 2 with differing oceanic component schemes are utilized from FRCGC (Luo et al. 2005)

and the University of Hawaii (Fu and Wang 2004). The atmospheric components of these are run at horizontal resolution of $2^\circ \times 2^\circ$ and about 3.75° , respectively. The Seoul National University model with the MOM 2.2 oceanic component is included as well (Kug et al. 2007).

Finally, the seven member models of the Development of a European Multimodel Ensemble System for Seasonal to Interannual Prediction (DEMETER) project were included in this suite. The DEMETER consists of a variety of models from different parts of Europe. A description of this project is provided in Palmer et al. (2004). There were nine members in an ensemble simulation for each model of the DEMETER family. An average over these nine members was taken before using these data in the superensemble. Specifics on each model in the DEMETER project may be found in Table 1.

The superensemble methodology has been used to provide better forecasts for a variety of phenomena, including, but not limited to, the South Asian monsoon and tropical cyclone track and intensity (Williford et al. 2003). Improvements in these forecasts coupled with Lorenz's deductions suggest the superensemble may be a very useful tool in the long-term prediction of anomalous precipitation events globally, including over tropical South America.

3.4. Downscaling Methodology

While this superensemble technique is now over ten years old, its reliance on statistics as opposed to parameterizations allow it to remain relevant and at the forefront of forecasting tools. It is possible to utilize any global coupled model as a member in these forecasts. Limiting factors in forecasting accuracy are nonlinear model parameterizations along with the lack of high-resolution data during initialization. The ability to use dependable data on smaller scales in the initialization and training phases of models would greatly aid in better forecasting.

The SSE aims to reduce these model biases previously referred to by Misra et al. (2003). This coupled with a statistical downscaling may lead to substantially better forecasts on a regional scale. This particular method of downscaling differs from statistical downscaling in a traditional sense. The process is explained here. Table 1 shows the spatial scales of member model forecasts in this study. Recent reanalysis by Rienecker et al. (2011) provides rainfall data on a finer grid scale. Utilizing a bilinear

interpolation, model rainfall datasets are interpolated to MERRA's grid resolution (0.5° latitude $\times$ 0.67° longitude). Kumar and Krishnamurti (2012) show:

$$R_{\text{obs}} = aR_{\text{mdl}} + b + \varepsilon, \quad (9)$$

where R_{obs} and R_{mdl} are the observed and interpolated model forecasts of rainfall, respectively; a and b are regression coefficients known as the slope and intercept of the least fitting squares; and ε is the error term. The main idea is to minimize the absolute value of $(|\varepsilon|)$. Regression coefficients are calculated independently using the cross-validation principle where the regression coefficient for calculated year is held aside while the remaining years in the dataset are used to calculate the regression coefficients. The downscaled model rainfall is then obtained using the aforementioned coefficients:

$$R_{\text{dscl}} = aR_{\text{mdl}} + b, \quad (10)$$

where R_{dscl} is the downscaled rainfall forecast of the model; here, a and b are calculated using Equation (9) at each grid point and separately for every month of the year. The information carried spatially and temporally by these coefficients are used to develop downscaled datasets from 1982 – 2001. It essentially provides data that includes model bias over the entire domain and subdomains.

3.5. Further SSE Refinement

In practice, selective inclusion (and exclusion) of particularly ineffective model members involved in the calculation of SSE coefficients, has proven effective in attempts to increase forecast skill (Krishnamurti and Kumar 2012). The number of models needed for minimal RMS error increases with the magnitude of the temporal scale. In order to illustrate the potential benefits on this study, the SSE forecasts including all member models will be referred to as SSE-A. On the other hand, forecasts derived from the alternative, modified SSE will be denoted by SSE-B. The DEMETER group models make up seven of the member models. The best performing remaining models were retained in the new SSE. The two least effective (SNU and UH) were omitted. The effectiveness of each model was based upon SSE weightings, RMS error, and anomaly correlations. Consequently, SSE-B is developed from only 11 member models.

CHAPTER 4. DESCRIPTION OF ANOMALIES

Precipitation climatology and anomalies are shown in Figures 1 and 2, respectively, from 1982 – 2001. The year 1984 (1992) shows anomalously positive (negative) precipitation values in observed data. These two years were chosen to undergo further scrutiny. In the Amazon basin, negative (positive) precipitation anomalies are typically caused by the late (early) onset and early (late) demise of the wet season. Using EOF analysis, Jones et al. (2012) discovered in fact that the median onset and demise of the wet season were late October and late April, respectively. The same study shows the demise of the wet seasons in 1984 and 1992 fit the predicted outcome with dates near May 1st and April 15th, respectively. The observed anomalies over the domain are described here.

Figure 1. MAM climatology of rainfall (mm/day) for the 1982 - 2001 period using MERRA precipitation. Downscaled rainfall (left) and coarse resolution rainfall (right) are shown. Note the similarities in anomaly positioning but the differences in resolution.

