ASSESSMENT OF STATISTICALLY DOWNSCALED CMIP5 SIMULATIONS OF THE NORTH AMERICAN MONSOON SYSTEM

A Thesis

by

MANUEL HERNANDEZ

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Oliver Frauenfeld
Committee Members,	Steven Quiring
	Gerald North
Head of Department,	David Cairns

August 2015

Major Subject: Geography

ABSTRACT

The North American Monsoon System (NAMS) has been projected to undergo a redistribution of precipitation due to enhanced spring convective barriers. Previous studies have utilized coarse-scaled global climate models (GCMs) from the Coupled Model Intercomparison Project (CMIP) Phase 5 to assess the impacts climate change over the southwest United States and northern Mexico. However, GCMs' spatial resolution limits the representation of regional to local scale processes, especially over complex terrain. Thus, the focus of this research is to determine whether statistically downscaled GCMs are viable tools for assessments of the NAMS. First, two reanalysis systems are compared to station observations over the NAMS region using a suite of goodness of fit measures. This evaluation seeks to illustrate the suitability of each product in quantifying the climate characteristics associated with the NAMS over the monsoon-affected region. Second, simulated temperature and precipitation characteristics of downscaled CMIP5 GCMs are assessed against the reanalysis product identified in our previous analysis. The use of downscaled CMIP5 model output has been introduced as a means for evaluating the climate system and their impacts on a regional to local scale. This second objective will demonstrate the improvements, or limitations, in using statistical downscaling to assess the NAMS.

Two highly resolved, gridded reanalyses, the North American Regional Reanalysis (NARR) and the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis (ERA-I) are assessed in their ability to simulate

temperature and precipitation in comparison to in-situ Global Historical Climatology Network (GHCN) observations. Results indicate similar temperature agreement for both reanalyses, while simulated precipitation is better captured in NARR. Precipitation demonstrated varying degrees of agreement, signifying difficulties when simulating precipitation. However, seasonal errors suggest better performance with NARR when compared to ERA-I, thus chosen as the superior reanalysis. Statistically downscaled CMIP5 output was then compared to NARR to explore the improvements, or limitations, in using downscaled climate model output. Various downscaled CMIP5 outputs, including from CNRM-CM5, HadGEM2-CC, and HadGEM2-ES, demonstrated improved representation of temperature and precipitation when compared to their coarsely resolved counterparts. Furthermore, the downscaled temperature and precipitation climatologies reveal better portrayal of the monsoon seasonality, capturing the onset and decay phase of the NAMS. However, all three downscaled models exhibit a warm, wet bias over complex terrain, indicating continued difficulties in reproducing observations. These results confirm enhancements achieved via statistical downscaling, and suggest this statistical approach to improve future climate projections. Thus, statistically downscaled model output serves a viable tool for improved climate projections for water resources and future adaptation.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Oliver Frauenfeld, and my committee members, Dr. Quiring, and Dr. North, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. I also want to extend my gratitude to the Climate Science Lab research group for their assistance and support in programming for this project.

Finally, thanks to my family for all their encouragement throughout my time at Texas A&M University.

NOMENCLATURE

AR	Assessment Report
AZ/NM	Arizona/New Mexico
BCCA	Bias-Corrected and Constructed Analogs
BCSD	Bias-Correction and Spatial Disaggregation
CDF	Cumulative Distribution Function
CMIP	Coupled Model Intercomparison Project
CNRM-CM	Centre National de Recherches Météorologiques Coupled Global Climate Model
CSIRO-Mk	Commonwealth Scientific and Industrial Research Organisation Mark
d	Index of Agreement
DCHP	Downscaled Climate and Hydrology Projections
ECMWF	European Centre for Medium-Range Weather Forecasts
ENSO	El Niño-Southern Oscillation
ERA-I	European Centre for Medium-Range Weather Forecasts Interim Reanalysis
Four Corners	Arizona, New Mexico, Colorado, Utah
GCM	Global Climate Model
GHCN	Global Historical Climatology Network
GHG	Greenhouse Gas

GoC	Gulf of California
GoF	Goodness of Fit
GoM	Gulf of Mexico
GPCC	Global Precipitation Climatology Centre
HadCM	Hadley Centre Coupled Model
HadGEM2-CC	Hadley Centre Global Environmental Model, version 2 – Carbon Cycle
HadGEM2-ES	Hadley Centre Global Environmental Model, version 2 – Earth System
ITCZ	Inter-Tropical Convergence Zone
IPCC	Intergovernmental Panel on Climate Change
JJA	June-July-August
JJAS	June-July-August-September
LLJ	Low-Level Jet
MAE	Mean Absolute Error
MAM	March-April-May
MAPE	Mean Absolute Percent Error
MJO	Madden-Julian Oscillation
NAMS	North American Monsoon System
NCDC	National Climatic Data Center
NCEP	National Centers for Environmental Prediction
NARR	North American Regional Reanalysis

NNR	NCEP-NCAR Reanalysis
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
PDO	Pacific Decadal Oscillation
R	Pearson's product-moment correlation coefficient
R^2	Coefficient of Determination
RCM	Regional Climate Model
RCP	Representative Concentration Pathways
RDAS	Regional Data Assimilation System
RMSE	Root Mean Square Error
SDB	Standard Deviation of Bias
SON	September-October-November
SST	Sea Surface Temperature
VAR	Variation Analysis

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

I.1 Overview

The North American Monsoon System (NAMS) plays an integral role in the livelihoods of communities within the regions of western Mexico and southwestern United States. The regional NAMS phenomenon provides up to 50-70% of the total rainfall within the southwest US and Mexico, feeding into economic and environmental resources including agricultural practices, water management, and fire season variability. As global temperatures continue to rise as a consequence of climate change, many stakeholders and populations depend on reliable projections regarding the future of these resources. Furthermore, profound socioeconomic effects can result from changes to the progression of the seasonal monsoon rains, underscoring the need for an accurate representation of the NAMS in, e.g., climate models. Both regional and global-scale climate models are commonly used to simulate the NAMS and its associated large-scale features. Many of these models in the monsoon region are sensitive to parameterizations such as the convective scheme and boundary layer treatment, which in turn have an effect on continental-scale simulations (Ropeleweski et al., 2005). Furthermore, the steep terrain in many parts of the monsoon-affected region is difficult to represent using dynamical models because of their coarse resolution (Adams and Comrie, 1997; Ropelewski et al., 2005; Vera et al., 2006). These limitations hinder key local processes from being simulated in global climate models (GCMs) including models' coarse

resolution and individual model biases. However, large-scale atmospheric features and their associated teleconnections with oceanic influences are well-captured in GCMs (Polade et al., 2013) and make them ideal for investigating future climate changes.

The NAMS exhibits both inter- and intra-seasonal variability and is affected by sea surface temperature (SST) anomalies, land surface conditions such as soil moisture and antecedent snowpack, the position and strength of tropical convergence zones, water vapor transport, and large-scale circulations (Vera et al., 2006). Seager et al. (2007) have shown a consensus between future model projections of a more arid climate in the southwestern United State and thus a reduction in precipitation within the mean monsoon season. Enhanced spring convective barriers can shift the monsoon season due to a reduction in early seasonal rainfall, followed by increased precipitation late in the monsoon season and a delayed demise (Cook and Seager, 2013). The convective barriers that inhibit convection and cloud formation in the spring and early summer are brought on by remote forcing based on tropospheric stability, and local forcing based on low-level and surface moisture (Seth et al., 2011; Cook and Seager, 2013). Coarse model resolution still remains an issue in simulating NAMS timing and seasonal precipitation distribution (Cook and Seager, 2013).

Monsoon research and applications are key for many stakeholders within this highly complex and multicultural/binational region. Seasonal forecasts of the onset and strength of the monsoon on an intra- and inter-annual basis increases society's ability to cope with climate variability and tailor adaptive practices for their own needs (Ray et al. 2006). Providing plausible predictions of future monsoon changes will allow

stakeholders and decision makers to use climate information to plan effectively within the numerous economic sectors, thereby providing support to vulnerable populations.

I.2 Seasonal Development

The NAMS, also known as the Arizona monsoon and Mexican monsoon, is a large-scale summertime precipitation phenomenon that is characterized by its wind reversals and distinct rainfall over the southwestern United States and western Mexico. The NAMS shares most of the basic characteristics of the Asian monsoon, but not to the same degree of intensity and longevity. The differential heating between the North American landmass and adjacent oceans leads to the associated seasonal wind shifts that switch from dry westerly winds aloft, to moist southeasterly winds. Through these changes in circulation patterns, the inflow of moisture into the southwest US gives way to dramatic increases in convective activity (Crimmins, 2006). Along with the land-sea thermal gradients, the development of a surface low-pressure system is accompanied by the development of an upper-level monsoon anticyclone, thus giving way to the seasonal precipitation (Ropelewski et al., 2005). Furthermore, the seasonal rainfall is coupled with the northward progression of the Inter-Tropical Convergence Zone (ITZC), with most of the total rain falling over oceans rather than land. Precipitation has been associated with large values of deep tropospheric heating and a tropical-type thermodynamic balance between diabatic heating and adiabatic cooling, which is controlled by static stability in the upper-half of the troposphere (Barlow et al., 1997). The NAMS provides precipitation to numerous regions within western Mexico and the

southwest United States, with different areas receiving a high percentage of their annual rainfall from the NAMS. For example, 40% of the annual precipitation in southern Arizona/New Mexico, 50% in western Mexico, and some areas in Mexico receive up to 70% of their annual precipitation during the monsoon season (Ropelewski et al., 2005). Large uplands such as the Colorado Plateau, Sierra Madre Occidental, and mountain ranges through Nevada, southwestern Arizona, and northwestern Sonora, play a key role in monsoonal deep convection and channeling moisture into regions of the southwest (Adams and Comrie, 1997). Furthermore, the peninsular ranges such as on Baja California, and the Sierra Nevada Range, play an important climatological role for the interior deserts, limiting penetration of marine moisture from the Pacific Ocean (Adams and Comrie, 1997). Also, the lower Colorado River valley and surrounding low desert areas play a critical role in the formation of the thermal low during the summer months (Adams and Comrie, 1997). Oceanic regions including the eastern Pacific Ocean, Gulf of California (GoC), and Gulf of Mexico (GoM), are critical moisture sources for the NAMS. Low-level moisture advection is primarily derived from the GoC and eastern Pacific, while mid- to upper-level moisture has been shown to originate from the GoM. Also, moisture recycling via evapotranspiration from the green-up in the Sierra Madre Occidental and anomalous wet conditions in the Great Plains during the early monsoon period can represent moisture sources. Furthermore, the spatiotemporal progression of the monsoon is driven by continental-scale interactions of the Pacific high pressure and Bermuda high-pressure systems (Carleton, 1987; Crimmins, 2006).

Prior to the onset of the NAMS, upper-level flow is predominantly zonal. Furthermore, enhanced low-level jet (LLJ)-related precipitation over the Great Plains is associated with suppressed rainfall over the desert regions of the southwest US and along the east coast (Higgins et al., 1997). The seasonal onset of the monsoon begins in late May/early June, as heavy precipitation spreads northward along the Sierra Madre Occidental. During this time, the zonal flow transitions to a monsoon-like anticyclone, centered near 15°N just to the south of Mexico (Ropelewski et al., 2005). Increased precipitation in northwest Mexico coincides with increased vertical transport of moisture by convection (Douglas et al., 1993). During this time, the subtropical high pressure is too far south, inhibiting any moisture from entering the southwest US. The abrupt change in large amounts of precipitation in the southwest US during the months of July to August is accompanied with the out-of-phase relationship of decreased precipitation over the Great Plains and enhanced rainfall on the east coast (Higgins et al., 1997). The rainfall characteristics over the Great Plains are associated with upper-tropospheric convergence as a result of the strengthened and expanded middle- and upper-tropospheric monsoon high, while the east coast's increased rainfall is due to a weak trough (Higgins et al., 1997). The increased precipitation over the southwest is associated with the evolution of an anticyclone at the jet stream level, a thermally induced trough over the desert regions, the continued displacement of the semi-permanent high-pressure systems, the formation of southerly low-level jets over the GoC, and the formation of the Arizona monsoon boundary (Vera et al., 2006). Changes in the environment's thermodynamic potential for convection during this time, which is

key in understanding the mechanisms behind deep convection in the monsoon region, are primarily driven by increases in low-level moisture since surface temperatures are already high (Barlow et al., 1997). Furthermore, vertical velocity fields have shown ascent over regions of monsoonal and coastal precipitation, as well as over the North American Plateau, with descent occurring over the west coast and central US during this time (Barlow et al., 1997). Also, moisture transport onto the landmass is accomplished by boundary layer flow from the GoC and the middle troposphere via southeasterly flow from the GoM (Vera et al., 2006). During the months of July to August, the NAMS has reached the mature phase and experiences the bulk of its monsoonal rainfall. The anticyclonic feature has strengthened with a strong ridge to the north, which dominates the atmospheric circulation pattern from the Pacific Coast eastward through the GoM (Ropelewski et al., 2005). By mid-September to October, the monsoon has reached its decay phase, with the ridge over the southwest U.S. weakening considerably due to the retreat of the subtropical high. Along with the weakened monsoonal high, the region begins to experience decreases in precipitation and a shift in precipitation events stemming from synoptic-scale frontal systems rather than localized convective instability (Vera et al., 2006). The official National Weather Service (NWS) season is a fixed date-range from June 15th to September 30th.

I.3 Causes of Variability

The NAMS precipitation has continuously varied due to several dynamical mechanisms that have an impact on different time scales (inter-annual and intra-

seasonal). One of the most noted mechanisms behind the variability stems from episodic moist tropical Pacific surge events termed gulf surges, which are channeled from the GoC into the southwest United States as a result of a low-level atmospheric pressure gradient (Adams and Comrie, 1997; Favors and Abatzoglou, 2012; Ropelewski et al., 2006; Zehnder, 2004). The pressure gradient that is driven by the thermal equilibrium between the tropical Pacific and the GoC becomes disrupted, primarily due to a moist air mass at the mouth of the gulf (Adams and Comrie, 1997). Furthermore, the southerly or southeasterly flow is accompanied by a rise in dew point, cooling, and a general increase in low cloud and convective activity. This destabilization of the boundary layer over the low elevation regions of Arizona promotes the development of deep convection (Zehnder, 2004). The GoC, confined between mountain ranges to the west and east, is able to channel moisture into low elevation regions in the desert southwest. The mechanism behind the initiation of a major gulf surge has been hypothesized to stem from the passage of easterly waves or Pacific hurricanes at the mouth of the GoC. Minor surges have been shown to occur from the passage of an easterly wave without the preceding upper-level trough (as seen in major surges), the passage of a mesoscale convective complex passing from the landmass onto the gulf, or backdoor cold fronts approaching from the north/northeast instead of from the typical west/northwest direction (Adams and Comrie, 1997). Gulf surges are generally confined below 700hPa and with associated wind speeds as high as 20 m/s at the 850 hPa level (Zehnder, 2004). Although past research has extensively investigated gulf surge events, the exact dynamics are still unknown. It is generally thought that gulf surges are some type of

coastally trapped disturbance, but a dynamical explanation of gulf surges is still unclear (Newman and Johnson, 2013; Zehnder, 2004).

On a daily basis, the convective maximum closely aligns with the orography of the region, e.g., mountain ranges. The daily maximum varies depending on the time of day, with convection initiated during the morning hours and transported from the GoC to the slopes of the Sierra Madre Occidental. Throughout the midday hours, convection is usually initiated along the western slopes and high ridges of Sierra Madre Occidental, while during the evening/nighttime hours the flow reverses out into the GoC from the Sierra Madre Occidental slopes. On an intra-seasonal basis, the NAMS precipitation is largely influenced by the location of the subtropical ridge which controls convective activity. If there is a southerly shift of the subtropical ridge, locations in the southwest are characterized by a strong drying effect, while a northerly shift is associated with increased moisture flow from the south (Adams and Comrie, 1997). The Madden-Julian Oscillation (MJO), a 30-60 day oscillation associated with velocity potential anomalies in the Pacific, has the ability to influence the NAMS by amplifying easterly waves, which are key in gulf surge initiation (Lorenz and Hartmann, 2006). The relative location of the upper-level monsoon anticyclone in the midlatitudes at the time of the gulf surge affects the response to the surge in the southwest United States (Higgins et al., 2004). Wetter-than-normal conditions occur in the southwest when the ridge axis is located to the east of the region, while drier-than-normal conditions are evident when the ridge axis is located to the west (Vera et al., 2006). The effects of the MJO seem to be region-limited, as positive zonal wind anomalies in the eastern tropical Pacific lead to

above-normal precipitation in northwest Mexico and Arizona (Lorenz and Hartmann, 2006). This connection can be regionally limited because these areas are affected by gulf surges, which is not the case over New Mexico. Furthermore, potential predictability of NAMS precipitation is improved when the El Niño-Southern Oscillation (ENSO) and MJO phases are known, since they are linked to adjustments of the precipitation patterns throughout the NAMS region.

On an inter-annual basis, influences of SSTs anomalies, land surface conditions, position and strength of the tropical convergence zones, water vapor transport, and large-scale circulation have been explored (Adams and Comrie, 1997; Barlow et al., 1998; Ropelewski et al., 2006; Vera et al., 2006). SSTs are an important feature causing differential sensible heating that produces sea level pressure gradients (Barlow et al., 1998). Around 80% of rainfall within the southwest US occurs after SSTs in the GoC reach 28.5°C or higher. Remote SST anomalies in the equatorial Pacific influence the monsoon through an anomalously dry monsoon season related to El Niño conditions, or an anomalously wet monsoon season throughout the core monsoon region in response to La Niña (Vera et al., 2006). This pattern is illustrated due to ENSO's ability to modulate the monsoonal ridge. Furthermore, Pacific SSTs can influence the timing of the monsoon. Pacific wintertime SSTs impact summer monsoon conditions through variations in the Pacific jet. This jet impacts the west coast's precipitation regime during the preceding winter, which in turn affects local land-based sources of memory, timing, and intensity of the monsoon (Higgins and Shi, 2000). Through this mechanism, early-onset monsoon is related to negative North Pacific SST anomalies and positive

subtropical SST anomalies, with the opposite conditions leading to a late-onset monsoon (NOAA/NWS, 2004). Castro et al. (2001) investigated the relationship between NAMS and tropical and north Pacific SSTs, revealing a connection with the monsoonal ridge, monsoonal onset, and the distribution of precipitation. As revealed through observational analysis, the north and tropical Pacific affect regional teleconnection patterns (Castro et al., 2001). The north Pacific SSTs control the northeast/southwest displacement of the monsoon ridge or trough, while ENSO controls north-south displacements (Castro et al., 2001). On a decadal scale, north Pacific SSTs are related to the Pacific Decadal Oscillation (PDO), which impacts precipitation conditions (drought vs. rainy conditions). Furthermore, the effects of remote SSTs on precipitation modulate the location of the subtropical ridge (large-scale circulation). The location of the ridge can inhibit moisture from entering the southwest US if its location is positioned too far south or west. The ridging and troughing that is influenced by SST anomalies gives way to the out-of-phase relationship between the Great Plains and southwest US. If the Great Plains region experiences anomalously wet conditions, the southwest US tends to experience dry conditions, and vice-versa.

Land surface conditions provide a memory that may influence variations in the monsoon (NOAA/NWS, 2004). Different land surface characteristics, such as antecedent snowpack, soil moisture, and seasonal green-up, influence the monsoon system. Previous studies have shown anomalously deep snowpack can be associated with decreased summer rainfall in the southwest US (Notaro and Zarrin, 2011; Vera et al., 2006). The snowpack-rainfall relationship exists due to solar energy being used for

snowmelt rather than atmospheric heating, as well as weakening the land-ocean thermal contrast. Previous modeling studies have shown that the monsoonal anticyclone is sensitive to the amount of soil moisture (NOAA/NWS, 2004). Furthermore, the monsoonal rainfall during the early summer in Mexico is accompanied by the rapid green up in vegetation along the slopes of the Sierra Madre Occidental, impacting surface temperatures and humidity (NOAA/NWS, 2004). The movement of the ITCZ has been shown to influence the precipitation regime of the NAMS. During El Niño events, the ITCZ shifts southward, resulting in an anomalously strong local meridional circulation and reduction in rainfall over Mexico. During La Niña events, the ITCZ shifts northward, resulting in a weakening of the local Hadley circulation, which increases NAMS precipitation (Vera et al., 2006). Furthermore, high SSTs can displace the ITCZ southward off the Mexican coast, while low SSTs weaken the convergence zone and displace it further north (NOAA/NWS, 2004).

An important aspect of the NAMS is the moisture content that is derived from the GoC and GoM. However, a debate still remains on how much each body of water contributes to the total NAMS moisture. Thermal gradients, driven by SSTs in the GoC, influence the formation of lower level jet episodes such as gulf surges. The Great Plains jet, weakened due to the continental-scale circulation features, is pushed into Mexico and the southwest as another source of mid- and upper-level moisture (NOAA/NWS, 2004).

Numerous influences, from remote SSTs to continental-scale circulation patterns, have a range of impacts on the precipitation regime of the NAMS. Profound

socioeconomic impacts, including water management practices, agricultural production, and timing of the fire season, can be affected by monsoonal variability as countless people are directly or indirectly dependent on the seasonal rains. With future climate changes and increases in greenhouse gases, characteristics of the NAMS are expected to change in a warming world, making it imperative to skillfully represent the monsoon variability.

I.4 Modeling Approaches

Representing the NAMS, which relies on correctly simulating intricate coupling of the atmosphere-ocean-land system, remains a challenge for state-of-the-art GCMs, increasing uncertainty in monsoon projections for the 21st century in part because of GCMs' coarse resolution. Many of these state-of-the-art GCMs are fully coupled (sea, ice, land, atmosphere components) and represent past and present climate reasonably well (Geil et al., 2013; Maloney et al., 2013; Seager et al., 2007). Furthermore, GCMs have been extensively used to investigate numerous climate phenomena at different temporal and spatial scales. In modeling the NAMS, efforts continue to integrate correct parameterizations and increased spatial resolution to better capture the key dynamics of the system. Although GCMs have been able to capture the general progression of the NAMS and its associated continental-scale features, limitations such as spatial resolution inhibit regional to local scale projections of precipitation. Many dynamical models are limited by different features of the NAMS, including realistic seasonal evolution and spatial distribution of rainfall due to the orography within the monsoon region. The

interaction of steep orography, diurnal land-ocean circulations, and atmospheric buoyancy are difficult to simulate due to the inadequate deep convection schemes in models (Ropelewski et al., 2006). Furthermore, precipitation rates in the core region of the NAMS have been shown to be sensitive to the choice of convective parameterization and boundary layer treatment, land surface treatment, and SSTs in the GoC (Ropelewski et al., 2006). That, in turn, has an effect on the continental-scale features, which can be sensitive to the precipitation rates in dynamical models. Also, the diurnal cycle and its associated processes and feedbacks are poorly represented in models, stemming from the limited resolution and their representation of high topography (Ropelewski et al., 2006; Vera et al., 2006).

The representation of key topographic features due to coarse spatial resolution is one of the limitations in global models. Regional climate models (RCMs), however, have a better depiction of monsoon precipitation because of a more realistic representation of diurnal cycle convection (Castro et al., 2012). Castro et al. (2012) identified shortcomings within GCMs since two conditions must be met in simulating the NAMS: mesoscale physical processes that lead to precipitation, and reasonable representation of the climatology and inter-annual variability of continental-scale circulation. GCMs are limited in representing the diurnal cycle of convection due to the poor representation of terrain forcing, mesoscale features, land-atmosphere coupling, and parameterized convective rainfall (Castro et al., 2012). Although RCMs provide a better representation of the NAMS, they remain computationally expensive to run due to

their fine spatial resolution and number of ensemble runs. GCMs provide a greater number of runs due to their coarser resolutions.

As part of the fifth phase of the Coupled Model Intercomparison Project (CMIP5), over a dozen international climate modeling groups have examined long-term simulations of the 20th and 21st century using representative concentration pathways (RCPs) for the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5). The Fourth Assessment Report (AR4) of the IPCC used 23 coupled ocean-atmosphere models, based on the CMIP phase 3, which showed that most GCMs were able to produce the monsoonal rain belt and its northward progression, but overestimated seasonal precipitation, especially during the monsoonal retreat in boreal fall (Cavazos and Arriaga-Ramirez, 2012; Lin et al., 2008). Recent focus has been on the CMIP5 GCMs and their ability to simulate the NAMS. Geil et al. (2013) found no significant improvement in the simulation of the magnitude of the mean annual cycle of monthly precipitation, but the timing of seasonal changes has improved, with 27% of CMIP5 models having a zero phase lag versus the CMIP3 models. However, the CMIP5 multimodel mean annual cycle has a wet bias with a late monsoonal decay. On a decadal basis, Langford et al. (2014) found climatological monsoon precipitation in CMIP5 models to be higher than in atmospheric reanalysis products. Higher-resolution models were able to perform better due to improved orographic representation on the west coast of Mexico. Furthermore, Langford et al. (2014) found that the small fraction of variance (20%) accounted for by tropical and North Pacific SSTs limits the predictability of decadal precipitation. They hypothesize that the remainder of the decadal variability is

due to land-atmosphere feedbacks in regions where SST forcing is limited (farther inland) or because of anomalies generated by internal atmospheric variability. Also, they suggest future studies should focus on land-atmosphere feedbacks involving soil moisture and aerosol forcing due to the absence of a connection between low-frequency precipitation and SST anomalies.

Using CMIP5 models, future projections of the southwest US produce a greater drying trend in the twenty-first century and a transition to a more arid climate is already underway (Langford et al., 2014; Seager et al., 2007). Furthermore, Seth et al. (2011) found the redistribution of precipitation from spring to summer can be driven by two mechanisms that lead to enhanced convective barriers. This forcing includes a local mechanism, where reduced evapotranspiration decreases surface moisture availability, and a remote mechanism, where greenhouse gas (GHG)-induced atmospheric warming increases tropospheric stability (Cook and Seager, 2013). Increases in tropospheric stability persist from winter into spring and are reinforced by a reduction in surface moisture conditions, suggesting both mechanisms are important (Seth et al., 2007). In terms of the NAMS, early season precipitation will decrease but is balanced by a change in late season precipitation increases. This decline in early precipitation is caused by increases in tropospheric stability forced by GHG warming in the troposphere (remote mechanism) and modulated by available moisture at the surface (local mechanism) (Cook and Seager, 2013). Once enough surface moisture is able to overcome the enhanced convective barrier, precipitation begins to increase. However, total NAMS precipitation changes are small and largely insignificant (Cook and Seager, 2013). Cook

and Seager (2013) present shortcomings affecting their results due to model biases including a substantial positive precipitation bias, as well as a delay in the monsoon retreat.

I.5 Objectives

The focus of this thesis is to determine whether statistically downscaled GCMs are viable tools for assessments of the NAMS. Coarse-scale GCMs remain a viable tool to investigate present and future climate but are limited due to their orographic representation, limiting local scale analysis and impacting future adaptation practices. Statistical downscaling of GCM simulations represents a way to take advantage of the computational efficiency of GCMs, while still obtaining the finer-resolution output similar to that of RCMs. The purpose of this study is to determine whether this downscaled model output is able to accurately portray historical climate throughout the southwest United States and northern Mexico. This research will quantify the representation of temperature and precipitation characteristics over the monsoon-affected region using CMIP5 downscaled model output for the 20th century.

The specific objectives to address the overarching focus are:

1) To evaluate two reanalysis systems in representing realistic observations over the NAMS region (Chapter II), and
2) to assess simulated temperature and precipitation characteristics of downscaled CMIP5 GCMs against the reanalysis product identified in #1 (Chapter III).

Temporally complete and highly resolved reanalysis products will be compared to in-situ observations in Chapter II to identify a reanalysis best suited for the subsequent model evaluation. The initial reanalysis assessment will illustrate the suitability of each reanalysis in quantifying the climate characteristics over the monsoon-affected region. Reanalysis products have been extensively utilized in previous studies including dynamical processes that cannot be ascertained by in-situ point-observations. Validation of the two reanalyses' performance will also provide confidence for other users who apply these products in their own research. In Chapter III, statistically downscaled CMIP5 output will be assessed based on the better-suited reanalysis product identified in objective 1. Using downscaled CMIP5 model output will demonstrate the improvements, or limitations, in using statistical downscaling to assess temperature and precipitation characteristics related to the NAMS. As the need for future assessments of regional to local scale climate conditions continues to grow, downscaling methodologies for coarsely resolved GCMs have been introduced as a means for evaluating the climate system. This provides key information to, e.g., decision makers and water managers in the southwest US and Mexico to better prepare for future climate change impacts, including mitigation and adaptation practices.

CHAPTER II

ASSESSMENT OF REANALYSIS PRODUCTS AGAINST GHCN OBSERVATIONS

II.1 Overview

Previous studies have utilized reanalysis products to investigate the state of the North American Monsoon System (NAMS), as well as for validation of general circulation models (GCMs). Reanalyses provide spatially and temporally complete output, whereas station observations are limited due to their inconsistencies and limited point observations. We assess the ability of two reanalyses, NARR and ERA-I, in representing air temperature and total precipitation in comparison to in-situ observations (GHCN) over the southwest United States and western Mexico from 1980 to 2010. Results demonstrate similar small errors in air temperature for both reanalysis products, based on a suite of goodness of fit metrics. Reanalysis precipitation yields greater disagreement with observations, with the greatest biases over the four corners region and southwestern Mexico. However, correlation coefficients and index of agreement, on an annual and seasonal basis, reveal NARR to have reduced differences in comparison to ERA-I. Grid cells with best and worst case differences between reanalyses and observations illustrate a better precipitation agreement with NARR in terms of magnitude and variability. Elevation differences are evident between both reanalyses and observed elevations, with topographic variability accounting for 24% and 69% of the variance in seasonal temperature bias for NARR and ERA-I, respectively. This analysis

demonstrates that NARR has smaller errors compared to ERA-I and will therefore be more suitable for the subsequent model evaluation.

II.2 Introduction

Previous studies have extensively used reanalysis products for climate monitoring, as they provide a dynamically consistent estimate of the climate state at different temporal and spatial scales (Dee et al., 2014). Many reanalysis efforts vary in terms of input variables (satellite observations, meteorological stations, aircraft measurements, ship reports, etc.), which are merged with state-of-the-art ocean-atmosphere-sea ice models to create a best estimate of the present atmosphere and land conditions (Burke et al., 2011). Furthermore, reanalysis products, sometimes spanning horizontally across the entire globe and vertically into the stratosphere, are used for numerous applications including driving land surface models and as boundary conditions for regional climate models (e.g., Brunke et al., 2011). The European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis (ERA-I; Dee et al., 2011) and the National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR; Mesinger et al., 2005) are two reanalyses evaluated in this study because they both have high spatial resolution in comparison to other similar products. NARR, an extension of the global low-resolution NCEP-NCAR Reanalysis (NNR), is a unique real-time high-resolution dataset, spanning from 1979 to the present, which uses detailed precipitation observations as one of its assimilated data products (Mesinger et al., 2005). Although near surface temperature is not assimilated into

NARR, it has been shown to produce a better temperature representation than NNR (Mesinger et al., 2005).

NARR uses the NCEP Eta Model (32 km/45 layer resolution) and the Regional Data Assimilation System, a recent version of the Noah land surface model, in addition to numerous observational datasets to provide a much improved analysis of land hydrology and land-atmosphere interactions (Mesinger et al., 2005). The ERA-I represents a global, high-resolution dataset that is spatially and temporally complete (i.e., 1979 to the present), with an improved atmospheric model and assimilation system relative to the predecessor 40-year reanalysis, ERA-40 (Dee et al., 2013). ERA-I is based on a 2006 version of the ECMWF Integrated Forecast Model with a spectral resolution of T255 (~80 km), 60 vertical levels, and a model top at 0.1 hPa (~64 km; Dee et al., 2014). The core component of the ERA-I data assimilation is based on the 12-hourly four-dimensional variation analysis with adaptive estimation of biases in satellite radiance data (Dee et al., 2011). ERA-I precipitation is forecasted separately, simulating stratiform precipitation from the prognostic parameterization of clouds and convective precipitation from the parameterization of convection solely (Bechtold et al., 2004; de Leeuw et al., 2014; Tompkins et al., 2007). Although forecasted separately, both parameterizations are linked by detrainment of convective clouds (de Leeuw et al., 2014). NARR produces precipitation through the assimilation of data from rain gauge observations, the Climate Prediction Center Merged Analysis of Precipitation, and the Parameter-elevation Regressions on Independent Slopes Model using numerous weighting schemes (Dee et al., 2011; Mesinger et al., 2005). Despite being directly

assimilated, precipitation is difficult to simulate within a reanalysis system for it is highly dependent on the model itself (Janowiak et al., 1998).

Temperature and precipitation variables are key in monsoonal processes and are largely associated with the term 'monsoon.' Previous analyses have utilized reanalysis products to investigate dynamical processes related to the NAMS, which impact precipitation characteristics over the southwest U.S. and western Mexico. Station observations only provide surface based data, whereas reanalysis offers a top-down overview of the state of the climate. Reanalyses, including ERA-I and NARR, have been used to investigate historical changes of the NAMS, as well as validating general circulation models to identify biases (Cook and Seager, 2013; Geil et al., 2013; Kumar and Merwade, 2011; Torrez-Alvarez et al., 2014). Furthermore, Lin et al. (2014) have demonstrated the ability of ERA-I to represent climatological spatial patterns and interannual variability in a multi-reanalysis investigation of global monsoons, including the NAMS. However, the study uses an index to investigate the mean monsoon precipitation variations over the Northern Hemisphere.

The goal of this portion of the study is to evaluate the ability of the NARR and ERA-I systems to realistically represent in situ observations. Although biases and errors have been recognized within reanalysis systems, observations also have shortcomings, as they are temporally sparse, spatially limited, and many only represent point-measurements. The goal of this objective is to provide a quantitative assessment of NARR and ERA-I in capturing the variability of station-based observations. The agreement will be assessed using a suite of statistical measures between reanalysis

products and observations. Conducting this comparison for monsoon-relevant surface variables (i.e., temperature and precipitation) within the NAMS region will allow us to assess which reanalysis is better-suited to capturing monsoonal processes.

II.3 Data and Methodology

II.3.1 Study Region

The core NAMS region, following Cook and Seager's (2013) definition, is 112°W-102°W to 18°N-33°N (Figure 2.1). We use this same area as our NAMS core region, as well as the extended region of 125°W-96°W, 15°N-42°N, to encompass a similar region used within the North American Monsoon Experiment (NAME) to investigate intraseasonal variability (Higgins, 2000). This expanded area will thus facilitate the understanding of larger-scale variability that impacts the NAMS.

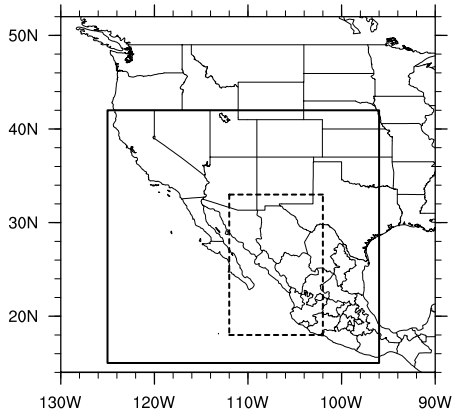


Figure 2.1. Domain of the NAMS core region (black dashed line) as identified by Cook and Seager (2013). The black solid box shows the extended study region also utilized for this portion of the study.

II.3.2 Observational and Reanalysis Data

To validate the reanalysis products, monthly observational temperature and precipitation data were obtained from the National Oceanic and Atmospheric Administration (NOAA)/ National Center for Environmental Information (NCEI) GHCN. GHCN version 2 represents historical precipitation data derived from numerous sources, including up to 20,000 stations globally (Lawrimore et al., 2011). The historical precipitation records underwent various quality control tests including visual inspection of the stations themselves, and quantitative controls including the cumulative sum test, an analogous test that looks for changes in the variance, and determination of station outliers using a variety of nonparametric statistics (Peterson and Vose, 1997). Operational climate monitoring for GHCN temperature has transitioned from version 2 to 3. GHCN version 3, an updated historical temperature record that uses version 2 as its platform, underwent new quality control processes and advanced techniques for removing data inhomogeneities (Menne and Williams, 2009). These new quality controls include checks for spatial inconsistencies, climatological outliers, series duplications, as well as automated pairwise comparisons for homogeneity adjustments (Lawrimore et al., 2011; Menne and Williams, 2009). Here, monthly mean total precipitation amounts are used from the GHCN version 2 for 1980 to 2010, while monthly mean air temperature measurements are obtained from version 3, derived from the daily maximum and minimum temperatures, for 1980 to 2010.

As mentioned in the prior section, ERA-I and NARR data will be compared against these GHCN observations. 2-meter air temperatures were obtained from the

monthly means of daily means from the ERA-Interim dataset, while total precipitation was procured from synoptic monthly means, both at a 0.75° spatial resolution from 1980 to 2010. 2-meter monthly air temperature and monthly average total precipitation at the surface were obtained from the NARR dataset at a 0.3° resolution.

II.3.3 Methodology

Evaluation of the two reanalysis products is based on a suite of statistical methods to investigate the agreement between GHCN observations and reanalysis temperature and precipitation. Because the reanalysis products have different spatial resolutions, they were first regridded to a uniform 1°×1° resolution using a bilinear interpolation scheme, which is a common regridding technique (e.g., Chen and Frauenfeld, 2014). Observations in Mexico are sparse, whereas the U.S. contains a greater amount. GHCN observations were also aggregated to a 1°×1° resolution following a similar analysis as Ma et al. (2009), whereby station observations were averaged within 1°×1° grid cells for comparison with the reanalysis products. Once averaged, grid cells with 20 or more years of data and at least three stations from 1980 to 2010 were retained. Although a higher number of stations per grid cell would be desirable, a higher threshold would severely limit the number of grid cells available for our study.

To quantify the ability of reanalysis products to capture temperature and precipitation variability, different goodness-of-fit (GoF) measures were used. The evaluations were performed on two different temporal scales: annual, based on the calendar year, and seasonal, focusing on the monsoon months of June-July-August-

September (JJAS). The goodness-of-fit measures for this study include Pearson's product-moment correlation coefficient (R), coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), mean absolute percent error (MAPE), standard deviation of bias (SDB), and index of agreement (d). Multiple metrics are used to investigate the agreement between reanalysis and observational products because individual metrics have their own biases and assumptions. Where applicable, statistical significance is assessed using the 90% confidence interval similar to previous assessments of NARR (Fall et al., 2010). Finally, elevation differences between observations and the reanalyses were correlated with annual and seasonal temperature and precipitation biases to explore the dependence of the biases on topography.

The correlation coefficient can be thought of as a dimensionless quantity that provides an absolute measure of the relationship (direction and degree) between two variables (Bulmer, 1979). A coefficient of determination value quantifies the fraction of the total variance of one variable that can be accounted for by the other. The correlation based on these measures assumes two variables to be linear, but their degree of association will be underestimated if it is not (Bulmer, 1979). Another major limitation of correlation-based measures is their greater sensitivity to outlying values than to observations near the mean (Legates and Davis, 1997). Thus, R and R^2 potentially provide a biased estimate of the agreement between the two variables. RMSE and MAE are metrics used to quantify the variations in errors between two sets of forecasts. Lower RMSE and MAE values indicate better agreement, whereas, greater differences in RMSE and MAE are indicative of greater variance in the individual errors. RMSE

values will be larger or equal to MAE because RMSE penalizes large errors while the MAE gives the same weight to all errors (Chai and Draxler, 2014). That, in turn, biases RMSE, as it is not a reliable measure of average error (Willmott and Matsuura, 2005). Furthermore, biases can be either positive or negative, and can cancel each other when summed (Ma et al., 2009). MAE takes the sum of the magnitudes (absolute values) of the errors and divides the total error by the number of observations, capturing the average magnitude error (Willmot and Matsuura, 2005). MAPE is another metric used to capture a forecast error, but can be biased if extreme values are present (zeros or near-zeros). Standard deviation expresses the extent to which a certain variable departs from the mean, as well as measures the random errors due to uncertainty in initial and boundary conditions or observations. SDB is computed both spatially and temporally between the station observations and reanalysis products, as in Ma et al. (2009). Index of agreement (d), developed by Willmott (1981), is a metric that is much less sensitive to the shape of the error-frequency distribution and, as a consequence, to errors concentrated in outliers (Legates and McCabe, 1999; Willmott et al., 2012). Values range from 0.0 to 1.0 with greater values signifying better agreement between reanalysis and observations. A disadvantage of the index of agreement is the assumption of error-free observations when comparing to model-based products (Ji and Gallo, 2006).

II.4 Results

II.4.1 Goodness-of-fit Measurements

Annual		R^2	RMSE	MAE	MAPE	SDB	d
Temperature	NARR	0.82	9.89	8.23	64	2.35	0.58
	ERA-I	0.79	7.90	6.42	50	2.15	0.64
Precipitation	NARR	0.006	38.04	31.48	61	4.62	0.35
	ERA-I	0.001	37.62	30.80	59	4.65	0.42

Seasonal		R^2	RMSE	MAE	MAPE	SDB	d
Temperature	NARR	0.18	2.26	1.74	8	1.21	0.54
	ERA-I	0.13	1.88	1.66	7	0.94	0.51
Precipitation	NARR	0.008	34.01	28.48	55	4.32	0.36
	ERA-I	0.005	30.58	25.72	50	4.07	0.39

Table 2.1. Goodness-of-fit metrics depicting the association of reanalysis temperature and precipitation variables against GHCN observations, averaged over the monsoon domain on an annual and seasonal basis.

Mean GoF measurements over the monsoon domain are shown in Table 2.1. In general, the statistical metrics reveal ERA-I to have a slightly greater agreement relative to NARR for temperature and precipitation on an annual basis. In contrast, GoF seasonal values depict slightly higher errors for ERA-I temperature compared to NARR, while precipitation remains in better agreement with ERA-I. On an annual basis (mean of all grid cells), R^2 shows ERA-I temperature to have a somewhat greater association with observations in comparison to NARR. A weak precipitation relationship is evident between both reanalysis products and GHCN, with NARR having a higher R^2 value. On

a seasonal basis, ERA-I temperature continues to have a higher relationship with observations, but is statistically insignificant, whereas NARR demonstrates a slightly lower significant relationship. However, R^2 values for annual ERA-I and NARR precipitation and seasonal ERA-I temperature and precipitation failed to pass the 90% significance interval, showing no association.

RMSE values indicate similar relationships, with lower errors for ERA-I temperature and precipitation on an annual basis, but greater disagreement for temperature seasonally. NARR RMSE values for temperature illustrate a decrease in error from its annual and seasonal results. MAE demonstrates a similar relationship for temperature and precipitation on an annual basis. MAPE values show that NARR and ERA-I seasonal precipitation (NARR = 180%; ERA-I = 187%), as well as annual temperature (NARR = 109%; ERA-I = 102%), poorly simulate the observations. On a seasonal basis, NARR temperature and precipitation MAPE depict lower errors when compared to ERA-I. Annual and seasonal SDB depicts ERA-I temperature and precipitation to align more closely with GHCN observations.

On an annual basis, index of agreement (d) reveals a higher temperature association between ERA-I and observations, while the opposite is true for precipitation (NARR = 0.42; ERA-I = 0.38). On a seasonal basis, similar index of agreement values are observed for both reanalysis products, with NARR having a higher association for temperature (NARR = 0.42; ERA-I = 0.39). In general, GoF values depict contrasting results for the reanalysis products in comparison to GHCN observations. This analysis was based on grid cells averaged over the entire monsoon domain, which can be skewed

by individual grid cell outliers. To better capture the spatial variability of the relationships, individual grid cell statistics are explored next.

Figure 2.2 shows significant annual grid cell correlations between reanalysis and observed precipitation and temperature. Annual temperature correlations reveal a similar spatial pattern for both NARR and ERA-I: strong agreement over the southwest U.S., indicating good agreement with GHCN in capturing temperature interannual variability (Figure 2.2). In contrast, NARR precipitation illustrates a stronger grid cell correlation with GHCN observations while ERA-I simulates a higher significant negative correlation, demonstrating ERA-I fails to capture the temporal variability in certain grid cells (Figure 2.2). No significant correlations over Mexico demonstrate the lack of observations, or the inability of reanalyses to simulate an accurate signal over the steep terrain in this area.

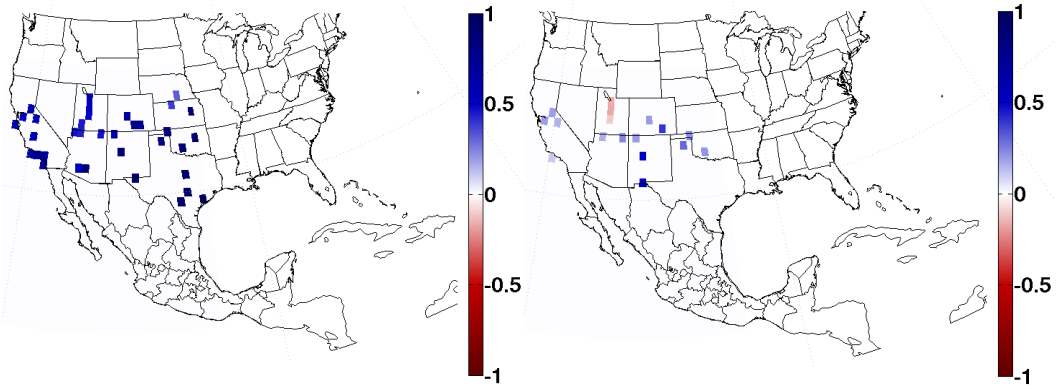


Figure 2.2. Statistically significant annual grid cell correlation coefficients for temperature (left) and precipitation (right) for NARR (a) and ERA-I (b) against GHCN station observations.