Figure 2. . Enlarged view of observed MERRA rainfall anomalies (mm/day). Positive anomalies are shown over Brazil during MAM of 1984 (top), while widespread negative values are evident in MAM of 1992 (bottom). Positive (negative) anomaly values are shown off the northwestern coast in the bottom (top) figure indicating enhanced (diminished) rainfall totals during documented El Niño (La Niña) events.

4.1. March – May 1984

The majority of the Amazon basin experienced above average rainfall during MAM of 1984 (to be referred to as Case 1). Figure 2 (upper) shows observed rainfall anomalies over the domain. The Nordeste region to the east of the Amazon delta showed the most pronounced anomalies with an expansive area of daily rainfall on the order of 3 mm/day above climatology. While much of the basin has positive anomalies of greater than 0.6 mm/day, a maximum extends from the aforementioned Nordeste maximum westward to secondary maximums in Tocantins, Mato Grasso, southern Para, and eastern Amazonas near Manaus. These three maxima all exceed 3 mm/day. Manaus averages just under 2000 mm/year of rainfall. A surplus of approximately 300 mm fell during this three-month period.

Areas of below normal rainfall are also seen during Case 1. The southeastern Brazilian coast near Rio de Janeiro and the majority of Bolivia both saw rainfall deficits around 1 mm/day. A significant negative anomaly is shown along the Caribbean coast in Venezuela, Guyana, and Suriname that exceeds 3 mm/day over a wide area including parts of the state of Para in Brazil. It is easy to point out the 6 mm/day difference in rainfall between northern and southern sections of Para.

The effect of the El Niño Southern Oscillation (ENSO) is also depicted. A developing La Niña is documented during 1984 (Chou et al. 2003). The negative rainfall anomalies over the coast of Ecuador extending westward in the South Pacific Ocean are signs synonymous with a negative ENSO event. The cooler waters lead to decreased precipitation and subsequent downward motion over the area. Increased precipitation over the Amazon basin and delta tend to accompany La Niña events (Marengo et al. 2001).

4.2. March – May 1992

The primary distinct feature (lower Figure 2) shows below average rainfall on the order of 2 – 3 mm/day during MAM of 1992, hereafter referred to as Case 2, over the northern two-thirds of Brazil. One of larger deficits (~3 mm/day) is centered just east of the Amazonian delta over the coastal region in the state of Maranhao. Negative anomalies of at least 3 mm/day extend over the states of Piaui, northern Ceara, northern Tocantins, and the southeastern half of Para. A second minimum was centered near

Manaus in eastern Amazonas also measured greater than 2.6 mm/day. Anomaly values of 2 mm/day extend northward into the state of Roraima. Negative anomalies extend as far west as the Andes Mountains in Colombia and Peru. The only positive anomalies in the domain between 10°S and 10°N are isolated to Ecuador into extreme northwestern Peru and a second over north central Venezuela. The Ecuador anomaly has values approaching +3 mm/day. This case occurs within the timeframe of a strong ENSO event. Subsidence over the Amazonian basin due to excess convection in the eastern Pacific Ocean may have led to the suppression of precipitation over the region. Large positive rainfall anomalies are noted over Ecuador on the edge of the domain, which is typical of an ENSO event. Not to be neglected are the effects of previous SST fluctuations in the Atlantic basin. Both of these are considered in the analysis of this anomaly.

The negative anomaly in northern Brazil extends southward near the coast. The states of Bahia, Minas Gerais, and Rio de Janeiro show negative anomalies ranging from 0.5 to 1.5 mm/day. Conversely, the majority of the southern third of Brazil had a precipitation surplus. The maximum positive anomaly of approximately 3 mm/day was located near the Argentina – Paraguay border in the state of Parana. This positive anomaly is rather isolated and inland except to the south in Rio Grande do Sul where it remains positive but is less than 1.5 mm/day. A secondary maximum anomaly of approximately 2 mm/day is shown east of the Andes Mountains over central Bolivia.

4.3. SST Influence

Sea surface temperatures and anomalies are very important to climate forecasting and may have significant effects on the climate locally and globally depending upon the severity of the anomaly. The ENSO phenomenon is a prime example of the interdependence of the ocean and atmosphere. During a particularly strong ENSO, rainfall over the Amazon basin is suppressed. Increased convection over the western Pacific off the coast of Ecuador leads to outflow aloft and descending air over the majority of Brazil. This factor alone may cause widespread negative rainfall anomalies over the northern two-thirds of the domain. We may not, however, ignore the effects of the Atlantic SST.

4.3.1. Case 1 – 1984

The tropical Atlantic Ocean on average is cooler than average during DJF and MAM of 1983 – 1984. Weak anomalies on the order of 0.15° to 0.3°C are present just east of the Lesser Antilles in averaged monthly SST data (Figure 3). In contrast, a large positive anomaly centered around $15 - 20^{\circ}\text{S}$ off the southeastern coast of Brazil reaches a magnitude of 1.2°C . This anomaly extends northward to the equator and could aid in rainfall enhancement in a manner similar to increased precipitation on the coast of Ecuador during an El Niño event. Conversely, the negative precipitation anomaly on the Caribbean coast looks to be influenced by nearby cooler waters of the North Atlantic.