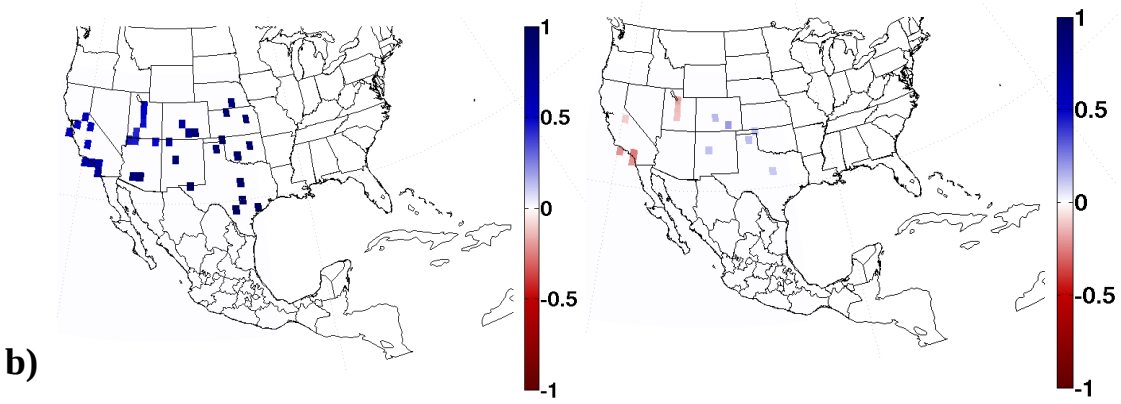


Figure 2.2 Continued. Statistically significant annual grid cell correlation coefficients for temperature (left) and precipitation (right) for NARR (a) and ERA-I (b) against GHCN station observations.

Individual grid cells of annual RMSE values for NARR and ERA-I are shown in Figure 2.3. A similar RMSE temperature spatial pattern is revealed across the southwest U.S. A high disagreement is observed in Utah and Colorado, with certain grid cells exceeding 20°C errors for both products. Many of the grid cells' RMSE values for temperature are within 2-6°C, but are skewed due to the higher disagreement in complex terrain as shown in the domain-averaged RMSE temperature mean in Table 2.1. Spatial RMSE patterns for precipitation are similar in both reanalyses. However, NARR precipitation RMSE (mm/month) shows greater errors over Colorado in comparison to ERA-I. In general, precipitation errors in both products are high, ranging between 75 to 100 mm/month.

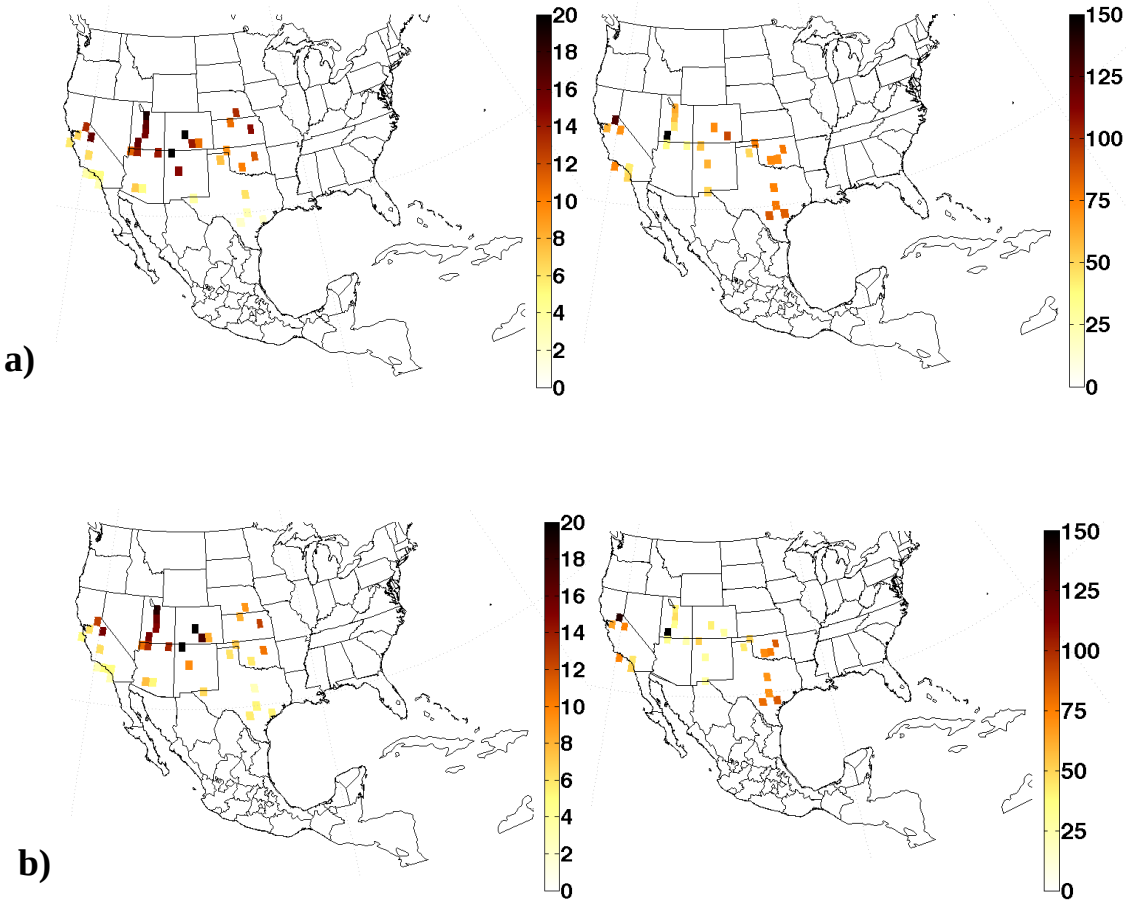


Figure 2.3. Annual temperature (left; °C) and precipitation (right; mm/month) RMSE for NARR (a) and ERA-I (b).

MAE reveals a similar pattern of errors as RMSE, including high temperature errors throughout the four corners region (not shown, Appendix 1). Precipitation MAE demonstrates a similar error over the Utah region for NARR (Figure 2.3), as well a reduced error in ERA-I within that same region (Appendix 1). MAE results further support our initial findings for the precipitation RMSE, showing observed precipitation to be a difficult variable to accurately capture in both reanalyses. To further quantify the distribution of errors, temperature and precipitation MAPE were assessed (Figure 2.4).

The general distribution of temperature MAPE again identifies high values over the four corners region, exceeding 100% in both products. Precipitation error in NARR exceeds 100% over the four corners region, while the error is reduced in ERA-I.

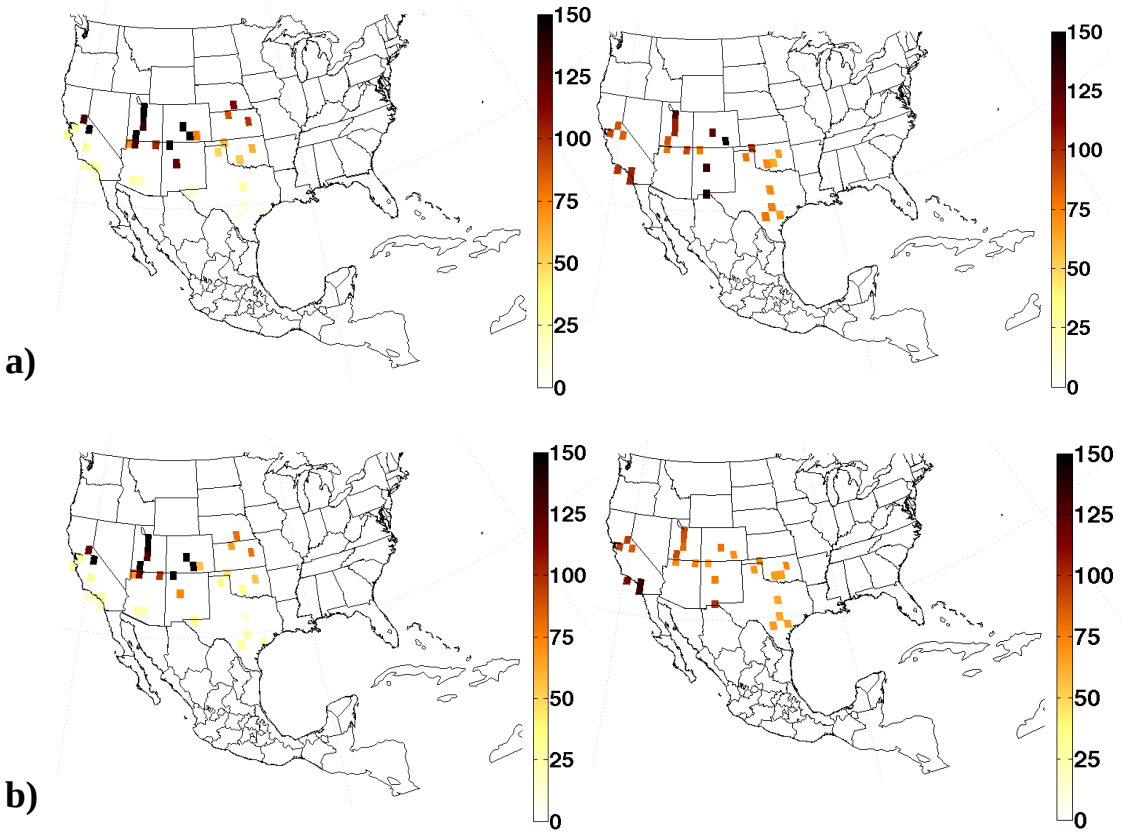
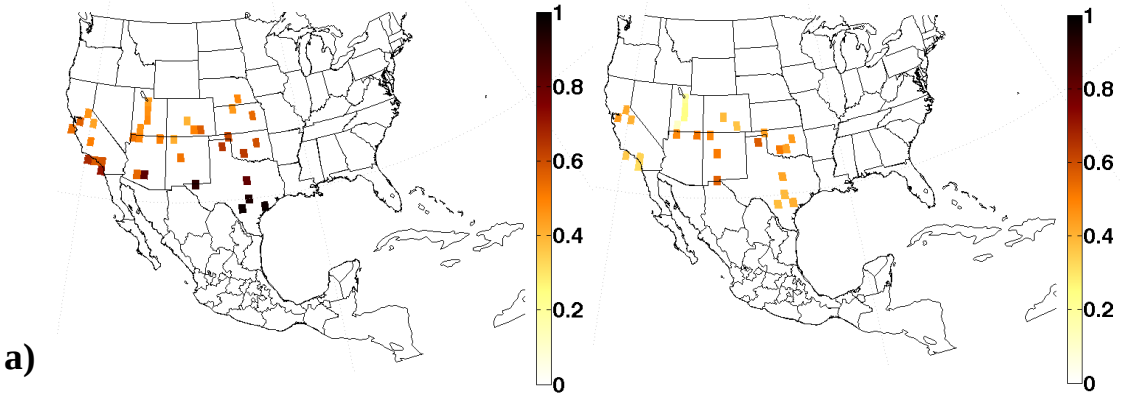


Figure 2.4. Annual temperature (left; °C) and precipitation (right; mm/month) MAPE for NARR (a) and ERA-I (b).

SDB plots for temperature and precipitation show a similar bias pattern in both reanalysis products when compared to previous analyses (Appendix 1). Spatial patterns of temperature illustrate a large bias in the four corners region, extending into Oklahoma

and Kansas. Precipitation patterns demonstrate better performance in ERA-I over the four corners region when compared to NARR.

Annual index of agreement (d) for temperature shows a similar spatial agreement with GHCN observations for both products (Figure 2.5). A greater agreement is demonstrated throughout Texas and the Arizona/Mexico border, with reduced agreement northward. In contrast, NARR precipitation demonstrates a greater association with observations in comparison to ERA-I. Both products simulate similar maximum and minimum values (max = 0.75, min = 0) but show different spatial agreements. No relationship was exhibited over the complex terrain of western Mexico, illustrating the limitations with either the in-situ observations or the simulated climate characteristics in the reanalyses.



a)

Figure 2.5. Annual temperature (left; °C) and precipitation (right; mm/month) index of agreement for NARR (a) and ERA-I (b).

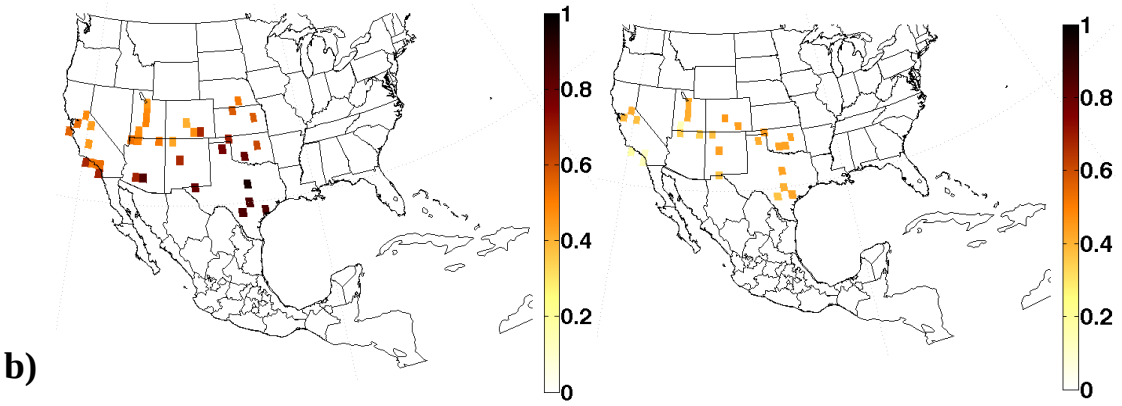


Figure 2.5 Continued. Annual temperature (left; °C) and precipitation (right; mm/month) index of agreement for NARR (a) and ERA-I (b).

Spatially, the seasonal results show lower agreement in contrast to annual temperature GoF measures such as correlation coefficients and index of agreement. Similarly to annual temperature correlations, seasonal temperature R-values illustrate a strong positive correlation over Texas, but weaken over the four corners region (Figure 2.6). This result illustrates that the reanalyses tend to capture the annual temperatures more closely than during the summer months. Furthermore, weaker correlations over the four regions highlight the difficulties of accurately capturing climate conditions over complex terrain. However, as both products do not assimilate temperature, our correlation results provide confidence in the performance of the reanlyses. Seasonal NARR precipitation patterns demonstrate a positive correlation over the four corners region, whereas ERA-I illustrates both weak positive and negative correlations. Precipitation correlations demonstrate the improved representation in NARR as it assimilates precipitation from numerous sources, while ERA-I is model dependent.

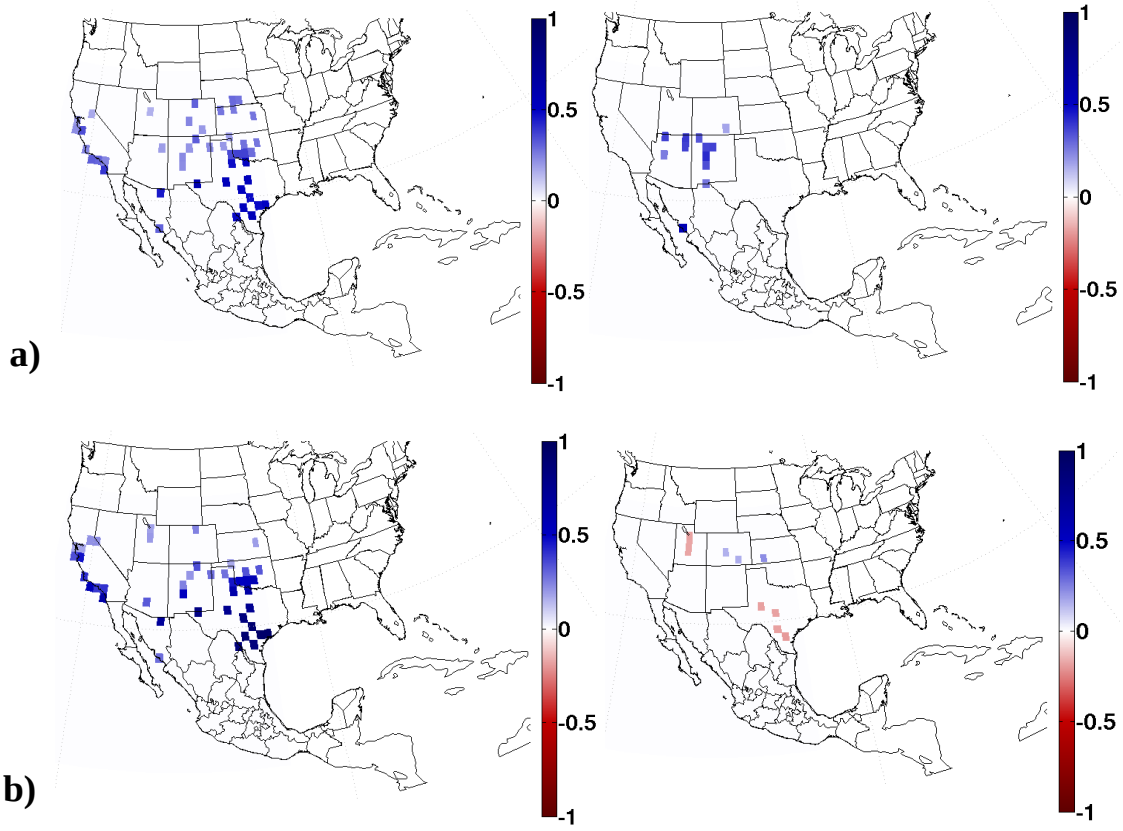


Figure 2.6. Statistically significant seasonal grid cell correlation coefficients for temperature (left) and precipitation (right) for NARR (a) and ERA-I (b) with GHCN station observations.

Seasonal temperature RMSE at the grid cell level shows a better agreement between reanalyses and observations in comparison to annual RMSE values, especially over the four corners region (Figure 2.7). Low errors are seen throughout the southwest; however, errors in NARR are slightly smaller in comparison to ERA-I. In contrast, NARR precipitation RMSE at the grid cell level demonstrates greater disagreement over the four corners region in comparison to ERA-I. Similar spatial disagreements over Mexico are evident in both reanalysis. Previous analysis has identified ERA-I as an

adequate product for reproducing the seasonal precipitation spatial distribution for global monsoons, but has been shown to do poorly over North America. This can be due to the zonal land-ocean and hemispheric thermal contrast (Lin et al., 2014).

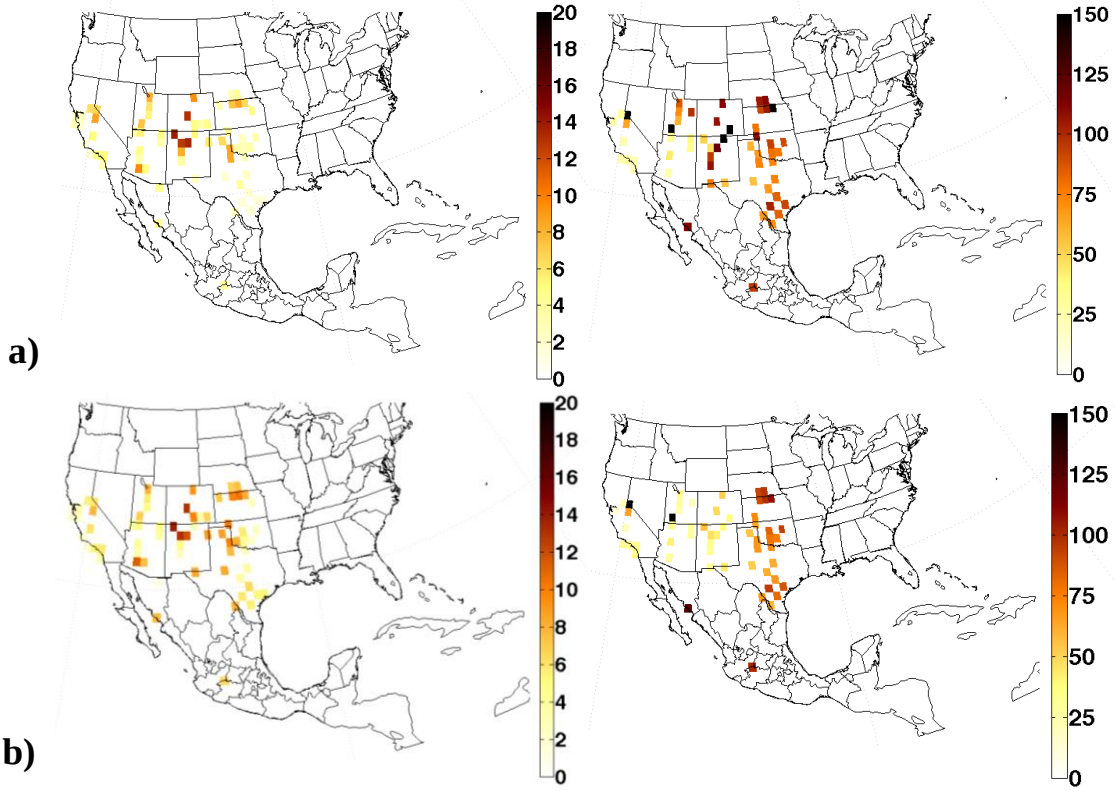


Figure 2.7. Seasonal temperature (left; °C) and precipitation (right; mm/month) RMSE for NARR (a) and ERA-I (b).

Similarly to seasonal temperature and precipitation RMSE, seasonal MAE spatial patterns depict a large error over the four corners region, with NARR precipitation simulating greater errors than ERA-I (Appendix 1). Both reanalyses demonstrate a large error over California for precipitation; an error only slightly captured by RMSE.