4.3.2. Case 2 – 1992

The largest negative rainfall anomalies exist on the coast of northeast Brazil just east of the Amazon delta. The proximity of these anomalies suggests another potential factor may have led to decreased rainfall other than ENSO. The SST anomalies over the Atlantic basin for the three months prior to the initialization until the end of MAM show persistent negative values between 0.5° to 1°C near the Brazilian coast (Figure 3). Outside the month of December 1991, the Atlantic Ocean showed distinct negative anomalies near the equator and Brazilian coast. It has been shown the anomalously cool SSTs tend to suppress convection in the region it occupies. More specifically, a relationship exists between positive Pacific SST anomalies and negative Atlantic anomalies in the tropics. Marengo et al. (2001) shows that in general this accompanies a late onset and early end to the monsoon. The lack of precipitation over the mainland would support this argument.

Figure 3. Kaplan Extended SST V2 anomalies (°C) for Case 1 (left column) from December 1983 (top left) to May 1984 (second panel from bottom, left). Anomalies for Case 2 from December 1991 (top right) to May 1992 (second panel from bottom, right). 6-month means from December to May for Case 1 and Case 2 are shown at bottom left and bottom right, respectively. Images are provided by the NOAA-ESRL Physical Sciences Division, Boulder, Colorado from their website at http://www.esrl.noaa.gov/psd/data/gridded/data.kaplan_sst.html.

Table 3. Reference tool showing position of observed, member models, ensemble mean, and modified superensemble (SSE-B) unless specified otherwise. SSE-A is not pictured.

MERRA (observed)	BMRC	CERF	ECMW
GFDL	INGV	LODY	MAXP
METF	NCEP	SINT	SNU
UH	UKMO	EM	SSE-B

CHAPTER 5. MEMBER MODEL, ENSEMBLE MEAN, AND SSE FORECASTS

From 1982 – 2001, seasonal precipitation forecasts during MAM over the domain utilizing the 13 coupled global circulation models, their EM, and the SSE were compared to the observed rainfall. Figure 4 is a guide that shows the layout of observed and model forecasts in a gridded format. Analysis shown in Figure 4 reveals SSE-B outperforms the vast majority of member models and their EM during each year. For example, the member models generally performed rather poorly during MAM of 1982 with many showing correlations close to or less than zero. SSE-B posts a correlation of 0.54 compared to a member model high of 0.34 by the LODY and an EM value of 0.06. In the year 2000, RMS error of the model EM was 2.11 with best performing model, SNU, posting a RMSE of 1.83. The SSE-B RMSE in the same year is 1.84. While this is not nearly an ideal forecast, it shows an example of performance near or above that of the individual member models. The improvements were not only evident in years of poor model performance. In 1983 and 1998, the EM correlations were 0.75 and 0.66, respectively, while the SSE-B improved anomaly correlations to 0.83 in both.

It is also important to note the differences in model performance in general between coarse resolution and downscaling forecasts. Several member models performed better with the coarse resolution data. The BMRC, ECMW, LODY all had lower RMS errors using the downscaled method. Both SSE versions made significant improvements with the downscaling technique. Following this method, SSE-B anomaly correlations are 0.62 and 0.80 for 1984 and 1992, respectively. A graphical representation of this is shown in Figures 5 – 8, while anomaly correlations and RMS errors are shown in Figure 9. Figure 5 shows the ability of SSE-B to forecast the strong precipitation gradient over Para and French Guiana. A small region of negative anomalies over the southeastern Brazilian coast in observed data are also shown in SSE-B. The maxima and minima over the Colombian and Ecuadorian coasts are picked up as well. In general, SSE-B duplicates the anomalies found in the observed data better than any of the member models or the EM. One caveat to this conclusion involves the combination of coarse resolution data and SSE-B. While the forecast is consistently better than SSE-A, the same level of improvements are not noted compared to the downscaled versions (Figure 8). The EM

Figure 4. Anomaly correlation (top) and RMS error (bottom) annually from 1981 – 2001 during MAM using the downscaling technique. SSE-B outperforms each member model, the EM, and SSE-A.

Figure 5. Forecasts of precipitation anomalies (mm/day) in MAM 1984 over the entire domain utilizing the downscaling method. Anomaly correlations of each model are included as the top number in each panel. RMS errors are displayed as the lower number in the same manner. The anomaly correlation and RMS error of the SSE-A (not pictured) are 0.40 and 1.41, respectively.

Figure 6. Same as Figure 5 except using coarse resolution data. The anomaly correlation and RMS error of the SSE-A (not pictured) are -0.53 and 2.13, respectively.

Figure 7. Forecasts of precipitation anomalies (mm/day) in MAM 1992 over the entire domain utilizing the downscaling method. Anomaly correlations of each model are included as the top number in each panel. RMS errors are displayed as the lower number in the same manner. The anomaly correlation and RMS error of the SSE-A (not pictured) are 0.69 and 1.18, respectively.

Figure 8. Same as Figure 7 except using coarse resolution data. The anomaly correlation and RMS error of the SSE-A (not pictured) are -0.04 and 1.49, respectively.

Figure 9. Anomaly correlations in mm/day (top) and RMS error (bottom) for precipitation forecasts for Cases 1 and 2, MAM 1984 and 1992 over the entire domain. Coarse and downscaled data are included.

did recreate the strong positive reflection over Nordeste but was inhibited itself by a large negative anomaly produced by most member models over the tropical Atlantic Ocean that is not present in observed data (Figure 6). The SSE-B, instead, nearly transposes this configuration.

In Case 2, member models generally performed well and the SSE-B does a particularly good job at graphically reproducing observed rainfall (Figure 7). Small scale maxima over the Andes mountains in Peru are not seen any member models nor the EM but are strikingly similar in the observed and the SSE-B fields. The tight gradients over Ecuador, Colombia, and Venezuela are displayed as well. A strong positive anomaly over Paraguay and extreme southern Brazil is missed by all models in magnitude. The SSE-B is the only model that spatially depicts this anomaly.

Figure 10 shows a scatter plot diagram relating observed MERRA rainfall anomalies to the anomalies member models, the EM, the SSE-A, and the SSE-B. Data points along an ordinate with slope equal to 1 are an exact hit at a particular grid point in the domain. Model forecasts are plotted along the x-axis and observed data are plotted along the y-axis. Data points along the y-axis except at the origin denote a predicted zero value by the model but a positive or negative anomaly in the observed data. A “perfect” forecast by any model would have all values along a straight line through the origin and slope equal to 1 with no scatter. A forecast with zero correlation would follow either axis. While this is not anticipated, it provides a barometer by which to interpret the results. Looking at the member models with the smaller correlations (i.e. UH, METF, and MAXP), there is a tendency for a thinner scatter that tightly surround the y-axis. Data points near or to the left of the y-axis at values greater than 10 mm/day show a significant underestimation of rainfall at a specific grid point over a 20-year period by a particular model. On the other hand, the better performing models, such as the ECMW and SINT, have a reduced number of these “misses” and also have more values near the slope = 1 line. For perspective, these better models still have mean anomaly correlations of less than 0.35, thus having a greater number of total “misses” than “hits”. The SSE-B has none of these significant underestimations on average. The profile visually tilts toward the slope = 1 line. A mean anomaly correlation of 0.69 suggests the aforementioned logic applies as expected.

Figure 10. Scatter plot of anomaly correlations using the downscaling method during MAM from 1982-2001. Model rainfall is compared to observed rainfall in mm/day. Each model is denoted in the lower left of each panel. Mean anomaly correlations are shown in the lower right of each panel.

Figure 11. Graphical representation of subdomains over South America.

CHAPTER 6. SUBDOMAIN FORECASTS

Model performance was also analyzed over the four smaller subdomains over South America located in Figure 11. The first region (Domain A) is located along Caribbean coast from Venezuela eastward to Guyana, Suriname, and French Guiana. The coordinates of Domain A are from 10°N 70°W to 2°N 50°W. Domain B is located over the Nordeste region of Brazil. Its coordinates are 2°S 50°W to 16°S 34°W. Domain C includes the greatest variations in elevation and is positioned over the upper Amazon basin including regions of Colombia, Venezuela, Peru, Brazil, and Bolivia. Paraguay and southeastern Brazil compose Domain D. The coordinates of Domain C and D are 5°N 75°W to 14°S 60°W and 15°S 60°W to 28°S 40°W, respectively.

Figure 2 shows observed rainfall anomalies over the entire domain. This shows that although Case 1 and 2 were generally above normal and below normal years, respectively, certain regions experienced differing precipitation amounts. The ability of AGCMs to predict the variations on smaller spatial scales with accuracy is very valuable, especially in localized industries. The use of higher resolution data may aid in developing these forecasts. As previously discussed, in 1984, the majority of Brazil saw a rainfall surplus but the Caribbean coast experienced significant deficits. A similar observation can be seen in 1992 where the southern third of the domain saw positive anomalies. Figure 12 compares the anomaly correlations of all four subdomains and the entire domain. SSE-B performs well over all four subdomains during the span of the study with all average anomaly correlations of at least 0.59. The member models, EM, and SSE-A all carry negative correlations in Domain D. The UH, METF, and MAXP models performed particularly poor in all domains compared to other member models which is a reflection of the results discussed surrounding Figure 10. A noticeable observation was model performance in Domain B. All of the member models except MAXP had positive correlations in this region. Figures 13 – 16 show the seasonal anomaly correlations and RMS error for each subdomain in MAM from 1981 – 2001. Table 4 shows a summary of the most (and least) effective seasons for the rainfall anomaly forecasts with respect for RMS error.

Figure 12. Average of seasonal anomaly correlations for each member model, the EM, SSE-A, and SSE-B over 20-yr period (1982 - 2001) during MAM over the entire domain and each subdomain.

Figure 13. Anomaly correlations (top) and RMS error (bottom) for model precipitation forecasts utilizing the downscaling method for Domain A.

Figure 14. Same as Figure 13 except for Domain B.

Table 4. Best and worst performing member models compared to SSE-B according to 20-year MAM mean RMS error (mm/day) in each subdomain. Values in parentheses are mean RMS errors for member models in indicated year.