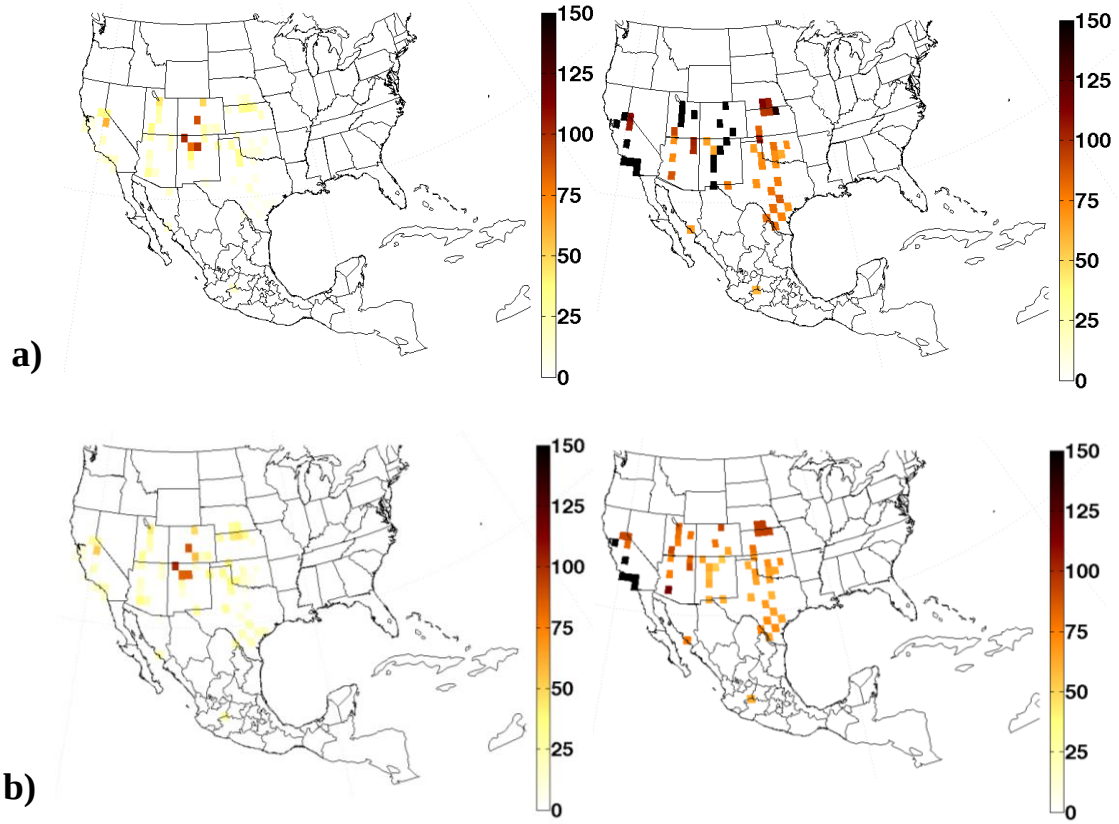


Figure 2.8. Seasonal temperature (left; °C) and precipitation (right; mm/month) MAPE for NARR (a) and ERA-I (b) in percent.

MAPE for seasonal temperatures illustrate a lower error over the southwest U.S. in NARR and ERA-I; however, they both exceed 60% error over Colorado (Figure 2.8). Seasonal precipitation also demonstrates a large region of error exceeding 100% over Colorado, while ERA-I hovers around ~75% error. In both reanalyses, precipitation exceeds 100% error over the west coast, demonstrating difficulties in accurately simulating precipitation. Also, precipitation SDB illustrates a similar bias over the Colorado region, while spatial temperature patterns show a reduced seasonal bias in comparison to annual plots (Appendix 1). Seasonal temperature index of agreement

illustrates a decreased error over the study region, but still shows NARR to have better agreement than ERA-I (e.g., Texas and four corners; Figure 2.9). Similarly, NARR precipitation index of agreement illustrates better agreement with GHCN observations over the four corners region and the Sierra Madre Occidental. However, the general variability in the index of agreement is similar in both products.

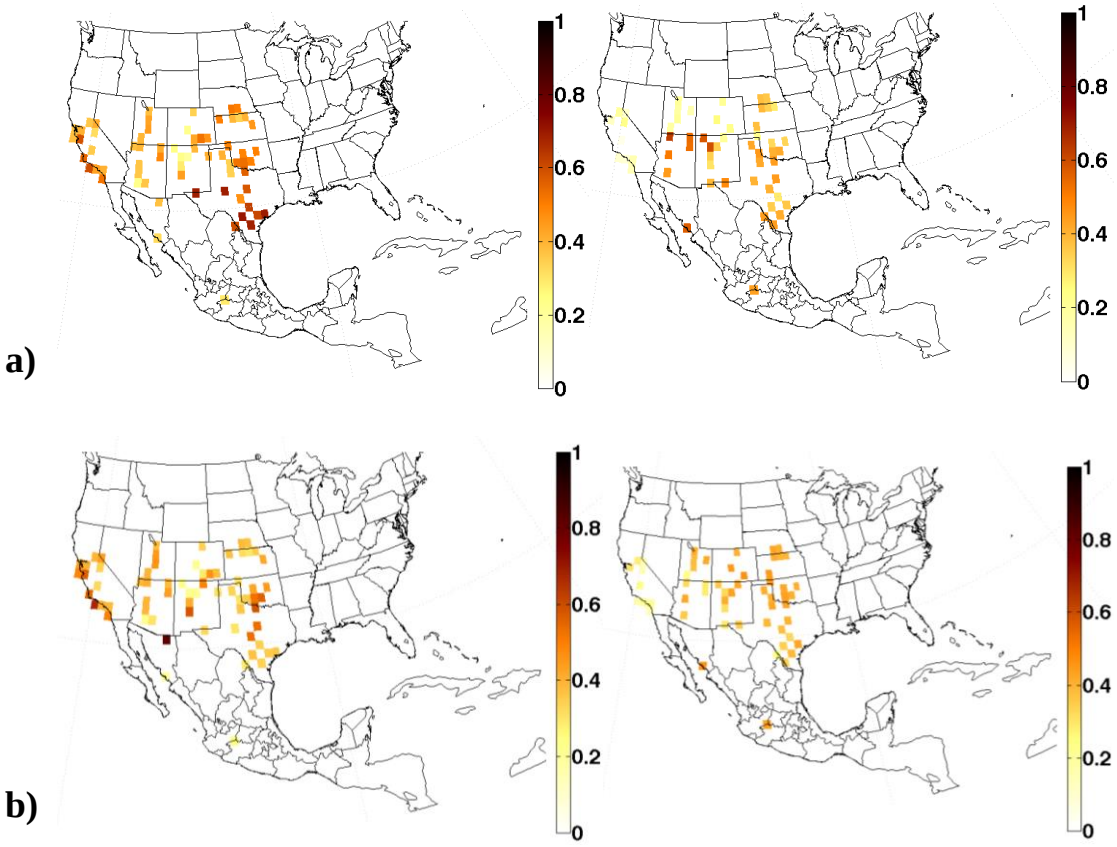


Figure 2.9. Seasonal index of agreement (d values) for temperature (left) and precipitation (right) for NARR (a) and ERA-I (b).

II.4.2 Temporal Variability

To further assess the spatiotemporal capabilities of the reanalysis products in simulating observations at the grid cell level, best and worst-case scenario grid cell time

series, using RMSE results as a determinant, were explored on an annual and seasonal basis. Previous statistical analyses on model evaluations have used RMSE to quantify model error (Geil et al., 2013), motivating our choice of this measure as a means of selecting best and worst case grid cells. Figure 2.10 shows the best and worse case scenario annual grid cell temperature for reanalysis and observations from 1980 to 2010. For temperature, both products capture the interannual variability but tend to underestimate the observed temperature. NARR's best case illustrates a temperature bias of up to 4°C, as seen in 1998, while ERA-I depicts a bias of about 1.4°C (Figure 2.10). The worst-case grid cell for ERA-I overestimates temperature by up to ~20°C, while NARR's worst-case overestimates temperature by ~23°C. In both worst-case scenarios, there is a considerable underestimation when compared to observed temperatures. However, annual ERA-I seems to show better agreement with observed temperatures when comparing "best" and "worst" case scenarios.

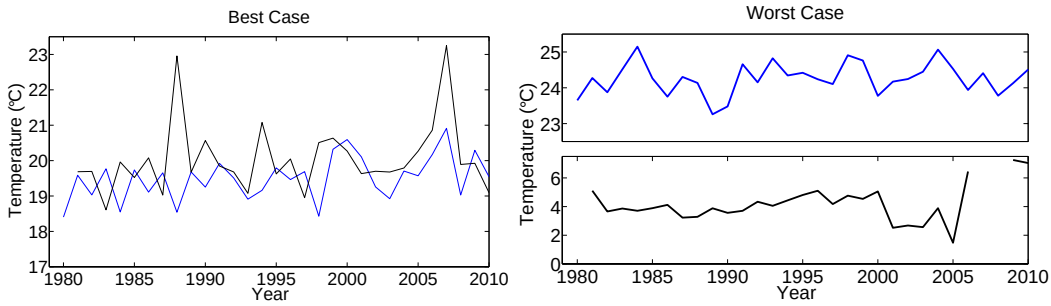


Figure 2.10. Best (left) and worst (right) case annual temperature grid cell time series for NARR (a; blue), ERA-I (b; red), and GHCN (black) using RMSE as a determinant.

b)

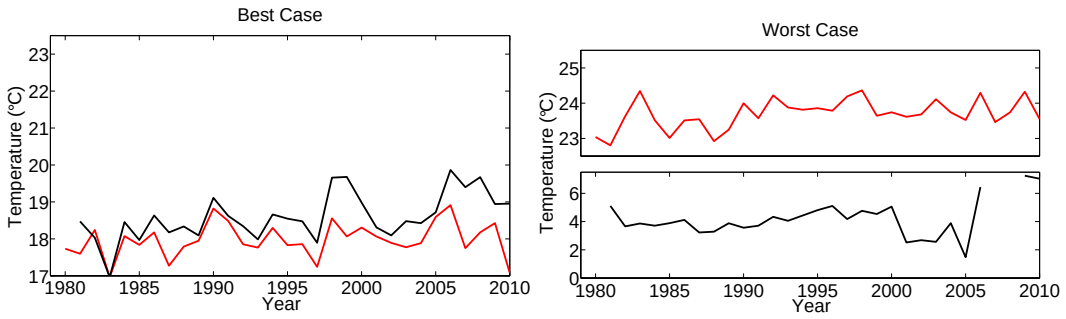


Figure 2.10 Continued. Best (left) and worst (right) case annual temperature grid cell time series for NARR (a; blue), ERA-I (b; red), and GHCN (black) using RMSE as a determinant.

In comparison, the best-case grid cell partially captures observed precipitation variability, as seen in figure 2.11. NARR's best-case scenario simulates the observed magnitude, however, a bias of up to 230 mm/month is exhibited. ERA-I's best case illustrates an underestimation in the earlier time period (~143 mm/month), while an increase in variability is shown at the end of 1990. For the worst-case scenario, NARR precipitation cannot capture the interannual variability, however, the observed magnitude is somewhat captured. The ERA-I worst-case scenario neither matches the interannual variability in the observations, nor the magnitude, with a bias of up to 640 mm/month.

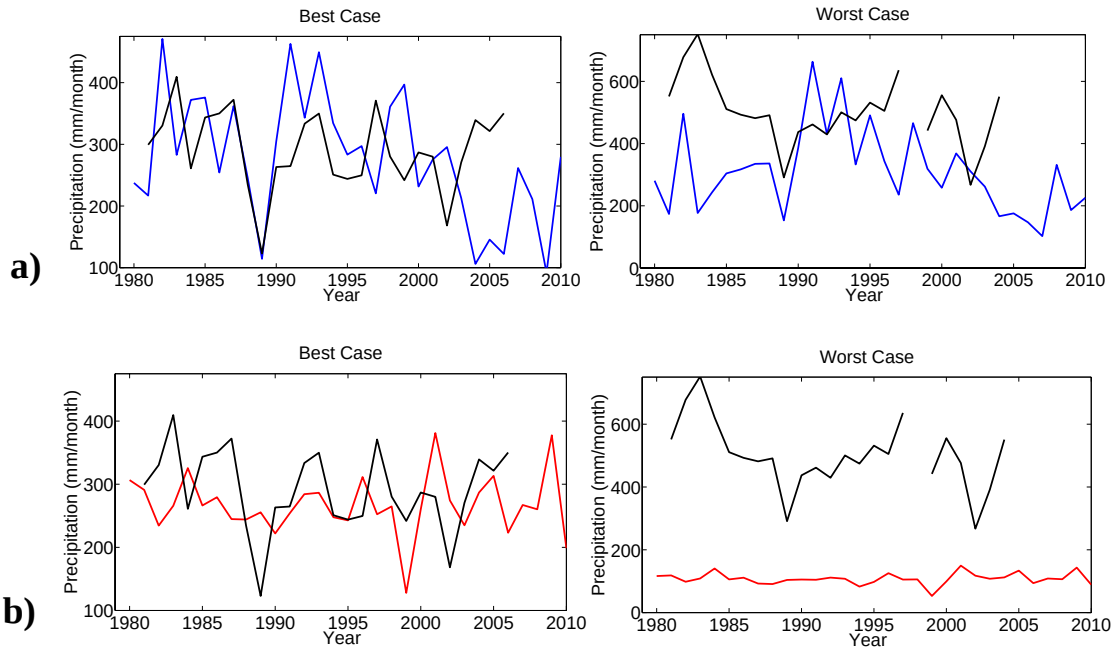


Figure 2.11. Best (left) and worst (right) case annual precipitation grid cell time series for NARR (a; blue), ERA-I (b; red), and GHCN (black) using RMSE as a determinant.

A similar analysis was conducted using Pearson's correlation coefficient as the determinant for best and worst-case grid cells (Appendix 2). Temperature results demonstrate ERA-I to have a temperature bias of 4 to 5°C for its best-case scenario. On the other hand, NARR's correlation-based best-case grid cell was identical to RMSE-based best-case grid cell, with a maximum temperature bias just under 4°C. Both products capture some of the temperature variability, however, NARR is able to fairly reproduce the observed magnitude when compared to ERA-I. Precipitation best case results demonstrate both products somewhat capture observed interannual variability, however, NARR is able to simulate the observed magnitude better than ERA-I.

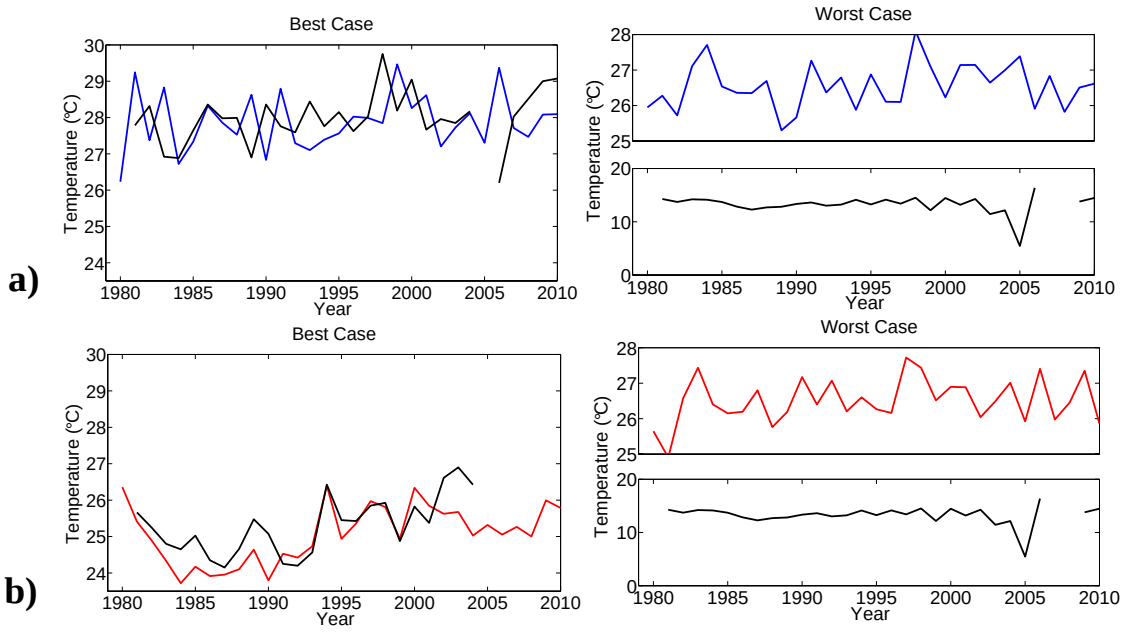


Figure 2.12. Best (left) and worst (right) case seasonal temperature grid cell time series for NARR (a; blue), ERA-I (b; red), and GHCN (black) using RMSE as a determinant.

Seasonal time series for the best and worst-case grid cells, using RMSE as a determinant, for NARR and ERA-I (blue and red) and GHCN (black) from 1980-2010 are shown in figure 2.12. On a seasonal basis, both reanalyses reproduce the temperature variability throughout the 31-year time period. However, NARR simulates a lag when compared to observations, especially around 1990, while ERA-I can reproduce the observed magnitude. Both products simulate similar temperature magnitudes and interannual variability in their worst-case, with a temperature overestimation by at least 10°C.

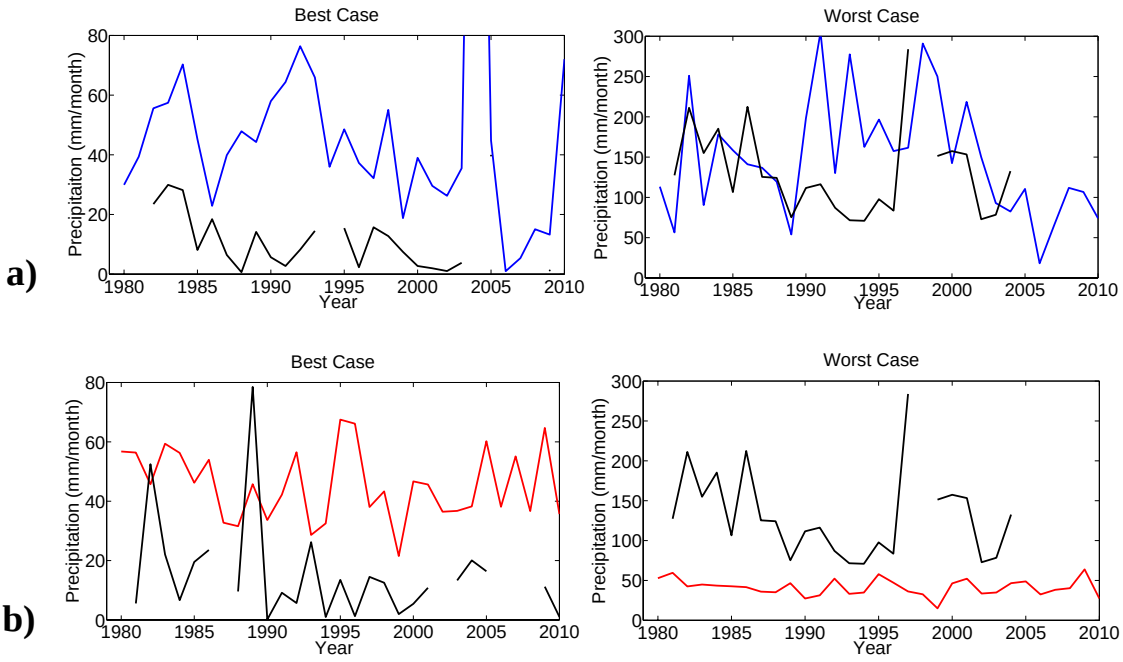


Figure 2.13. Best (left) and worst (right) case seasonal precipitation grid cell time series for NARR (a) and ERA-I (b) using RMSE as a determinant.

Precipitation, although highly variable, is better represented in NARR in comparison to ERA-I (Figure 2.13). NARR seasonal precipitation captures the variability in observations; however, both products overestimate the observed precipitation. Worst grid cell time series shows NARR overestimating precipitation but with a decreased bias when compared to ERA-I, which misses much of the seasonal variability in observations.

Best-case temperature grid cells, using correlation coefficients as the determinant, for both reanalyses demonstrates a colder bias when compared to observations. However, NARR illustrates a smaller bias by about ~3°C in comparison to the ~5°C from ERA-I (Appendix 1). However, worst-case grid cells demonstrate

increased temperatures in NARR and decreased temperatures in ERA-I when compared to observations. Best-case precipitation grid cells illustrate both products to underestimate precipitation (Appendix 1).

II.4.3 Elevation Comparisons

Finally, simulated temperature and precipitation biases were quantified in relation to elevation biases to examine the role elevation plays in reanalyses' shortcomings. Elevation biases were calculated for each reanalysis based on station elevation averages within each grid cell. There is a greater disagreement over steep terrain, e.g., over Colorado and New Mexico, but the bias decreases towards the western states. Both reanalyses show similar results, with slight grid cell elevation differences. Correlation coefficients between elevation differences and temperature and precipitation biases on an annual and seasonal basis are shown on Table 2.2. Results illustrate a significant negative relationship between NARR's temperature bias and elevation differences, both on an annual and seasonal basis. Thus, NARR tends to underestimate temperature as elevation differences increase. Furthermore, results show that 21% of the variance in annual temperature error, as well as a 24% variance on a seasonal basis, can be attributed to elevation. Land surface temperatures are not directly assimilated into NARR because of the 3DVAR scheme (Chuang and Manikin, 2001), making temperature dependent on NARR's model configuration. In contrast, ERA-I demonstrates a significant positive correlation between elevation differences and seasonal temperature biases, showing an increase in elevation error can lead to an increase in temperature bias. Also, results illustrate that 69% of the variance in

temperature bias is explained by the elevation difference. Although decreasing the elevation bias will decrease the temperature error, both reanalyses' shortcomings cannot be fully explained by elevation biases.

Annual		R	P-Value	R^2
NARR	Temperature	0.45	<0.1	0.21
	Precipitation	0.18	0.35	0.03
ERA-I	Temperature	0.29	0.25	0.08
	Precipitation	0.11	0.61	0.01

Seasonal		R	P-Value	R^2
NARR	Temperature	0.49	< 0.1	0.24
	Precipitation	0.28	0.14	0.08
ERA-I	Temperature	0.83	< 0.1	0.69
	Precipitation	0.30	0.17	0.09

Table 2.2. Correlation coefficients between elevation differences and temperature and precipitation biases, on an annual and seasonal basis.

II.5 Discussion and Conclusions

This objective examined the ability of two reanalysis products—NARR and ERA-I—in simulating GHCN observed precipitation and temperature from 1980 to 2010 using different statistical metrics and time series. The results indicate:

1. Based on mean goodness-of-fit measurements, ERA-I and NARR temperature show similar relatively low errors when compared with in-situ observations. ERA-I annual precipitation shows a slightly better association with observations, but errors demonstrate the inability of reanalyses to accurately capture

precipitation amount. On a seasonal basis, NARR exhibits better agreement with temperature in contrast to ERA-I, whereas precipitation errors gave no clear indication of the better reanalysis.