	Domain A	Domain B	Domain C	Domain D
Best performing member model	SINT	ECMW	MAXP	CERF
Worst performing member model	UH	MAXP	METF	INGV
Best year	1991 (1.02)	1992 (1.21)	1986 (0.83)	1985 (0.48)
SSE-B in best year	0.85	0.70	0.59	0.33
Worst year	1988 (3.00)	1990 (2.62)	2001 (2.79)	2001 (1.59)
SSE-B in worst year	2.59	2.21	2.31	1.38*

*One of only two years with RMSE above 1.00 in Domain D (1998)

Of note is the general excellent performance of member models in Domain D. The mean RMSE for all models was between 0.89 and 0.93. This is significant when considering the next lowest mean RMSE from any member model in any other subdomain was 1.27 by MAXP in Domain C. As alluded to in Table 3, the SSE-B had precipitation forecasts that produced RMS errors in Domain D that were as low as 0.33 (Figure 16). While member models had the highest anomaly correlations in Domain B, the RMS errors in that region were not as impressive. The worst model performance was found in Domain A. It is in this region where we found the aforementioned deviations from the larger anomaly spatial trends in Cases 1 and 2.

In regards to these two cases, it is possible to investigate the differences in coarse resolution initializations versus the downscaled results reviewed in detail above. Figure 17 shows RMS errors for all four subdomains comparing coarse and downscaled precipitation forecasts. The larger trend of mixed results for member models continues throughout the smaller domains except in Domain C where downscaled forecasts are better across the board. This is also another opportunity to point out two repeating occurrences in this study. The model performance in Domain D compared to other domains; the consistent manner in which SSE-B either rivals or in many cases outperforms the best member model and the ensemble mean.

Figure 15. Same as Figure 13 except for Domain C.

Figure 16. Same as Figure 13 except for Domain D.

Figure 17. RMS errors (mm/day) for MAM precipitation forecasts (1984 and 1992) in the four subdomains – Domain A (upper left), Domain B (upper right), Domain C (lower left), and Domain D (lower right). While the downscaled SSE-A forecasts are generally better than the member models, the downscaled SSE-B forecasts are clearly better in every domain as evidenced by the last group in each chart.

CHAPTER 7. CONCLUDING REMARKS

Statistical downscaling has been shown to improve forecast skill (Chakraborty & Krishnamurti, 2009; Kumar & Krishnamurti, 2012; Nicholas & Battisti, 2012). While several studies have examined dynamic downscaling over South America, the limited literature in the study of statistical downscaling lends the opportunity to contribute an additional perspective in the matter. NASA's MERRA reanalysis is used in this study as a high-resolution alternative to original model data initializations. Brands et al. (2011) explored the challenges of using reanalysis data, specifically NCEP/NCAR and ERA-40, as the basis of statistical downscaling modeling. Reanalysis is used extensively throughout by many authors in statistical downscaling projects. All reanalysis datasets possess inhomogeneities that will cause issues in model performance. These errors arise primarily over the oceans but also over land regions. This is primarily due to errors in measurement, which is partly due to lack of observational data. The remote regions of the Amazon and the Andes Mountains are regions that fit these criteria. When looking at the model RMS error for Domain C, a substantial improvement was seen from coarse resolution forecasts to the downscaled forecasts. All but three of the member models used either NCEP or ERA-40 reanalysis in the atmospheric component as initial conditions. The higher resolution of the MERRA reanalysis helped to improve these forecasts. This may be due in part due to better handling of the topography over the region.

The superensemble technique provides a tool capable of reducing model bias while utilizing the downscaling method during the initialization phase. While the full 13-member superensemble did not consistently perform better than the ensemble mean forecast, the modified version of the superensemble was very accurate compared to the member models and their ensemble mean. This follows the findings by Krishnamurti and Kumar (2012) where the superensemble RMS error decreases as more models are included and the temporal scale increases but only to 12 member models for a 15 yr length dataset. The modified superensemble performed better across the majority of the board in all domains, including the larger domain, when comparing anomaly correlation and RMS error. Scatter plots comparing observed rainfall anomalies to forecasted

anomalies reveal that the superensemble had very few significant “misses” at any particular grid point over a 20-year period. The highest RMS error in any MAM in any year by the superensemble was 2.59. This occurred during 1988 in Domain A where the mean model RMS error was 3.00. In the same year and domain, all member models had negative anomaly correlations while the modified superensemble forecasted a 0.83 correlation. The magnitude of these improvements can be utilized in practice to pinpoint regions of potential anomalous rainfall.

With this study, we seek to contribute to the fledgling population of statistical downscaling studies over South America. While this study is not conducted in same manner as a traditional statistical downscaling, it does give some insight into alternative methods that may be used to develop more accurate forecasts on a seasonal timescale. Also, as the climate community has begun to lean away from deterministic measures as forecast validation techniques, future work may consider utilizing more probabilistic techniques as well.