2. Individual grid cell statistics show a similar spatial temperature agreement in both reanalyses, with large forecast errors over the four corners. Seasonally, NARR temperature demonstrates better performance over ERA-I, based on RMSE and index of agreement.
3. Annual precipitation illustrates similar spatial errors between NARR and ERA-I. However, certain metrics, such as RMSE and MAPE, demonstrate a greater agreement between ERA-I and observed precipitation, while correlation coefficient and index of agreement illustrates NARR to have better agreement with observations.
4. Temperature time series for the best-case grid cell time series, on an annual and seasonal basis, show both reanalyses capture the interannual and seasonal variability. However, both products tend to underestimate temperature. Precipitation time series for the best-case grid cell, both on an annual and seasonal basis, show NARR to generally better capture the interannual and seasonal variability in comparison to ERA-I.
5. Worst-case grid cells illustrate a large bias in both precipitation and temperature.
6. Elevation differences demonstrate an elevation gradient going from east to west. NARR and ERA-I exhibiting higher elevations in the Rocky Mountains and

lower elevations in the southern Great Plains and west coast in comparison to station based observations.

7. Furthermore, a significant correlation was found between temperature errors in NARR (annual and seasonal) and ERA-I (seasonal) and elevation biases. Also, elevation biases account for some of the variance in the temperature biases in NARR and ERA-I. Thus, decreasing the elevation error can have an impact on simulated temperature.

This study can be used to inform which reanalysis product may be best to use for research focused on NAMS variability. It is also used here to select the more suitable reanalysis to represent spatially and temporally complete observations for our subsequent model comparison. Based on the mean statistics, ERA-I is best to use for temperature and precipitation variability. However, grid cell outliers significantly impact the mean statistics presented in Table 2.1. Although individual grid cells also demonstrate similar temperature errors for both reanalyses, NARR is also able to simulate interannual and interseasonal variability, based on grid-cell time series results. Both products can thus resolve temperature fairly well. Precipitation, on the other hand, is a variable that many reanalysis and model products cannot resolve accurately due to, e.g., spatial resolution issues and precipitation parameterizations. Similar to temperature, mean precipitation metrics show that ERA-I has a better agreement with GHCN observations. Grid cell statistics depict NARR precipitation to have a stronger relationship with observations in comparison to ERA-I, especially based on index of agreement and correlation coefficients. The influence of grid cell outliers thus skewed

the domain-averaged annual/seasonal metric means. Furthermore, grid-cell time series showed NARR's ability to capture the interannual and interseasonal precipitation variability when compared to observations. ERA-I precipitation captures the general interannual variability in its best-case grid cell, but fails to capture the magnitude. Area-averaged GoF statistics show little differences between the two reanalyses, demonstrating the need to evaluate the agreement at the grid-cell level, rather than for regional averages. NARR's ability to better simulate precipitation likely stems from its assimilation of numerous precipitation products, and its high spatial resolution over North America. Thus, based on results presented in our study, NARR is better able to simulate precipitation and temperature over the North American Monsoon region.

CHAPTER III

ASSESSMENT OF STATISTICALLY DOWNSCALED CMIP5 MODELS

III.1 Overview

Projections of the North American Monsoon System (NAMS) have utilized the Coupled Model Intercomparison Project (CMIP) Phase 5 to assess the impacts of climate change over the southwestern United States and Mexico. However, the global climate models in CMIP5 have limitations due to their sensitivity to their respective convective schemes and orographic representations. Different methodologies have been introduced to counter the spatiotemporal biases within coarse-scaled models, including dynamical and statistical downscaling. This study focuses on statistical downscaling, and assesses the ability of this downscaled CMIP5 output to accurately capture NAMS processes. Temperature and precipitation output from 1979 to 1999 is evaluated against the North American Regional Reanalysis (NARR) over both the core monsoon region (112°W - 102°W to 25°N-33°N), and also an extended region (125°W - 95°W to 25°N - 40°N) that better captures regional-scale impacts that could be related to the NAMS. Statistically downscaled temperature shows good agreement with NARR but simulates a warm bias over the Sierra Madre Occidental during the monsoon season (JJAS). Similarly, downscaled model output reveals a precipitation bias over a similar region, potentially due to simulated orographic features important to the NAMS. Skill scores reveal the weakest temperature agreement between the downscaled model output and NARR to be during the monsoon season, however, the pre-monsoon season illustrates

the greatest agreement for precipitation. In general, statistically downscaled CMIP5 output better captures the NAMS spatiotemporal variability when compared to the original coarse global climate models output, and is therefore effective in projecting future climate over the monsoon region.

III.2 Introduction

Climate models are a key tool in investigating earth system processes and their variability through different climate forcings. Global climate models (GCMs) have been extensively utilized to evaluate climate projections at various times ranging from future to historical and paleo, including by the Intergovernmental Panel of Climate Change (IPCC) through the CMIP. Evaluation of climate models between the Fourth (CMIP3) and Fifth (CMIP5) Assessment Reports (AR) has shown continued development and improvements through the inclusion of biogeochemical cycles important in the climate system (Flato et al., 2013). Furthermore, the ability of climate models to simulate variables such as surface temperature has improved in many important aspects relative to AR4, but remains a challenge with slight improvements for variables such as large-scale precipitation patterns (Flato et al., 2013). Some of the systematic errors can be attributed to model resolution and the depiction of complex topography. Coarse resolution issues and their impacts on regional to local scales can be addressed through the application of regional climate downscaling. Statistical downscaling involves deriving empirical relationships linking large-scale atmospheric variables (predictors) and local/regional climate variables (predictands), which can then be applied to equivalent predictors from

climate models (Flato et al., 2013). Downscaling methods, while having shown improvements due to their output at a higher resolution, vary in terms of location, time frame, and methodology, making them difficult to evaluate. Furthermore, the reliability of downscaled output depends on the quality of the downscaling method and the large-scale boundary conditions, which are provided by coarse GCMs (Flato et al., 2013). This, in turn, questions the ability of downscaled models to accurately represent historical observations, and provide credible climate projections.

The NAMS has exhibited no distinct precipitation trends in the last half of the 20th century; however, decreasing monsoonal streamflow in western Mexico has nonetheless been reported due to changing precipitation or antecedent hydrological conditions (Anderson et al., 2010; Arriaga-Ramirez and Cavazos, 2010; Christensen et al., 2013; Gochis et al., 2007). Furthermore, a systematic delay in monsoon onset, peak, and decay has been shown (Grantz et al., 2007). GCMs have demonstrated difficulties in simulating key NAMS features, but do capture some features of the seasonal cycle. Geil et al. (2013) found no improvements in the magnitude of the mean annual cycle of monthly precipitation over the NAMS core region between CMIP3 and CMIP5. Furthermore, the multi-model mean annual cycle has exhibited a wet bias and fails to capture the retreat of the NAMS accurately (Geil et al., 2013). However, a 27% improvement in the timing of seasonal precipitation changes is evident in CMIP5 models in comparison to CMIP3, as well as adequate monsoon representation during the early and mid-season period, albeit with a large-scale circulation pattern bias (Geil et al., 2013). An additional significant issue is the inability of some GCMs to even simulate a

monsoon signal. This results in inconsistencies in future monsoon projections in the CMIP5 models, showing a general precipitation decrease in the core monsoon region but with little homogeneity. Also, CMIP5 models produce no consensus for future changes in monsoon timing (Christensen et al., 2013). Cook and Seager (2013) have shown small changes in future total monsoon precipitation but rather significant declines in early monsoon precipitation (June - July) and increases in late monsoon season precipitation (September - October).

The Downscaled Climate and Hydrology Projections (DCHP) represent a unique collaborative effort to provide statistically downscaled output for CMIP3 and CMIP5 models. Monthly bias-correction and spatial disaggregation (BCSD) and daily bias-corrected and constructed analogs (BCCA) techniques are applied to historical and future climate projections from GCM output in the CMIP3 and CMIP5 archive. DCHP recognized water managers' needs to assess the implications of future climate change for the management of their systems, as well as when climate change vulnerabilities and impacts may exceed thresholds, triggering the need for adaptive intervention (Reclamation, 2013). Furthermore, downscaled projections have been utilized because of GCMs' inability to provide an adequate local-scale assessment due to their coarse resolution.

To create monthly BCSD projections, a two-step process was implemented including bias-correction and spatial disaggregation. Historical simulations from GCM output were compared to observations to identify and remove biases for the bias-correction step. Three datasets are gathered including observed historical data, derived

from Maurere et al. (2002) for the 20th century, simulated historical conditions that span from 1950 to 1999, and the GCMs' simulated future climate conditions (Reclamation, 2013). GCMs are initially regridded to a common spatial resolution of 1°. To identify biases within the historical time frame, cumulative distribution functions (CDFs) are used based on observational grid cells, months, and variables (Reclamation, 2013). The paired CDFs are combined to create "quantile maps," where, at each percentile, one can assess the bias between the simulated GCM output and observations (Reclamation, 2013). The next step entails adjusting simulated historical and future outputs based on the quantile maps to correct any biases. The result of this methodology is an adjusted GCM dataset that is statistically consistent with observations over the given time frame (Reclamation, 2013). Next, the spatial disaggregation of the GCM output is downscaled from the 1° bias-corrected to a 1/8° downscaled resolution. This is achieved by merging the historical spatial climatology with the spatially disaggregated changes of the given time step measured from that climatology (Reclamation, 2013). Factor values are computed for every grid cell that reflects departures from the observed spatial climatology at every grid cell, and translated to the targeted downscaled resolution using a modified inverse-distance-squared interpolation based on Shepards' (1984) computer mapping algorithm (Reclamation, 2013). Finally, the downscaled-resolution factor values are merged with the downscaled-resolution observational spatial climatologies to produce downscaled-resolution adjusted GCM values (Reclamation, 2013).

Geil et al. (2013) have demonstrated the limitations using CMIP5 coarsely resolved models in representing NAMS precipitation characteristics. Downscaling

methodologies provides an avenue to attempt to resolve resolution issues and, in turn, key variables in the NAMS. The goal of our study is to evaluate the ability of statistically downscaled models in their ability to realistically represent a highly resolved reanalysis product. Our results can provide a foundation for future projections employing downscaling methodologies and their ability to capture the climate system.

III.3 Data and Methodology

III.3.1 Study Region

Two domains were chosen to assess precipitation and temperature characteristics in CMIP5 model output relative to observations from the North American Regional Reanalysis (NARR). The core NAMS domain (112°W - 102°W to 25°N-33°N; Figure 3.1) was chosen based on previous analysis from Cook and Seager (2013), but constrained to areas north of 25°N due to data availability. An extended NAMS domain (125°W - 95°W to 25°N - 40°N; Figure 3.1) is also chosen to further assess synoptic-scale processes up- and downstream from the core domain. Furthermore, the North American Monsoon Experiment employs a multi-scale analysis including a regional study area similar to our extended domain to better understand the intraseasonal variability of the NAMS (Higgins et al., 2000).

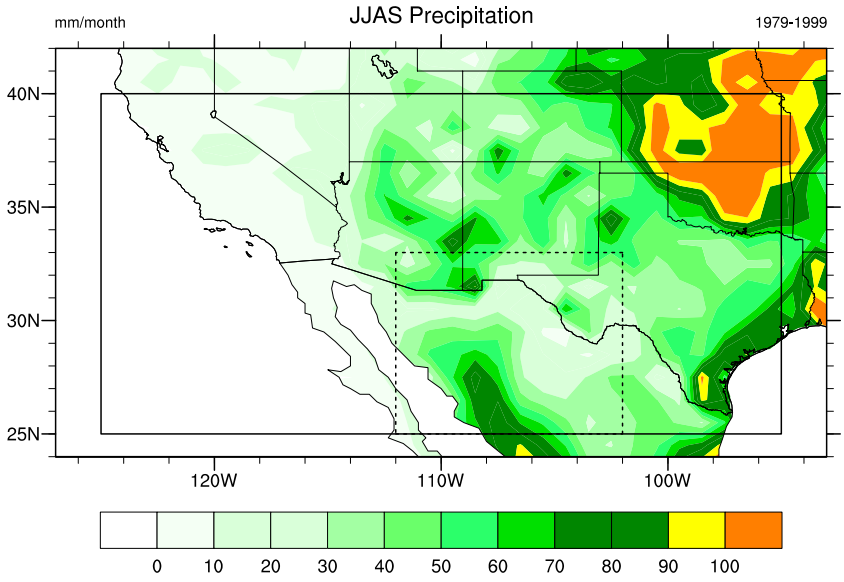


Figure 3.1. Seasonal climatology (JJAS) of the NAMS, based on NARR precipitation at 1° spatial resolution from 1979 to 1999. The dashed line represents the core NAMS region (112°W-102°W to 25°N-33°N) while the solid line illustrates the extended NAMS region (125°W-95°W to 25°N-40°N).

III.3.2 Reanalysis and Model Products

To assess the ability of statistically downscaled CMIP5 models in simulating the NAMS, monthly 2-meter air temperature and precipitation rate were obtained from NARR. NARR is based on the NCEP Eta Model (32 km /45 layer resolution) and numerous observational datasets to provide a much-improved analysis of land hydrology and land-atmosphere interactions (Mesinger et al., 2005). Three CMIP5 models are assessed based on previous findings from Geil et al. (2013) and Cook and Seager (2013), who evaluated CMIP5 model output for both historical and future time frames.

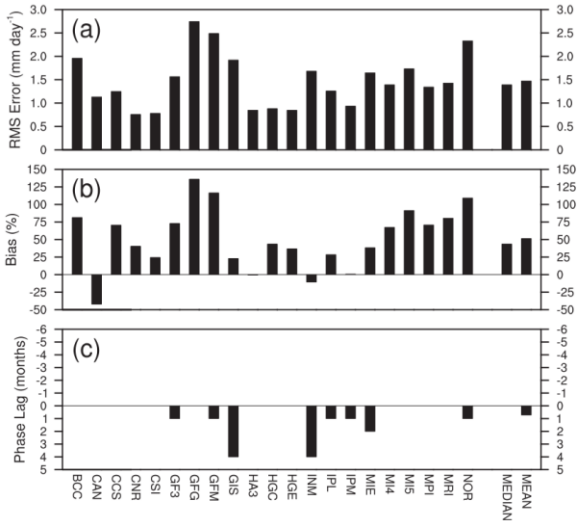


Figure 3.2. Comparison of the annual cycle of modeled and observed precipitation using root-mean-square error (RMSE) of monthly mean rainfall, percent bias of annual rainfall, and phase lag in number of months based on the defined core NAMS domain of Geil et al. (2013).

Figure 3.2 illustrates better performance of 5 of the 21 CMIP5 GCMs based on their low RMSE values and percent biases with no phase lag between 1979 to 2005: the Centre National de Recherches Météorologiques Coupled Global Climate Model, version 5 (CNRM-CM5; CNR), Commonwealth Scientific and Industrial Research Organisation Mark, version 3.6.0 (CSIRO-Mk3.6.0; CSI), Hadley Centre Coupled Model, version 3 (HadCM3; HA3), Hadley Centre Global Environmental Model, version 2 – Carbon Cycle (HadGEM2-CC; HGC), and Hadley Centre Global Environmental Model, version 2 – Earth System (HadGEM2-ES, HGE) (Geil et al., 2013). However, Cook and Seager (2013) identified 11 CMIP5 models based on their ability to simulate the NAMS seasonal cycle from 1980 to 1999 (Figure 3.3). Three of these higher resolution models, including CNRM-CM5, HadGEM2-CC, and HadGEM2-ES, were

able to simulate the seasonal progression of the NAMS in comparison to an observational climatology over their defined core monsoon domain. Thus, these three models are chosen here for subsequent evaluation.

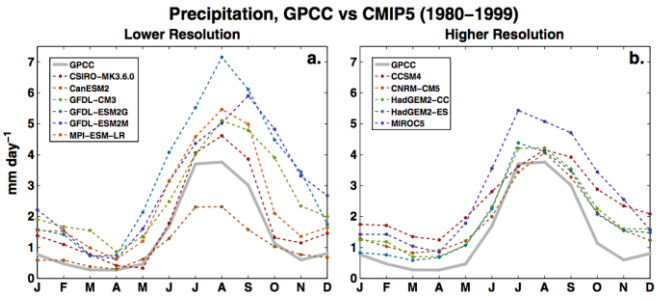


Figure 3.3. Precipitation comparisons of modeled and observed precipitation climatologies for the core NAMS region with lowest and highest resolution models displayed in (a) and (b), respectively (Cook and Seager, 2013).

CNRM-CM5 was developed specifically to contribute to CMIP5. It currently has a 1.4° and 1° horizontal resolution in the atmosphere and ocean respectively (Voldoire et al., 2012) and a new dynamical core based on a two time-level semi Lagrangian numerical integration scheme (Voldoire et al., 2012). Improvements over the previous versions were made, including a reduced bias in the mean surface temperature; however, precipitation and radiative biases still remain in certain regions.

The suite of HadGEM2 versions was created with specific model configurations incorporating varying levels of complexity, but with a common physical framework. These include the Carbon-Cycle (HadGEM2-CC) and Earth-System (HadGEM2-ES) versions (Martin et al., 2011). HadGEM2-CC includes the same processes as

HadGEM2-ES, except for the "chemistry" process. Both models have an atmospheric horizontal resolution of 1.875° × 1.25° and an ocean horizontal resolution of 1° (Martin et al., 2011).

The DHCP project provides statistically downscaled gridded CMIP5 output for climate and hydrology variables. Historical monthly bias-corrected GCM output at 1° was obtained for precipitation rate and near-surface air temperature for CNRM-CM5, HadGEM2-CC, and HadGEM2-ES. The data were downloaded for 1979 -1999, to overlap with the available NARR variables.

III.3.3 Methodology

To investigate the ability of statistically downscaled GCM output to simulate the NAMS precipitation and temperature characteristics, various statistical metrics and observational comparisons were quantified over the historical time period. Because of different spatial resolutions between model and observational output, NARR was regridded to a resolution of 1° through a bilinear interpolation scheme, a common regridding technique (e.g., Chen and Frauenfeld, 2014). Monthly and seasonal climatologies for temperature and precipitation were created to investigate models' ability to capture the seasonal progression of the monsoon system over the core monsoon domain, as well as the extended region. To quantify the spatio-temporal agreement between observations and model simulations, we calculated the root-mean-square difference (RMSD), Pearson correlation coefficients (R), linear regressions, and seasonal to yearly differences. When applicable, Student's t-tests were used to assess significance of differences, based on a 90% significance level similar, to previous assessments of

temperature trends and NARR (e.g., Fall et al., 2010). To evaluate the GCMs' skill in reproducing the seasonal and yearly climatologies, skill scores were calculated as defined by Taylor (2001):

$$S = \frac{(1+R)^4}{4(SDR + \frac{1}{SDR})^2}$$

The skill score (S) is calculated using the pattern correlation (R) between the model and observations and the ratio of spatial standard deviation (SDR) in the model against observations (Chen and Frauenfeld, 2014). Due to the limited number of grid cells, the test for statistical significance proposed by Song and Zhou (2014) was not implemented.

III.4 Results

III.4.1 Climatologies

We first investigate the seasonal spatial temperature and precipitation patterns during March-April-May (MAM), June-July-August (JJA), and September-October-November (SON), as well as the monthly climatologies. Similar patterns are evident in all models, thus only those from CNRM-CM5 are shown. Figure 3.4 shows the seasonal JJA patterns for simulated and observed temperature (left) and precipitation (right). The original (coarsely resolved) CNRM-CM5 output (Figure 3.4a) is also included to elucidate potential improvements of the downscaled output. In general, the spatial patterns in the downscaled output more closely resemble those from NARR. Furthermore, the decreased temperatures, due to the high terrain over the core monsoon region, are resolved in both the downscaled and original data.

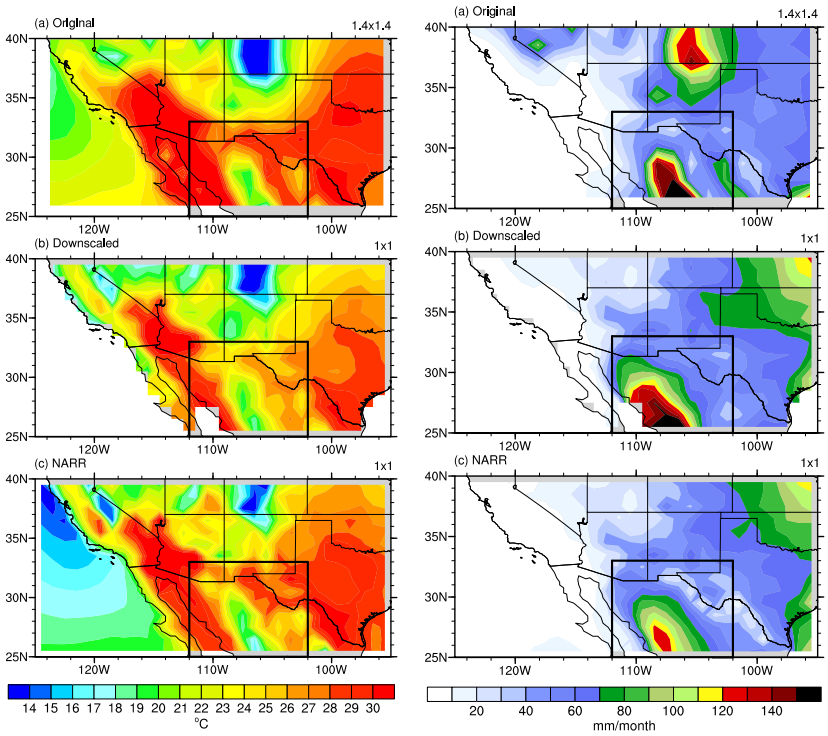


Figure 3.4. JJA temperature (left) and precipitation (right) spatial patterns for (a) original CNRM-CM5 output, (b) statistically downscaled CNRM-CM5 output, and (c) NARR observations from 1979 to 1999.

The temperature bias surrounding the Mohave Desert in the original data is also evident in the downscaled output, and a similar spatial pattern is observed in HadGEM2-CC and HadGEM2-ES (not shown, see Appendix 2). However, the downscaled CNRM-CM5 output captures only a fraction of the spatial temperature patterns east of the core domain, whereas the original output does capture the broad temperature pattern. Statistically significant differences between the downscaled CNRM-CM5 output and NARR depict a temperature bias over the core domain, as well as lower temperatures throughout New Mexico and Texas (Figure 3.5). Both HadGEM2s also show a similar temperature bias throughout the region (Appendix 2).

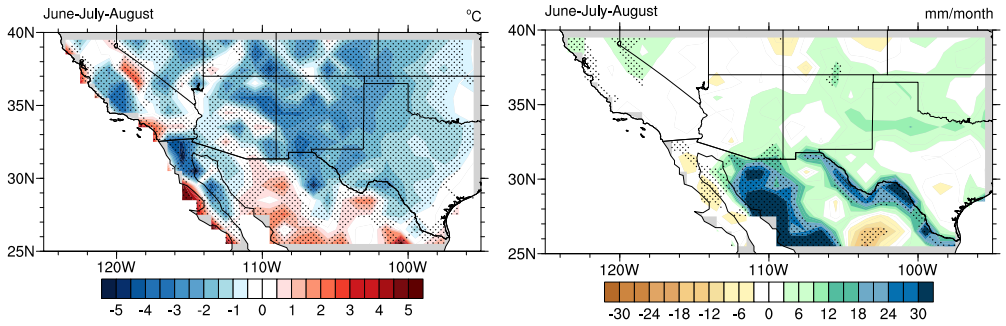


Figure 3.5. JJA temperature (left) and precipitation (right) spatial differences between downscaled CNRM-CM5 output and NARR with 90% statistical significance (stippled).

MAM spatial patterns are well captured by the downscaled GCM output, including the representation of decreased temperatures throughout the New Mexico, Arizona, Colorado, and Utah border region. Furthermore, the decreased temperature features are generally overestimated in the original resolution data (Appendix 2). SON spatial temperature patterns are adequately captured in both the downscaled and original GCM output, with a substantial cold bias throughout the southwest U.S. (Appendix 2).

Simulated JJA spatial precipitation patterns are well captured in the downscaled GCM output when compared to the observations, especially throughout the core monsoon domain (Figure 3.4). The seasonal progression of the monsoon system is replicated in both the downscaled and original CNRM-CM5 output, but a precipitation bias remains, as previously identified (Cook and Seager 2013, Geil et al., 2013). Furthermore, the precipitation bias is evident in the seasonal precipitation differences (Figure 3.5) over the Sierra Madre Occidental. However, the coarse resolution HadGEM2 models cannot replicate the local-scale precipitation characteristics of the monsoon system, simulating a large precipitation bias throughout the Sierra Madre

Occidental (Appendix 2). However, while all downscaled GCMs provide a similar precipitation pattern throughout Colorado and New Mexico during the summer season, a precipitation bias is also evident in the original GCMs. A reduction in the precipitation bias is found throughout the west and southwest U.S. in the downscaled GCM output for the pre-monsoon season (MAM) in comparison to the original GCM output. This is evident based on the seasonal differences between the downscaled GCMs and NARR output (Appendix 2). Broad spatial patterns during SON show a realistic reduction in precipitation throughout the core monsoon domain for the downscaled GCMs. Yet, a significant pattern is observed in parts of the core monsoon region for CNRM-CM5 and HadGEM2-ES, as well a large band south of the Rio Grande River for all downscaled GCMs (Appendix 2). The precipitation gradient over the Rio Grande is possibly an artifact in NARR. Both original HadGEM2s display broad precipitation patterns throughout the southwest, similar to those in NARR; however, CNRM-CM5 cannot duplicate the spatial pattern corresponding to the decay phase of the monsoon season.

Seasonal temporal patterns were explored, averaged for the core and extended domains for the summer months (Appendix 2). Over the core region, the observational temperature magnitude (~25°C) is well captured throughout the time period, whereas the coarsely resolved models fail to simulate the magnitude throughout the time series. However, both downscaled and original model output show little interannual agreement with NARR, an issue possibly driven by the steep terrain within the core region. Interestingly, both HadGEM2 models show increasing temperatures starting in 1990. Similarly, temperature time series results for the extended domain demonstrate a lack of

interannual agreement between downscaled and NARR output, but tend to capture the magnitude. Simulated precipitation results over the core and extended domain show little agreement with NARR in terms of their interannual variability. NARR's precipitation magnitude is partially captured by the downscaled output, including the earlier time period in CNRM-CM5 (1979 – 1985) and mid-period of HadGEM2-ES (1988 – 1992).

General descriptive statistics for annual and seasonal temperature in the core and extended domain are explored (Figures 3.6, Appendix 2). For annual temperature in the core domain, the three downscaled models' 75th percentiles fall within NARR's 25th percentile; however, only the HadGEM2-ES median falls within the reanalysis' 25th percentile. Thus, the results show an underestimation of temperature in the three downscaled models; however, HadGEM2-ES shows greater range of temperatures in the 25th percentile in comparison to NARR. In comparison, only the downscaled models' 5th percentile falls with NARR's 25th percentile over the extended domain, meaning a greater underestimation of temperature. During the summer months over the core domain, all models show a median of approximately 25.8°C, similar to the NARR temperature (26.1°C). However, all three models show the temperature data to be skewed below the median, illustrating a larger range. The extended domain illustrates a warm bias, with the HadGEM2s' 25th percentile overlapping with NARR's interquartiles. During the pre-monsoon season (MAM), NARR illustrates temperatures skewed below the median, however, interquartiles and averages fall within 18–19°C in the core domain, and 14–15.5°C in the extended domain for reanalysis and all three models. During SON, NARR denotes temperatures to be skewed below the median,

showing higher temperatures within NARR, as well as an average temperature of 19.2°C. This result is not evident in the downscaled output, which instead shows a decreased median in all models, with the lowest average model temperature at 17°C (HadGEM2-ES).

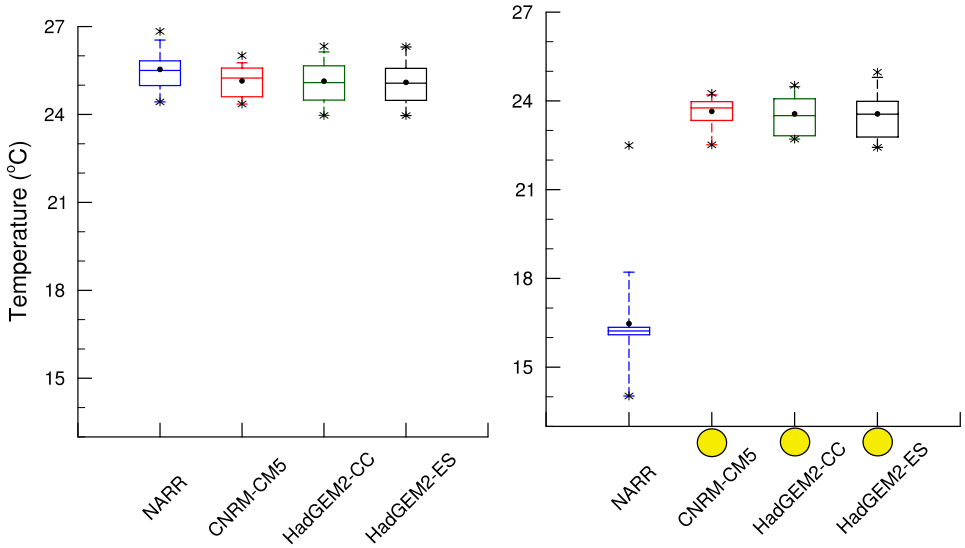


Figure 3.6. JJAS temperature for the core (left) and extended (right) domain. The 25th and 75th percentiles are shown by the edges of the boxes, the whiskers represent the 5th and 95th percentiles, and the medians are plotted as horizontal lines within the boxes. The asterisks illustrate the maximum and minimum temperature, average represents a black dot, while the yellow dots represent significant differences at the 90% level.

During the monsoon season (JJAS), the core domain temperature illustrates all downscaled models perform well in comparison to NARR (Figure 3.6; left). NARR's average and median temperature is around 25.5°C, while model average and median temperature is 0.5°C less than NARR. In comparison, the extended monsoon domain fails to reproduce the NARR temperature. However, individual JJAS temperature

outliers can skew the overall result, as evident from the maximum and minimum temperature values.

Annual precipitation over the core domain illustrates a wet bias in all downscaled models (Appendix 2). NARR indicates a median of about 28 mm/month, whereas the models over-predict the reanalysis observations by 6–9 mm, with their average near the median. The maximum annual precipitation of NARR and all downscaled models agrees well (~43–44 mm/month); however, the precipitation distribution within the downscaled model differs from NARR. Furthermore, all downscaled model output illustrates a wet bias of about 7 mm/month, based on precipitation averages. MAM precipitation in the core domain shows differences with the downscaled models persist. The HadGEM2s demonstrate similar 95th percentile results as NARR, illustrating higher precipitation amounts. Furthermore, the model medians indicate a slight overestimation in precipitation (~3–7 mm/month). The extended domain during MAM shows an overestimation of the average and median NARR value (~29 mm/month) by 4–6 mm/month, and greater maximum and minimum precipitation values. This similar overestimation in JJA is evident for both the core and extended domains.

During the post-monsoon season (SON), the downscaled models demonstrate good agreement between NARR and downscaled models in the core domain, with a slight average overestimation ranging from 4 to 8 mm/month. Over the extended domain, the downscaled models still show a wet bias, with an average overestimation ranging from 8 to 10 mm/month. This overestimation illustrated in both JJA and SON is also evident during the monsoon season (JJAS) in the core and extended domain (Figure

3.7). Although the downscaled models capture the higher precipitation over the core domain, they overestimate the average and median by ~12–15 and ~10–13 mm/month, respectively. Furthermore, the range of precipitation differs considerably in the model output when compared to NARR, as well as overestimating the median by a considerable amount (CNRM-CM5 = ~18 mm/month, HadGEM2-CC = ~21 mm/month, HadGEM2-ES = ~15 mm/month).

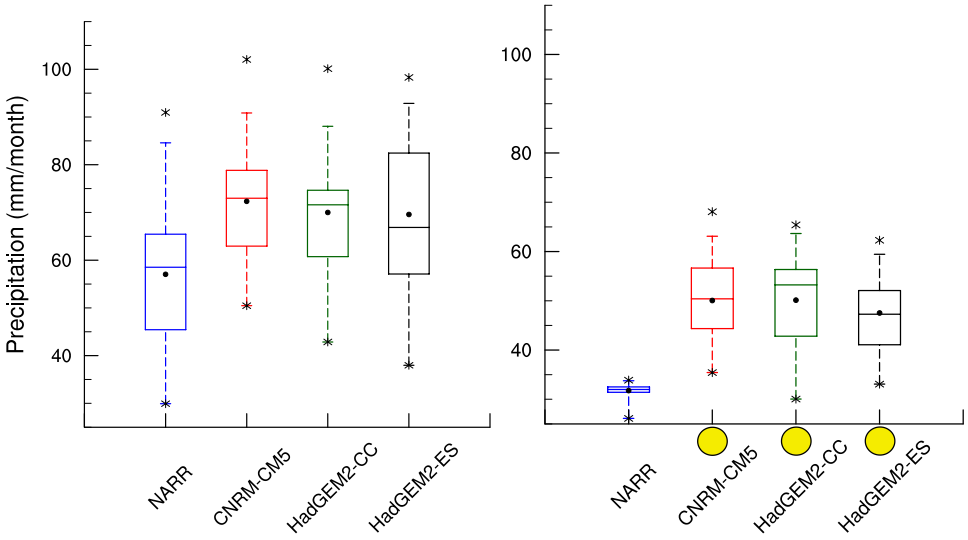


Figure 3.7. Similar to figure 6, except for precipitation (mm/month).

III.4.2 Monthly Climatologies

All grid cells were averaged within both the core and extended domains, and monthly temperature and precipitation climatologies of the annual cycle for the three GCMs and NARR were created (Figure 3.8 – 3.9). The monthly temperature cycle in NARR throughout the core and extended monsoon domains are well captured by the

downscaled GCMs (Figure 3.8). Temperature in the original GCMs' output (red) is overestimated in the summer months for both domains, with both HadGEM2s overestimating by up to ~4°C and CNRM-CM5 by ~2–3°C. To further illustrate the performance of the downscaled GCMs, statistically significant differences between the downscaled GCM and NARR output were computed for the core and extended monsoon domain (Appendix 2). Significant differences are evident over fall and early winter months for the three downscaled models, underestimating the transition to lower temperatures by -1°. During the monsoon season, temperature differences show little temperature bias, which is no statistically. In contrast, slight increasing temperature biases are observed during the monsoon season, with the greatest bias during June and July in HadGEM2-ES (0.89°C) and CNRM-CM5 (0.93°C), respectively.

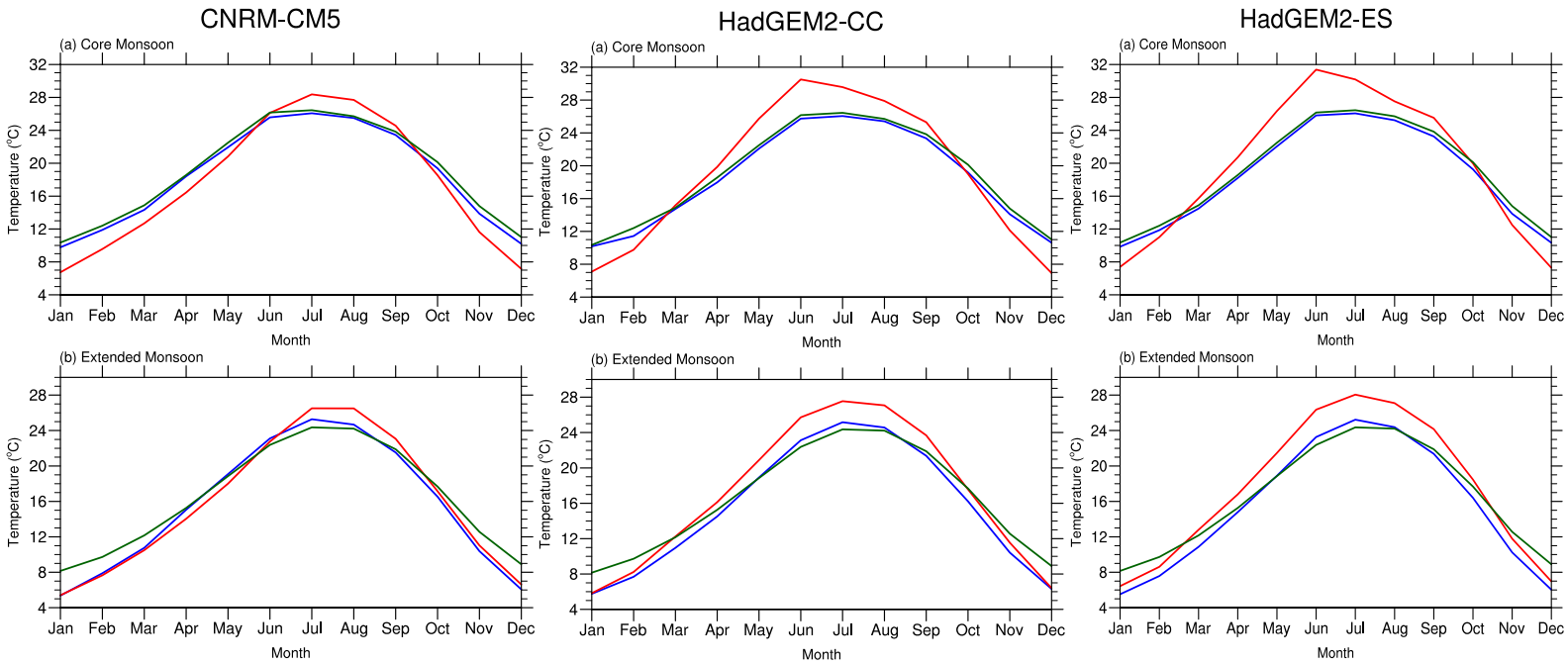


Figure 3.8. Monthly temperature climatologies for the (a) core and (b) extended monsoon domain from 1979 to 1999 for NARR (green line), original GCMs (red), and downscaled GCMs (green). CNRM-CM5, HadGEM2-CC, and HadGEM2-ES are shown from left to right.

Figure 3.9 depicts monthly climatologies for precipitation for all three GCMs. Within the core domain (Figure 3.9a), the seasonality of the monsoon system is evident, including its rapid onset (June) and decay (September). All three downscaled GCMs (blue) capture this seasonal progression, but exhibit a wet bias during the summer months. Furthermore, winter precipitation, which can contribute up to ~30% of annual precipitation over the southwest U.S., is well captured by the downscaled GCMs. However, the peak in precipitation during the monsoon season has shifted in the downscaled GCMs by a month when compared to NARR. This seasonal shift could be due to the observational product used in the bias-correction of the downscaling methodology. In contrast, coarsely resolved GCMs have varying degrees of agreement with NARR precipitation. CNRM-CM5 original output overestimates winter and spring precipitation by up to ~20 mm/month and has an early onset of the monsoon system. In comparison, the HadGEM2s overestimate precipitation throughout all months but capture the seasonality of the monsoon system. Within the extended domain (Figure 3.9b), a better representation of precipitation is evident between the downscaled GCMs' and NARR output when compared to the original GCM output. However, the downscaled GCM precipitation does not simulate the NARR spring magnitude, but has improved the high precipitation bias in April and June (red). Downscaled GCM differences with respect to NARR reveal the greatest bias to be during the monsoon season, particularly during July (20 – 22 mm/month).

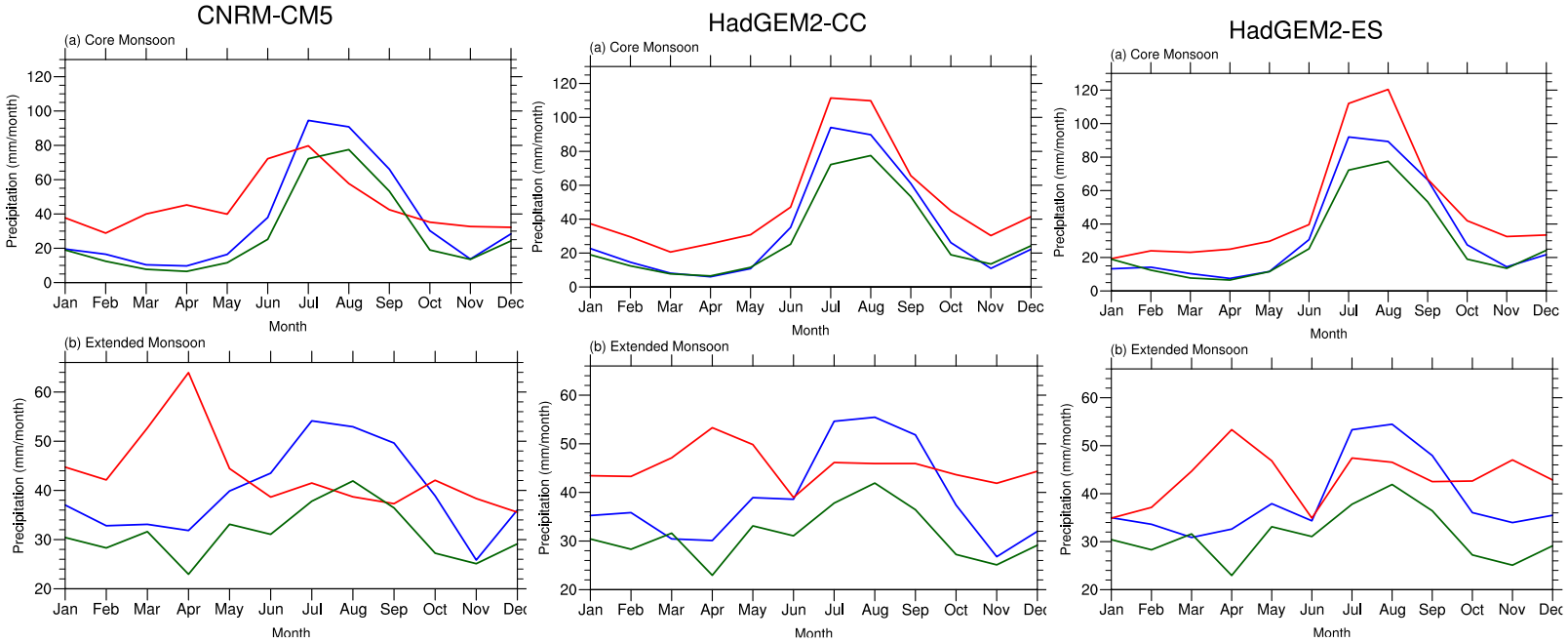


Figure 3.9. Monthly precipitation climatologies for the (a) core monsoon domain and (b) extended monsoon domain from 1979 to 1999 for NARR (green line), original GCM output (red line), and downscaled GCM output (blue line). CNRM-CM5, HadGEM2-CC, and HadGEM2-ES are shown from left to right.

Although the downscaling methodology tries to correct these biases based on observations, the representation of precipitation still remains an issue due to the uncertainties within the individual GCMs. Still, the bias correction methodology has shown improvements in the shortcomings within the coarsely resolved GCMs, evident through Figure 3.9. Also, the greatest bias is in CNRM-CM5 over the core domain, with the greatest difference in precipitation in 8 out of the 12 months (Appendix 2). For the extended domain, a consistent wet bias is observed throughout most of the months for all downscaled models, with a significant bias over the monsoon season (Appendix 2). The extended domain takes into account larger regions over the United States, including regions of complex terrain over Colorado and Utah, which can impact simulated precipitation.

Monthly interannual variability for temperature was also explored for all three GCMs to assess potential improvements in the downscaled output (Figure 3.10; Appendix 2). During the end of fall to early spring, downscaled CNRM-CM5 temperature output demonstrates better agreement with NARR in comparison to the original CNRM-CM5 output, which shows a cold bias during this time (Appendix 2). Furthermore, the increased temperatures during the summer months are well captured in the downscaled GCM output, whereas the original coarse model has increased temperatures (Figure 3.10). However, CNRM-CM5 downscaled temperature does not capture the interannual variability in NARR temperatures.