Mendes and Marengo (2010) showed that a downscaled temporal neural network outperforms statistical models in daily precipitation forecasts over the Amazon basin for a 30-year period ending in 1999. Valverde Ramírez et al. (2006) compare multiple linear regression (MLR) techniques with artificial neural network (ANN) techniques using data from meteorological stations in southeastern Brazil as opposed to reanalysis. They found that ANN techniques, which are described as “universal approximators” by Hornik et al. (1989), yielded better forecasts than MLR techniques. Krishnamurti et al. (2009) utilized the expansive APHRODITE dataset in a superensemble downscaling project with favorable results over monsoon Asia. Considering the success of these studies, the fusion of ANN downscaling techniques and a high-resolution land-based dataset with superensemble methodology would plausibly yield even higher forecast accuracy. A comparison of such a study with the results presented here would only lead to a better understanding of the options and benefits of downscaling, not only over South America, but also, in varying spatial and temporal scales.

REFERENCES

- Barnett, T. P. and Coauthors, 1994: Forecasting global ENSO-related climate anomalies. *Tellus A*, **46**, 381–397, doi:10.1034/j.1600-0870.1994.t01-3-00005.x.
- Berbet, M. L. C., and M. H. Costa, 2003: Climate change after tropical deforestation: seasonal variability of surface albedo and its effects on precipitation change. *J. Climate*, **16**, 2099–2104.
- Brands, S., J. Gutiérrez, S. Herrera, and A. Cofiño, 2011: On the Use of Reanalysis Data for Downscaling. *J. Climate*, doi:10.1175/JCLI-D-11-00251.1.
- Chakraborty, A., and T. N. Krishnamurti, 2006: Improved seasonal climate forecasts of the South Asian summer monsoon using a suite of 13 coupled ocean-atmosphere models. *Mon. Wea. Rev.*, **134**, 1697–1721.
- Chakraborty, A., and T. N. Krishnamurti, 2009: Improving Global Model Precipitation Forecasts over India Using Downscaling and the FSU Superensemble. Part II: Seasonal Climate. *Mon. Wea. Rev.*, **137**, 2736–2757, doi:10.1175/2009MWR2736.1.
- Chan, S. C., and V. Misra, 2011: Dynamic Downscaling of the North American Monsoon with the NCEP–Scripps Regional Spectral Model from the NCEP CFS Global Model. *J. Climate*, **24**, 653–673, doi:10.1175/2010JCLI3593.1.
- Chou, C., J.-Y. Tu, and J.-Y. Yu, 2003: Interannual Variability of the Western North Pacific Summer Monsoon: Differences between ENSO and Non-ENSO Years. *J. Climate*, **16**, 2275–2287, doi:10.1175/2761.1.
- Delworth, T. L. and Coauthors, 2006: GFDL’s CM2 Global Coupled Climate Models. Part I: Formulation and Simulation Characteristics. *J. Climate*, **19**, 643–674, doi:10.1175/JCLI3629.1.
- Druyan, L. M., M. Fulakeza, and P. Lonergan, 2002: Dynamic Downscaling of Seasonal Climate Predictions over Brazil. *J. Climate*, **15**, 3411–3426, doi:10.1175/1520-0442(2002)015<3411:DDOSCP>2.0.CO;2.
- Fu, R., B. Zhu, and R. E. Dickinson, 1999: How Do Atmosphere and Land Surface Influence Seasonal Changes of Convection in the Tropical Amazon? *J. Climate*, **12**, 1306–1321.
- Fu, R., R. E. Dickinson, M. Chen, and H. Wang, 2001: How Do Tropical Sea Surface Temperatures Influence the Seasonal Distribution of Precipitation in the Equatorial Amazon? *J. Climate*, **14**, 4003–4026, doi:10.1175/1520-0442(2001)014<4003:HDTSST>2.0.CO;2.

- Fu, X., and B. Wang, 2004: Differences of Boreal Summer Intraseasonal Oscillations Simulated in an Atmosphere–Ocean Coupled Model and an Atmosphere-Only Model*. *J. Climate*, **17**, 1263–1271, doi:10.1175/1520-0442(2004)017<1263:DOBSIO>2.0.CO;2.
- Gadgil, A., and A. Dhorde, 2005: Temperature trends in twentieth century at Pune, India. *Atmos. Environment*, **39**, 6550–6556, doi:10.1016/j.atmosenv.2005.07.032.
- Horel, J. D., A. N. Hahmann, and J. E. Geisler, 1989: An investigation of the Annual Cycle of Convective Activity over the Tropical Americas. *J. Climate*, **2**, 1388–1403, doi:10.1175/1520-0442(1989)002<1388:AIOTAC>2.0.CO;2.
- Hornik, K., M. Stinchcombe, and H. White, 1989: Multilayer feedforward networks are universal approximators. *Neural Networks*, **2**, 359–366.
- Jones, C., L. M. V. Carvalho, and B. Liebmann, 2012: Forecast Skill of the South American Monsoon System. *J. Climate*, **25**, 1883–1889, doi:10.1175/JCLI-D-11-00586.1.
- Juang, H.-M. H., and M. Kanamitsu, 1994: The NMC Nested Regional Spectral Model. *Mon. Wea. Rev.*, **122**, 3–26, doi:10.1175/1520-0493(1994)122<0003:TNNRSM>2.0.CO;2.
- Kousky, V. E., 1988: Pentad Outgoing Longwave Radiation Climatology for the South American Sector. *Revista Brasileira de Meteorologia*, **3**, 217–231.
- Krishnamurti, T. N., and Y. Ramanathan, 1982: Sensitivity of the Monsoon Onset to Differential Heating. *J. Atmos. Sci.*, **39**, 1290–1306, doi:10.1175/1520-0469(1982)039<1290:SOTMOT>2.0.CO;2.
- Krishnamurti, T. N., and V. Kumar, 2012: Improved Seasonal Precipitation Forecasts for the Asian Monsoon Using 16 Atmosphere–Ocean Coupled Models. Part II: Anomaly. *J. Climate*, **25**, 65–88, doi:10.1175/2011JCLI4126.1.
- Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble. *Science*, **285**, 1548–1550.
- Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. E. LaRow, D. R. Bachiochi, C. E. Williford, S. Gadgil, and S. Surendran, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196–4216.
- Krishnamurti, T. N., L. Stefanova, A. Chakraborty, T. S. V. V. Kumar, S. Cocke, D. R. Bachiochi, and B. Mackey, 2002: Seasonal Forecasts of Precipitation Anomalies for North American and Asian Monsoons. *J. Meteor. Soc. Japan*, **80**, 1415–1426.

- Krishnamurti, T. N., A. K. Mitra, T. S. V. V. Kumar, W. T. Yun, and W. K. Dewar, 2006: Seasonal climate forecasts of the South Asian monsoon using multiple coupled models. *Tellus A*, **58**, 487–507, doi:10.1111/j.1600-0870.2006.00184.x.
- Krishnamurti, T. N., A. K. Mitra, A. Chakraborty, and M. Rajeevan, 2009: Improving Global Model Precipitation Forecasts over India Using Downscaling and the FSU Superensemble. Part I: 1–5-Day Forecasts. *Mon. Wea. Rev.*, **137**, 2713–2735, doi:10.1175/2009MWR2568.1.
- Kug, J.-S., I.-S. Kang, and D.-H. Choi, 2007: Seasonal climate predictability with Tier-one and Tier-two prediction systems. *Climate Dyn.*, **31**, 403–416, doi:10.1007/s00382-007-0264-7.
- Kumar, V., and T. N. Krishnamurti, 2012: Improved Seasonal Precipitation Forecasts for the Asian Monsoon Using 16 Atmosphere–Ocean Coupled Models. Part I: Climatology. *J. Climate*, **25**, 39–64, doi:10.1175/2011JCLI4125.1.
- Li, W., and R. Fu, 2004: Transition of the large-scale atmospheric and land surface conditions from the dry to the wet season over Amazonia as diagnosed by the ECMWF re-analysis. *J. Climate*, **17**, 2637–2651.
- Luo, J.-J., S. Masson, S. Behera, S. Shingu, and T. Yamagata, 2005: Seasonal Climate Predictability in a Coupled OAGCM Using a Different Approach for Ensemble Forecasts. *J. Climate*, **18**, 4474–4497, doi:10.1175/JCLI3526.1.
- Marengo, J., B. Liebmann, V. E. Kousky, N. P. Filizola, and I. C. Wainer, 2001: Onset and End of the Rainy Season in the Brazilian Amazon Basin. *J. Climate*, **14**, 833–852, doi:10.1175/1520-0442(2001)014<0833:OAEOTR>2.0.CO;2.
- Marengo, J., W. R. Soares, C. Saulo, and M. Nicolini, 2004: Climatology of the Low-Level Jet East of the Andes as Derived from the NCEP–NCAR Reanalyses: Characteristics and Temporal Variability. *J. Climate*, **17**, 2261–2280, doi:10.1175/1520-0442(2004)017<2261:COTLJE>2.0.CO;2.
- Meehl, G. A., and A. Hu, 2006: Megadroughts in the Indian Monsoon Region and Southwest North America and a Mechanism for Associated Multidecadal Pacific Sea Surface Temperature Anomalies. *J. Climate*, **19**, 1605–1623, doi:10.1175/JCLI3675.1.
- Mendes, D., and J. Marengo, 2010: Temporal downscaling: a comparison between artificial neural network and autocorrelation techniques over the Amazon Basin in present and future climate change scenarios. *Theor. Appl. Climatol.*, **100**, 413–421, doi:10.1007/s00704-009-0193-y.