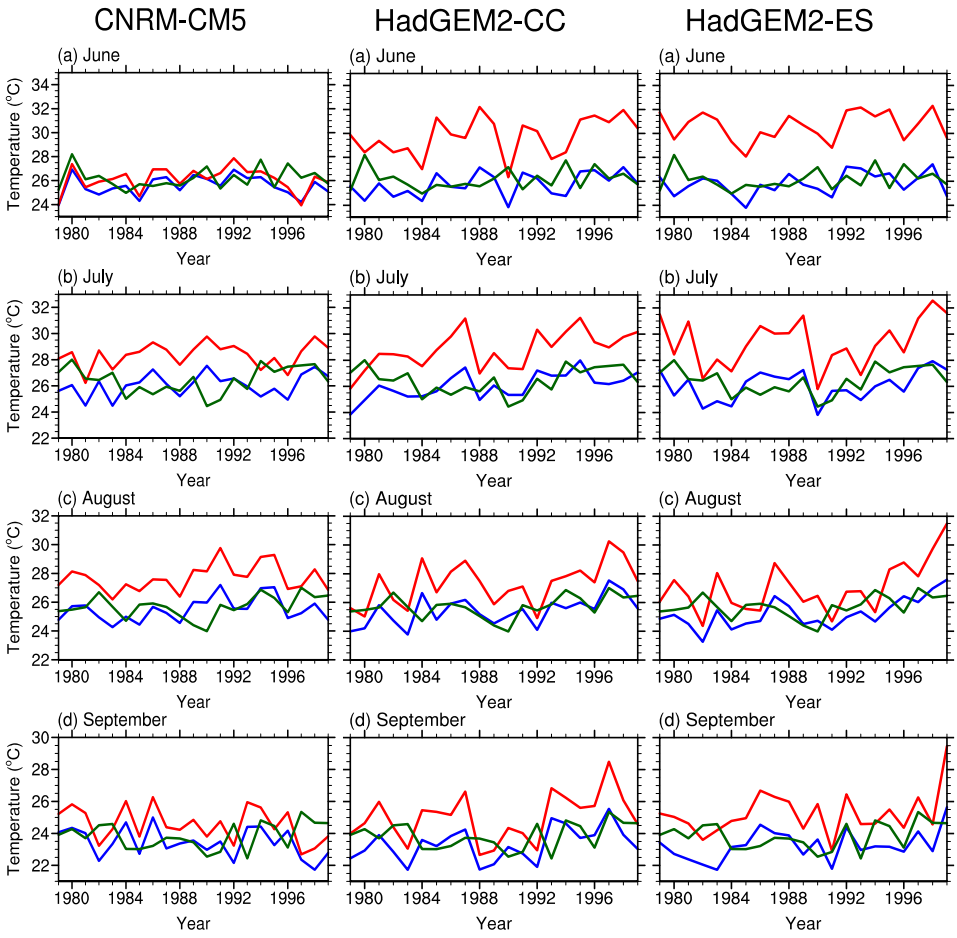


Figure 3.10. Interannual temperature variability from 1979 to 1999 for the core monsoon domain in NARR (green), downscaled output (blue), and coarsely resolved output (red) for the monsoon season.

Similarly, HadGEM2-CC downscaled output shows a much better agreement in simulating the magnitude of temperature in comparison to the coarse model (Figure 3.10, Appendix 2). The downscaled HadGEM2-CC does capture the interannual variability in some months, including November starting in 1992; however, temperature is either overestimated or underestimated within the downscaled output during the monsoon months. Interestingly, there is near-perfect agreement in interannual variability

during October between the downscaled and original model output, similarly to the month of June in CNRM-CM5. Finally, HadGEM2-ES downscaled output provides a better representation of temperature in comparison to the coarse model, as was the case for both previous models (Figure 3.10, Appendix 2). HadGEM2-ES original model output simulates summer temperatures poorly, whereas the downscaled output improves on the representation of temperatures when compared to NARR. Furthermore, some of the interannual variability is captured during the summer months, e.g., in July and September. However, as for the other two models, the downscaled HadGEM2-ES still has difficulty in reproducing interannual variability when compared to the NARR output.

Interannual temperature variability for the extended monsoon domain for CNRM-CM5 depicts both improvements and shortcomings in the downscaling output (Appendix 2). During the summer months, better representation of the simulated temperature is evident when compared to observations, as well as improved downscaled output in contrast to the original CNRM-CM5 (e.g., July through September). However, October through March depicts a near-perfect agreement of the downscaled and original model output, with a cold bias when compared to observations. For HadGEM2-CC, a similar near-perfect agreement between downscaled and original simulations is evident during December through February (Appendix 2). Furthermore, there is better agreement between NARR and the original model output during October and November when compared to downscaled output. Yet, the summer months are well simulated by the downscaled HadGEM2-CC, with some of the interannual variability captured by the

model during the monsoon season. However, a bias is observed throughout most of the months, a consistent pattern in the downscaled output from the other models. Similarly, HadGEM2-ES shows better agreement between the original and NARR output than the downscaled data (Appendix 2). The monsoon season is well captured by the downscaled data, but the high-resolution output still has difficulty in reproducing the interannual variability seen in the NARR temperatures.

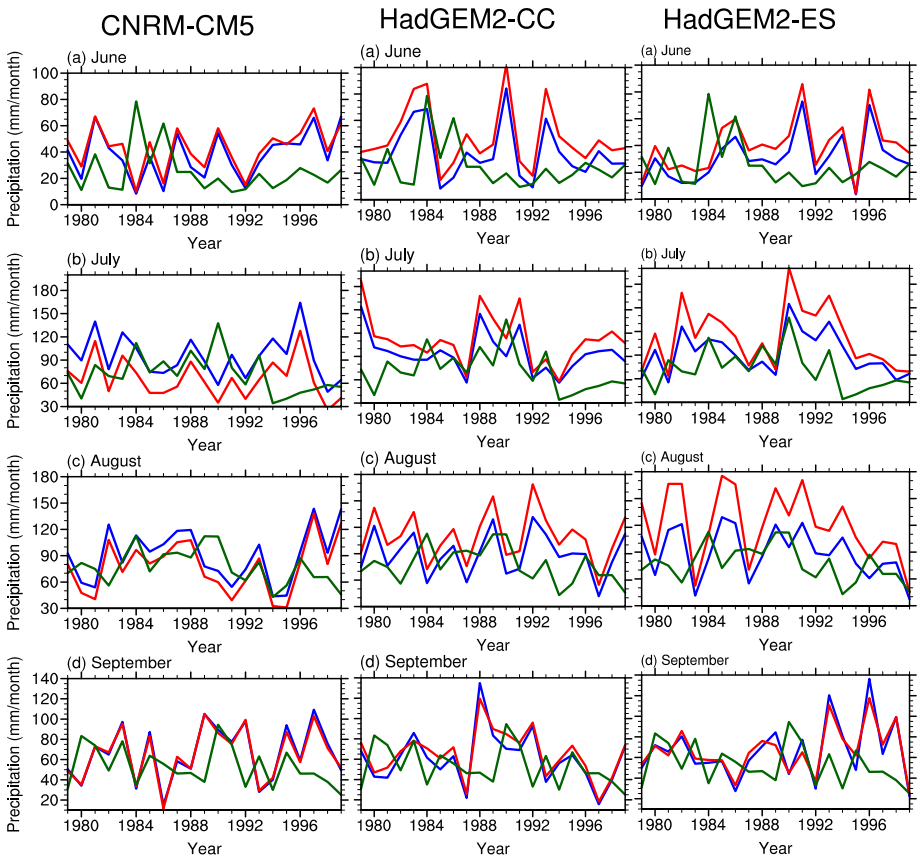


Figure 3.11. Interannual precipitation variability from 1979 to 1999, for the core monsoon domain in NARR (green), downscaled output (blue), and coarsely resolved output (red) during the monsoon season.

Interannual variability of precipitation for the monsoon season over the core domain is shown for all three GCMs' in figure 3.11. CNRM-CM5 demonstrates high interannual variability during most months, including the fall and winter seasons (Appendix 2). Although the model is inconsistent with NARR's interannual variability, it has a smaller wet bias when compared to the original model. During the monsoon season, the downscaled and original model output are similar, with the original model better simulating precipitation during July and August. Through the downscaling technique, HadGEM2-CC reduces the wet bias in the original model output, capturing the general precipitation. Yet, an issue remains with the downscaled HadGEM2-CC in simulating NARR's interannual variability. During the monsoon season, the model can capture some of the large precipitation increases in NARR during June, but the wet bias still remains an issue during the months of July and August (Figure 3.11). In comparison, the downscaled HadGEM2-ES does fairly well in capturing the interannual variability in July and August (Figure 3.11). Improvements in the downscaled HadGEM2-ES are further illustrated during April and May, when the NARR variability is well captured in comparison to the original model output. For precipitation over the extended domain, the downscaled CNRM-CM5 is improved during April and May, but aligns closely with the coarse model output during other months (Appendix 2). During the monsoon season, the downscaled output cannot capture the NARR interannual variability. A similar level of agreement is evident in both HadGEM2 models (Appendix 2). During the monsoon season, the downscaled HadGEM2s show no improvements relative to the original model output when compared to NARR.

III.4.3 Spatial Temperature Patterns

Monthly climatology differences between downscaled model simulations and NARR output are assessed for their spatial agreement throughout the monsoon domain. Similar temperature bias patterns are observed for all three downscaled models during the monsoon season (Figure 3.12, Appendix 2). Smaller temperature biases are found during the winter to spring months, with a significant warm bias over the Sierra Madre Occidental and a cold bias east of the range, in Texas. During the monsoon season, this cold bias spreads over the entire southwest, with an increasing temperature bias over the Sierra Madre Occidental (Figure 3.12). The temperature bias over the core domain is similar to the seasonal differences also seen in figure 3.5. The temperature increases in the downscaled output could be a product of the orographic representation in the model.

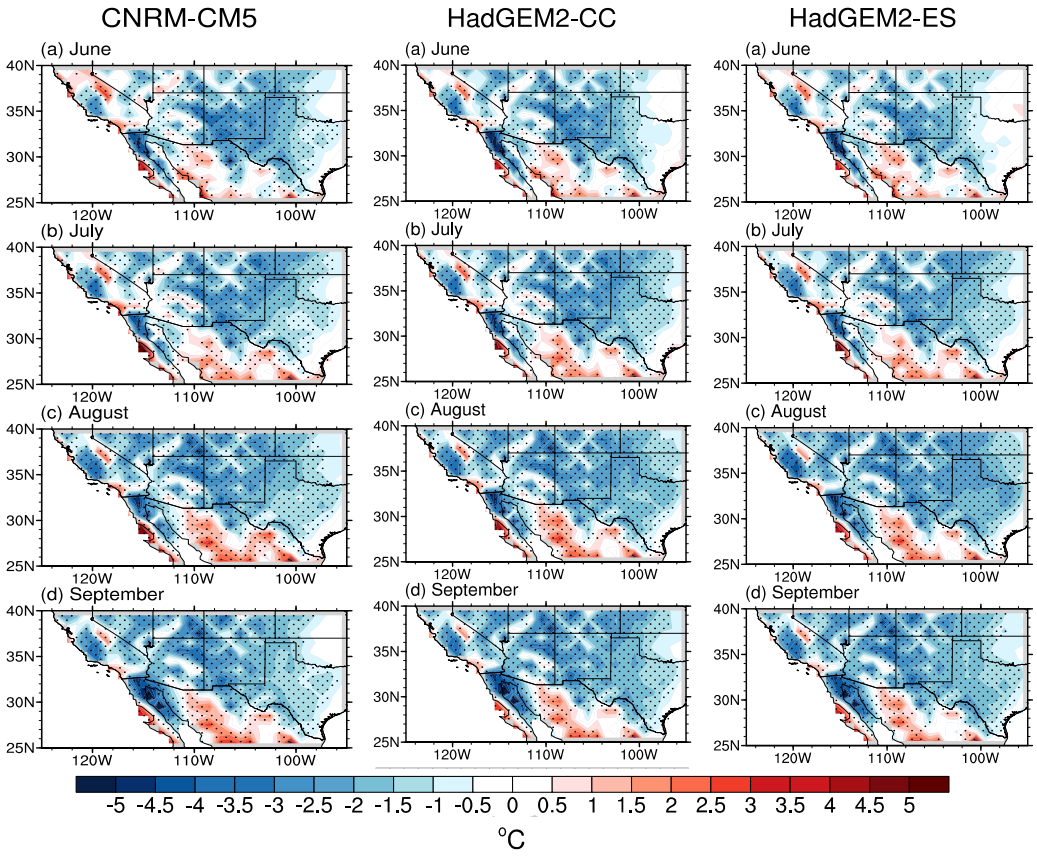


Figure 3.12. Temperature differences for (a) June, (b) July, (c) August, and (d) September between the downscaled simulated output and NARR with a statistical significance of 90% (stippled) during the monsoon season.

Monthly temperature RMSD was calculated to evaluate grid cell errors of the downscaled model output (Figure 3.13, Appendix 2). Small local-scale differences are observed in all three downscaled models, but the overall spatial temperature error patterns are similar in all three models. Small RMSDs are evident throughout most months, with the greatest RMSD values in late winter and spring (~5°C). During the monsoon season in the core domain, low RMSD values are observed. This suggests skillful representation of temperature in downscaled model output. High RMSD values

(~10°C) are observed throughout all months over central California and at the mouth of the Gulf of California (GoC).

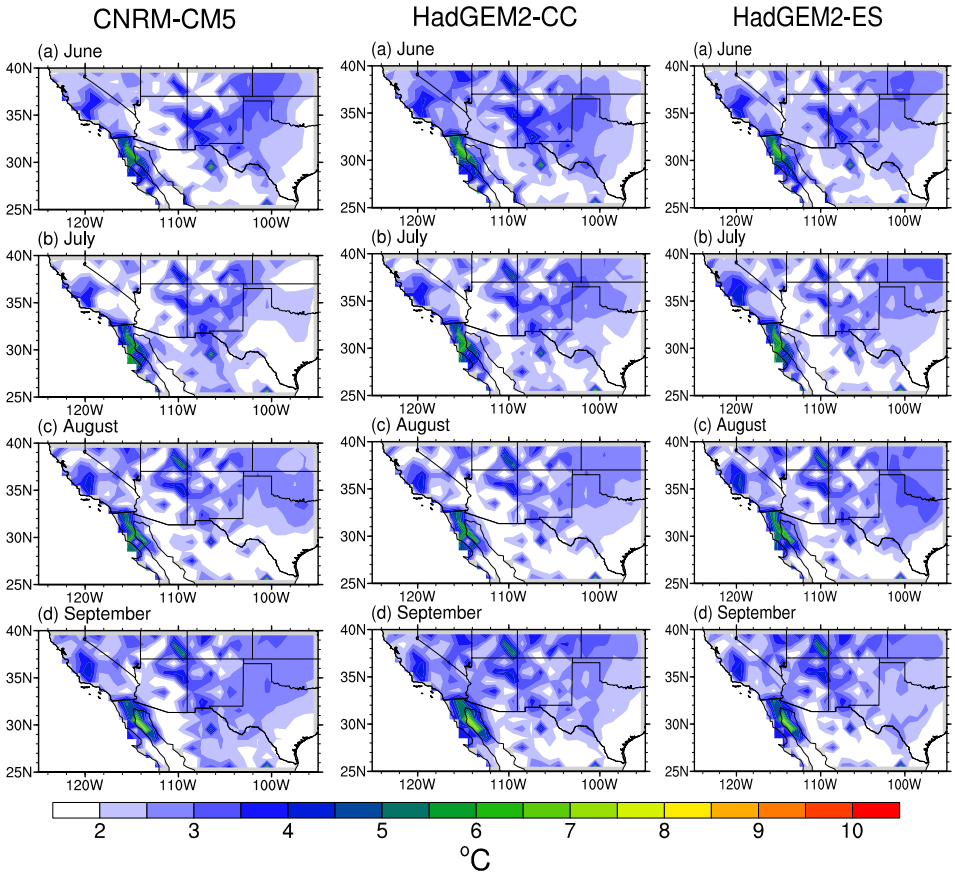


Figure 3.13. Temperature Root Mean Square Difference (RMSD) for (a) June, (b) July, (c) August, and (d) September between the downscaled simulated output and NARR with a statistical significance of 90% (stippled) during the monsoon season.

Monthly temperatures are correlated between downscaled model output and NARR (Figure 3.14, Appendix 2). For CNRM-CM5 downscaled output, there are both positive and negative correlations throughout the months. The highest positive correlations are demonstrated during September over the mouth of GoC (~0.65),

demonstrating the ability of CNRM-CM5 in simulating temperature interannual variability as shown by NARR. Significant positive correlations over the core domain are evident during the month of June; however, September illustrates a positive-negative pattern. Furthermore, a negative correlation over the southwest U.S. shows the inability of CNRM-CM5 to simulate temperature variability as represented by NARR, which assimilates rawinsondes. CNRM-CM5 uses a global spectral atmospheric model (ARPEGE-Climat) that is derived from the ARPEGE/Integrated Forecast System numerical model that includes temperature as a prognostic variable (Voldoire et al., 2011). The different temperature assimilations between the reanalysis and model thus can vary, further exacerbated by downscaling methodologies.

The downscaled HadGEM2-CC model illustrates a different correlation pattern when compared to CNRM-CM5 (Figure 3.14). The seasonal temperature correlations show either little to no positive signal during the earlier time period of the monsoon season; however, August illustrates a high correlation with NARR over the monsoon domain. Furthermore, the positive correlation during the month of August is identified for all three downscaled models, with CNRM-CM5 and HadGEM2-ES illustrating a positive correlation mainly over the Chihuahua Desert. Although the HadGEM2s share similar configurations, the correlation patterns differ considerably. This is evident in the months of June and July, where HadGEM2-CC demonstrates opposite correlations seen in HadGEM2-ES. HadGEM2-CC includes configurations from the atmosphere and ocean, coupled with the carbon cycle configuration that includes both the terrestrial carbon cycle and ocean biogeochemistry (Martin et al., 2011). The Earth-system model

(HadGEM2-ES) not only includes the previous mentioned configurations, but also takes into account tropospheric chemistry to quantify the Earth system feedbacks (Martin et al., 2011). Thus, the different configurations implemented within the HadGEM2s apparently produces contrasting results, evident through our correlation analysis. However, HadGEM2-ES illustrates a strong correlation during the month of September (0.84), similar to results shown in HadGEM2-CC (0.68) (Appendix 2). However, HadGEM2-ES illustrates a positive correlation over the core domain in July, demonstrating that the Earth system model is capturing part of the temperature variability.

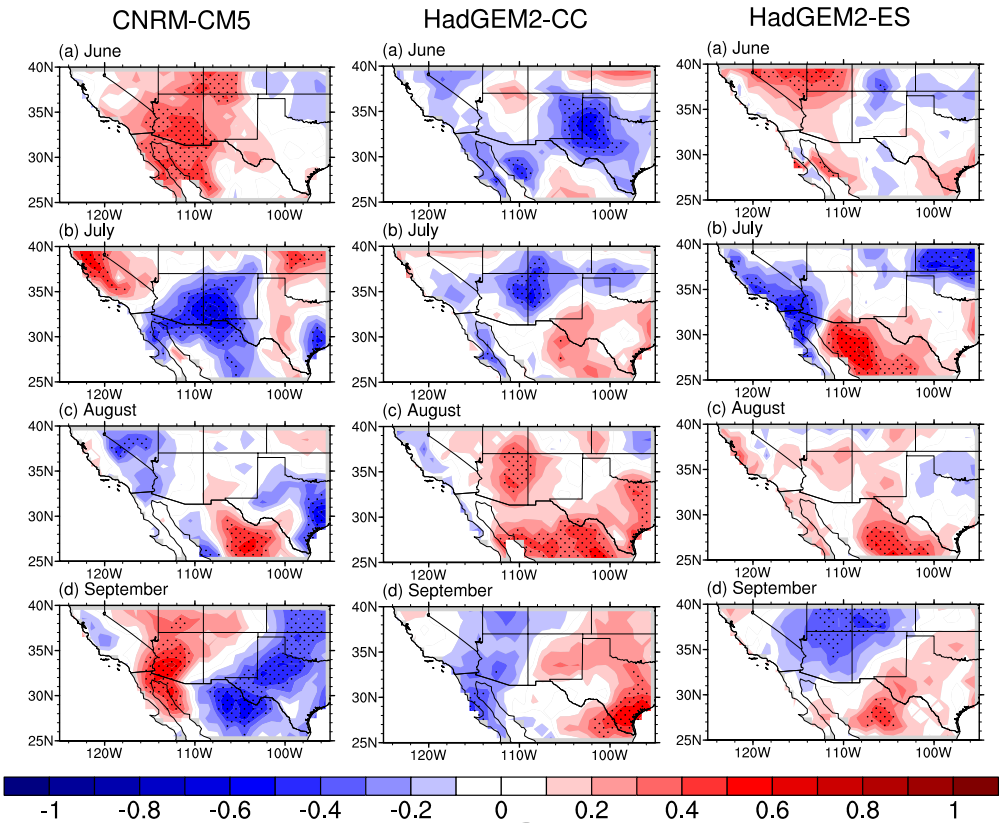


Figure 3.14. Temperature correlations for (a) June, (b) July, (c) August, and (d) September between downscaled simulated output and NARR with a statistical significance of 90% (stippled) during the monsoon season.

To further elucidate the relationship between NARR and modeled temperatures, linear trends were explored in NARR and compared to those from the downscaled model output. Initially, an area-average of the two domains was analyzed to obtain an estimate of the overall monthly linear trends. For the core monsoon domain, NARR exhibits statistically significant warming during May and November of 0.09°C yr^{-1} and 0.07 yr^{-1}, respectively (Table 3.1). No other significant increases in temperature are seen. The overall 1979 to 1999 NARR trend indicated no significant change, a similar result illustrated by the downscaled models. The monthly trends are not well captured by both HadGEM2s. HadGEM2-CC depicts a significant warming trend during the summer and fall, while HadGEM2-ES shows a significant warming in late summer, late fall, and winter. However, both models do capture the observed warming in November. CNRM-CM5 downscaled output shows significant warming in May (0.06°C yr^{-1}) similar to NARR, but fails to capture any other significant trends in other months.

	NARR											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.05	0.04	0.03	-0.03	0.09	0.02	0.03	0.04	0.03	0.02	0.07	-0.04
P-Val	0.17	0.32	0.37	0.47	0.01	0.57	0.50	0.14	0.36	0.47	0.03	0.31

	CNRM-CM5											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.01	0.04	0.06	0.06	0.06	0.01	0.05	0.04	-0.05	0.03	0.00	0.02
P-Val	0.85	0.46	0.17	0.18	0.05	0.73	0.14	0.16	0.12	0.32	0.98	0.57

	HadGEM2-CC											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.01	0.11	0.00	-0.02	0.05	0.07	0.10	0.09	0.07	0.07	0.16	-0.02
P-Val	0.90	0.02	0.95	0.58	0.26	0.06	0.00	0.01	0.07	0.07	0.00	0.67

	HadGEM2-ES											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.07	0.10	0.10	0.05	0.02	0.04	0.06	0.10	0.06	0.05	0.09	0.08
P-Val	0.22	0.05	0.08	0.19	0.72	0.21	0.18	0.00	0.08	0.29	0.04	0.06

Table 3.1. Monthly 1979–1999 temperature trends for NARR, CNRM-CM5, HadGEM2-CC, and HadGEM2-ES over the core domain. Statistically significant trends (90% level) are shaded.

Linear temperature trends for the extended domain for all four datasets are shown in table 3.2. The linear trends for NARR show a significant warming for the months of February (0.07 °C yr^{-1}) and May (0.04 °C yr^{-1}). CNRM-CM5 identifies significant warming during the spring season (0.6 – 0.7 °C yr^{-1}). However, the HadGEM2s exhibit a significant warming through most of the summer to early winter, a warming that is not seen in NARR.