- Misra, V., P. A. Dirmeyer, and B. P. Kirtman, 2003: Dynamic Downscaling of Seasonal Simulations over South America. *J. Climate*, **16**, 103–117, doi:10.1175/1520-0442(2003)016<0103:DDOSSO>2.0.CO;2.
- Nicholas, R. E., and D. S. Battisti, 2012: Empirical Downscaling of High-Resolution Regional Precipitation from Large-Scale Reanalysis Fields. *J. Appl. Meteor. Climatol.*, **51**, 100–114, doi:10.1175/JAMC-D-11-04.1.
- Nobre, P., A. D. Moura, and L. Sun, 2001: Dynamical Downscaling of Seasonal Climate Prediction over Nordeste Brazil with ECHAM3 and NCEP's Regional Spectral Models at IRI. *Bull. Amer. Meteor. Soc.*, **82**, 2787–2796, doi:10.1175/1520-0477(2001)082<2787:DDOSCP>2.3.CO;2.
- Nobre, P., J. Marengo, I. F. A. Cavalcanti, G. Obregon, V. Barros, I. Camilloni, N. Campos, and A. G. Ferreira, 2006: Seasonal-to-Decadal Predictability and Prediction of South American Climate. *J. Climate*, **19**, 5988–6004, doi:10.1175/JCLI3946.1.
- Palmer, T. N. and Coauthors, 2004: Development of a European Multimodel Ensemble System for Seasonal-To-Interannual Prediction (Demeter). *Bull. Amer. Meteor. Soc.*, **85**, 853–872, doi:10.1175/BAMS-85-6-853.
- Philander, S. G., 1990: El Nino, La Nina, and the Southern Oscillation. S. George Philander. Academic Press, San Diego, CA, 1989. x, 293 pp., illus. 59.50. Int. Geophys. Series, vol. 46. *Science*, **248**, 904–905.
- Rienecker, M. M. and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.
- Robertson, A. W., S. Kirshner, and P. Smyth, 2004: Downscaling of Daily Rainfall Occurrence over Northeast Brazil Using a Hidden Markov Model. *J. Climate*, **17**, 4407–4424, doi:10.1175/JCLI-3216.1.
- Saha, S. and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517, doi:10.1175/JCLI3812.1.
- Solman, S. A., and M. N. Nuñez, 1999: Local estimates of global climate change: a statistical downscaling approach. *Int. J. Climatol.*, **19**, 835–861.
- Stefanova, L., and T. N. Krishnamurti, 2002: Interpretation of Seasonal Climate Forecast Using Brier Skill Score, The Florida State University Superensemble, and the AMIP-I Dataset. *J. Climate*, **15**, 537–544, doi:10.1175/1520-0442(2002)015<0537:IOSCFU>2.0.CO;2.

- Valverde Ramírez, M. C., N. J. Ferreira, and H. F. de C. Velho, 2006: Linear and Nonlinear Statistical Downscaling for Rainfall Forecasting over Southeastern Brazil. *Wea. Forecasting*, **21**, 969–989, doi:10.1175/WAF981.1.
- Wang, B. and Coauthors, 2008: Advance and prospectus of seasonal prediction: assessment of the APCC/ClipAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Climate Dyn.*, **33**, 93–117, doi:10.1007/s00382-008-0460-0.
- Wang, G., O. Alves, A. Zhong, N. Smith, A. Schiller, G. Meyers, F. Tseitkin, and S. Godfrey, 2004: POAMA: An Australian ocean–atmosphere model for climate prediction. *Extended Abstracts, 15th Symposium on Global Change and Climate Var.*, Vol. 1 of, Seattle, WA, Amer. Meteorol. Soc.
- Wang, H., and R. Fu, 2002: Cross-equatorial flow and seasonal cycle of precipitation over South America. *J. Climate*, **15**, 1591–1608.
- Wilby, R. L., and T. M. L. Wigley, 1997: Downscaling general circulation model output: a review of methods and limitations. *Prog. in Phys. Geogr.*, **21**, 530–548.
- Williford, C. E., T. N. Krishnamurti, R. C. Torres, S. Cocke, Z. Christidis, and T. S. V. V. Kumar, 2003: Real-Time Multimodel Superensemble Forecasts of Atlantic Tropical Systems of 1999. *Mon. Wea. Rev.*, **131**, 1878–1894, doi:10.1175//2571.1.
- Wolff, J.-O., E. Maier-Reimer, and S. Legutke, 1997: *The Hamburg Ocean Primitive Equation Model: HOPE*.
- Wu, G., and Y. Zhang, 1998: Tibetan Plateau Forcing and the Timing of the Monsoon Onset over South Asia and the South China Sea. *Mon. Wea. Rev.*, **126**, 913–927, doi:10.1175/1520-0493(1998)126<0913:TPFATT>2.0.CO;2.
- Yun, W. T., L. Stefanova, A. K. Mitra, T. S. V. V. Kumar, W. K. Dewar, and T. N. Krishnamurti, 2005: A multi-model superensemble algorithm for seasonal climate prediction using DEMETER forecasts. *Tellus A*, **57**, 280–289, doi:10.1111/j.1600-0870.2005.00131.x.
- Zhong, A., H. H. Hendon, and O. Alves, 2005: Indian Ocean Variability and Its Association with ENSO in a Global Coupled Model. *J. Climate*, **18**, 3634–3649, doi:10.1175/JCLI3493.1.

BIOGRAPHICAL SKETCH

Bradford Johnson was born in October of 1981 in Baton Rouge, Louisiana. He graduated from Istrouma Senior High and Technology Magnet School in 1999. He earned a B.S. degree Meteorology from Iowa State University in the Spring of 2004. He is married to his wife, Lakeisha, and has one daughter, Tahira.