	NARR											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.06	0.07	0.03	-0.02	0.05	0.00	-0.01	0.01	-0.01	0.00	0.04	-0.03
P-Val	0.12	0.08	0.26	0.57	0.08	0.96	0.80	0.53	0.69	0.90	0.13	0.35

	CNRM-CM5											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.01	0.07	0.07	0.07	0.06	0.01	0.02	0.02	-0.03	0.05	0.01	0.00
P-Val	0.81	0.18	0.06	0.04	0.02	0.59	0.34	0.34	0.35	0.19	0.82	0.93

	HadGEM2-CC											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.02	0.10	-0.03	-0.03	0.01	0.04	0.06	0.06	0.07	0.05	0.13	-0.02
P-Val	0.66	0.04	0.44	0.49	0.67	0.18	0.04	0.02	0.04	0.13	0.01	0.64

	HadGEM2-ES											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (°C/yr)	0.08	0.07	0.06	0.06	0.03	0.06	0.08	0.12	0.08	0.09	0.08	0.14
P-Val	0.13	0.14	0.16	0.12	0.45	0.10	0.02	0.00	0.01	0.00	0.06	0.01

Table 3.2. Monthly 1979–1999 temperature trends for NARR, CNRM-CM5, HadGEM2-CC, and HadGEM2-ES over the extended domain. Statistically significant trends (90% level) are shaded.

Figure 3.15 shows the temperature trends over the 21-year period for NARR and downscaled model output for the monsoon season. Significant warming trends during the monsoon season occurred over the core domain during August and September for NARR. Also, the largest warming trend for NARR occurred during February over Oklahoma and Kansas, while a significant cooling trend occurred during December throughout the southwest U.S., evident in the area-averaged results (Table 3.2). CNRM-CM5 downscaled output captures the spatial warming trend during February throughout Kansas and Oklahoma, as well as the warming over the Sierra Madre Occidental in

August (Figure 3.15, Appendix 2). However, it fails to capture any of the significant warming or cooling during late fall and early winter in NARR. Regions of significant warming towards the western U.S. are demonstrated during March and May, a signal not evident in the reanalysis product. In comparison, HadGEM2-CC downscaled model illustrates the largest warmest trend during November, evident in Table 3.2 with an area-averaged warming of 0.13°C yr^{-1} (Appendix 2). During most of the monsoon season, a large region of significant warming is evident, especially throughout the core domain (Figure 3.15). The significant cooling in December is not well captured in the region, while the large magnitude of warming in February in NARR is shifted towards Arizona and west Texas in HadGEM2-CC.

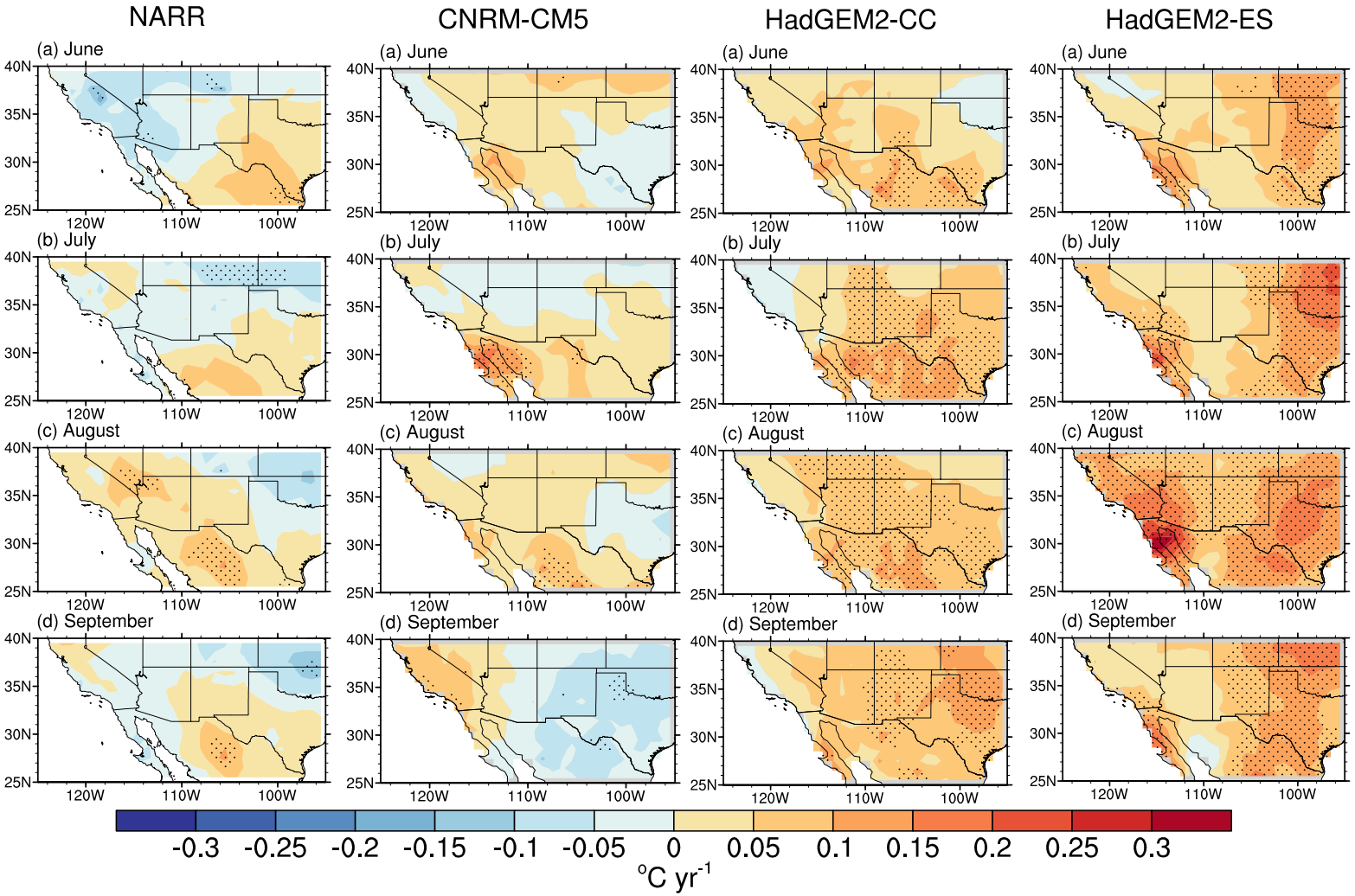


Figure 3.15. Monthly temperature trends for (a) June, (b) July, (c) August, and (d) September. Stippled regions indicate statistically significance (90% level) during the monsoon season.

In general, the spatial warming in the model is overestimated compared to NARR. The general trend for HadGEM2-ES shows significant warming during many of the months (Figure 3.15, Appendix 2). A large region of significant warming is simulated during the months of July, September, and October, which is not evident in NARR. The November warming pattern over the core monsoon domain is partly captured in HadGEM2-ES; however, the largest warming trend in February is not well captured in the model in terms of its spatial distribution. Systematic biases in temperature means and extremes appear in all CMIP5 models, which may be related to the land hydrology that affects the surface energy budget in drier regions (Sheffield et al., 2013). Furthermore, previous analyses have shown western North American to have a warm bias both in CNRM-CM5 and HadGEM2-ES during the summer months (Sheffield et al., 2013). The trend results support the temperature biases found within previous CMIP5 analyses, illustrating potential in identifying the shortcomings in simulating accurate climate conditions and systematic biases.

III.4.4 Spatial Precipitation Patterns

To further assess the ability of the downscaled model output to simulate NARR precipitation, we calculate the differences between the model and observations (Figure 3.16, Appendix 2). The spatial precipitation difference patterns are similar for all models, with only slight local differences. A spatial dry bias throughout the southern Chihuahua Desert is evident during January in CNRM-CM5 and HadGEM2-ES, while HadGEM2-CC illustrates no significant difference (Appendix 2). However, the HadGEM2s simulate a dry bias throughout Texas during March that does not appear in

CNRM-CM5, while a wet bias is observed throughout all downscaled models over west Texas during April. During the initial stages of the monsoon season (June), a significant precipitation bias is seen over the Sierra Madre Occidental and Sierra Madre Oriental mountains, reaching up to ~30 mm/month (Figure 3.16). Also, HadGEM2-ES demonstrates a strong significant dry bias over Oklahoma and Texas. During the early stages of the monsoon season, an inverse effect occurs as precipitation builds over Mexico and precipitation decreases in the U.S. Great Plains. All three downscaled models show this seasonal pattern (Figure 3.3, Appendix 2), however, HadGEM2-ES overestimates the lower precipitation in the U.S. Great Plains. The persistent wet bias over the mountainous regions of Mexico continues until September, a signal that has been observed in many CMIP5 models. Furthermore, these biases have also been noted in other monsoon systems, partly related to the vegetation distribution configurations of the HadGEM2s (Martin and Levin, 2011). During the winter months, a wet bias is simulated over the coastal ranges of California, as well as a dry bias over the Sacramento Valley. The wet and dry biases are largely over steep terrain, an issue many models face due to poor orographic representation. Previous analysis has shown that most CMIP5 models capture the Intertropical Convergence Zone (ITCZ), a key component in the progression of the NAMS (Sheffield et al., 2013). Furthermore, HadGEM2-ES has shown to have the lowest RMSE scores over a similar core domain for annual precipitation (Sheffield et al., 2013). However, a percent bias of 40.4, 44.8, and 37.9 over the core monsoon region was found for CNRM-CM5, HadGEM2-CC, HadGEM2-ES, respectively (Sheffield et al., 2013). Our results support previous analyses of the

overestimation of precipitation due to the timing of the monsoon system. However, all three models have shown to improve the precipitation shifts from their predecessor, CMIP3 (Geil et al., 2013, Sheffield et al., 2013).

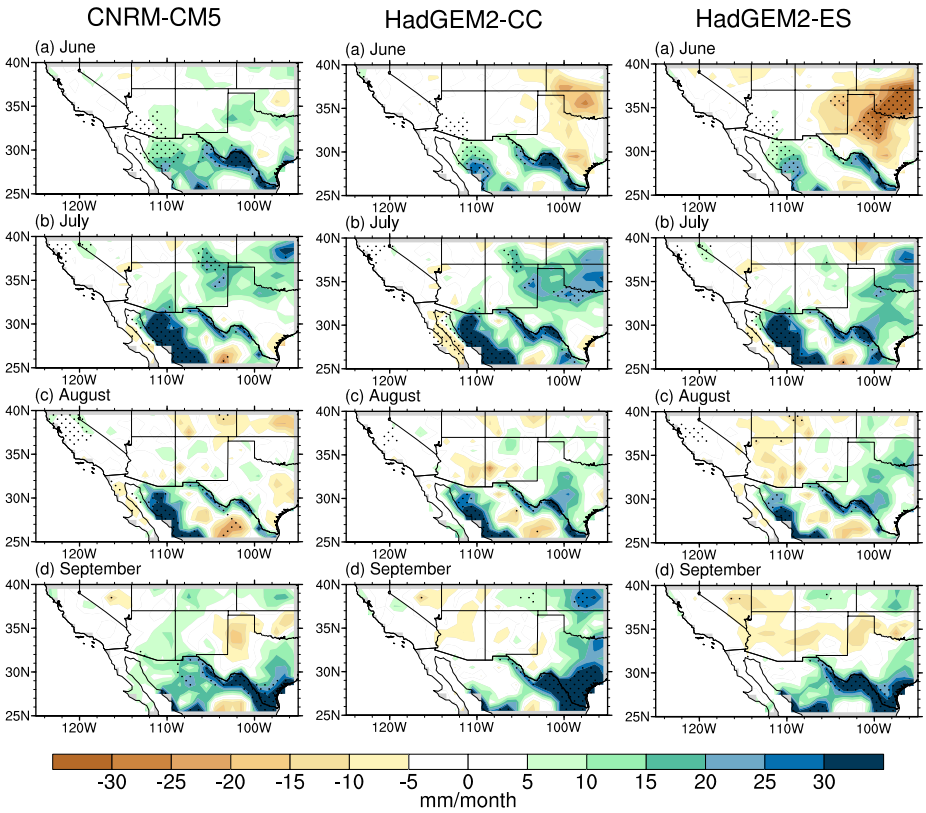


Figure 3.16. Monthly precipitation differences (mm/month) for (a) June, (b) July, (c) August, and (d) September between the downscaled simulated output and NARR with a statistical significance of 90% (stippled).

The differences between modeled precipitation simulations and NARR based on RMSD illustrate a higher disagreement for the downscaled model output, compared to temperature RMSD. Greater disagreement between all downscaled model and reanalysis output is evident over northern California from November to March, exceeding RMSD

values of 160 mm/month (Appendix 2). During the pre-monsoon season, in CNRM-CM5, the RMSD in precipitation illustrate values between 0 to 40 mm/month over the southwest U.S., with increased disagreement over Texas and Oklahoma by about 80 to 100 mm/month. This disagreement over Texas is less pronounced in HadGEM2-CC during April, while HadGEM2-ES shows a more pronounced disagreement. During the monsoon season, substantial differences are observed over the steep terrain of the Sierra Madre Occidental, with values reaching up to ~120 mm/month (Figure 3.17). As the monsoon system retreats, HadGEM2-CC exhibits the greatest disagreement over Texas. During the winter season, isolated spurious differences appear, which are evident in all the models.

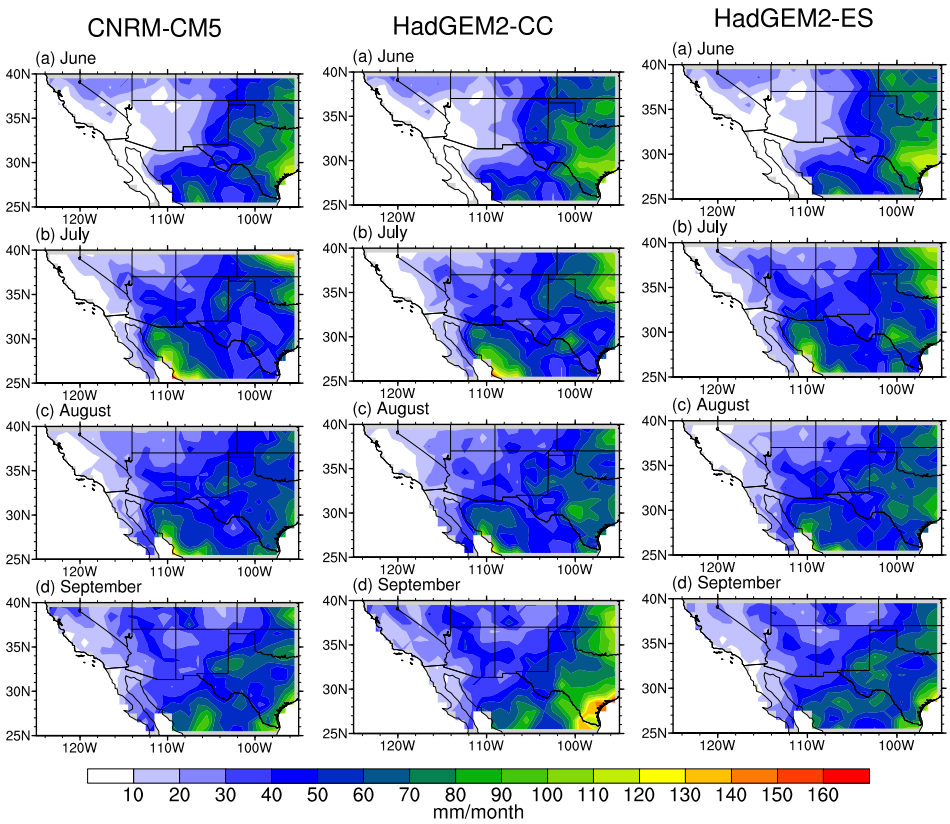


Figure 3.17. Precipitation root mean square difference (mm/month) for (a) June, (b) July, (c) August, and (d) September between downscaled model output and NARR during the monsoon season.

Pearson correlation coefficients between NARR and downscaled model precipitation for the monsoon season are shown in figure 3.18. Unlike temperature correlations, the spatial patterns from all three downscaled models differ in terms of their magnitudes and patterns. CNRM-CM5 exhibits a good agreement with NARR, as also shown in the monthly climatologies. During January, March, and May, a strong positive correlation is demonstrated across the southwest U.S (Appendix 2). Positive correlations in March extend over multiple states across the northern parts of the extended domain, while positive correlations are constrained over the Arizona/New Mexico (AZ/NM)

border during May. Furthermore, the highest (0.82) and lowest (–0.58) correlations are also found during May. During the initial stages of the monsoon season (June), no significant relationship is demonstrated over the core domain (Figure 3.18). However, a negative correlation is demonstrated in July, illustrating that CNRM-CM5 cannot capture the seasonal rains that usually occur during this time period in the southwest U.S. By August and September, little to no significant relationship is seen; however, a significant positive relationship during September over the southwest does show that CNRM-CM5 can capture the decreasing precipitation, or decay phase of the monsoon system. Early fall to mid-winter illustrates a general negative relationship with NARR, again revealing shortcomings in CNRM-CM5.

The downscaled HadGEM2-CC illustrates different correlation patterns when compared to CNRM-CM5 monthly correlations (Figure 3.18, Appendix 2). A positive relationship is found during April over the Chihuahua Desert and Sierra Madre Oriental mountains, similar to CNRM-CM5. Also, it is during this month that HadGEM2-CC demonstrates the highest correlation (0.79). In contrast to CNRM-CM5, the largest negative relationship (–0.68) is demonstrated during the month of July over Oklahoma (Figure 3.18). During July and August, a significant positive signal is illustrated throughout the Sierra Madre Oriental Mountains in HadGEM2-CC, while June and September demonstrate a negative relationship over Texas and western U.S., respectively. Little to no significant correlation are found over the Sierra Madre Occidental and AZ/NM region during the monsoon season, suggesting that the model cannot fully represent precipitation in steep terrain, despite the statistical downscaling.

HadGEM2-ES illustrates a different pattern of correlation when compared to CNRM-CM5 and HadGEM2-CC. Significant positive correlations are demonstrated throughout the Sierra Madre Oriental and Sierra Madre Occidental during the mid-winter and spring months (Appendix 2). However, the maximum significant correlation is in April (0.90), similar to HadGEM2-CC. During the monsoon season, the only significant relationship is illustrated during the month of July over the Sierra Madre Occidental (Figure 3.18), suggesting the model can simulate the interannual precipitation that progress over the complex terrain.

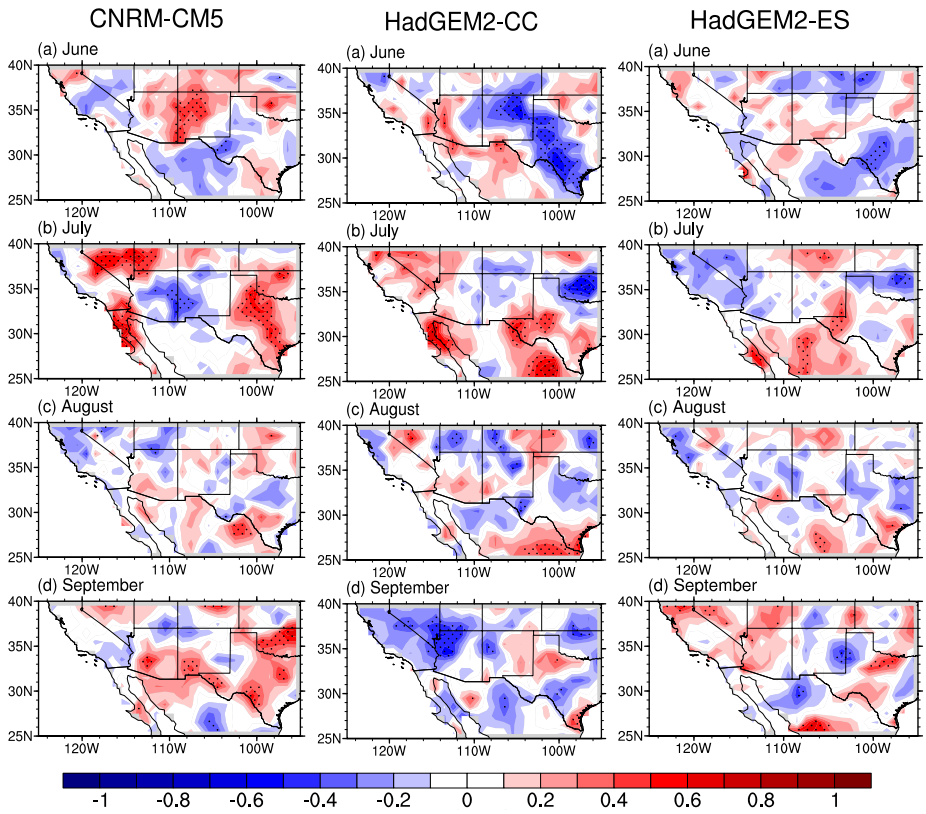


Figure 18. Monthly precipitation correlations for (a) June, (b) July, (c) August, and (d) September between the downscaled model output and NARR with a statistical significance of 90% (stippled).

To further quantify the performance of the downscaled models, monthly area-averaged linear trends for precipitation over the core and extended domain are calculated (Tables 3.3 and 3.4). NARR indicates significant decreases in January (1.33 mm yr^{-1}) and October (1.02 mm yr^{-1}). The drying trend observed in January is partially captured (0.99 mm yr^{-1}) but shows a wetting trend during November (1.02 mm yr^{-1}) for CNRM-CM5. No significant drying trend is evident in either HadGEM2 models.

	NARR											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	-1.33	-0.16	-0.42	-0.34	-0.21	-0.65	-1.26	-1.14	-1.09	-1.02	-0.20	-0.77
P-Value	0.05	0.58	0.19	0.22	0.57	0.31	0.18	0.12	0.14	0.09	0.62	0.36

	CNRM-CM5											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	-0.99	-0.31	0.47	-0.40	0.28	0.81	-0.91	0.66	0.94	0.16	1.02	0.83
P-Value	0.07	0.52	0.14	0.14	0.41	0.21	0.38	0.55	0.35	0.84	0.03	0.38

	HadGEM2-CC											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	0.04	0.32	0.07	-0.05	-0.19	-0.48	-1.28	-0.12	-0.51	-0.42	-0.20	0.16
P-Value	0.96	0.48	0.83	0.80	0.37	0.52	0.16	0.90	0.60	0.52	0.51	0.82

	HadGEM2-ES											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	0.25	-0.12	0.07	-0.18	0.10	1.04	-0.35	-1.42	0.92	0.40	-0.14	0.45
P-Value	0.66	0.80	0.85	0.53	0.77	0.11	0.73	0.16	0.38	0.63	0.76	0.49

Table 3.3. Monthly precipitation trends (mm/month), from 1979 to 1999, for NARR, CNRM-CM5, HadGEM2-CC, and HadGEM2-ES over the core domain. Statistically significant trends (90% level) are shown in salmon shading.

For the extended domain, NARR illustrated no significant wetting or drying trend in any month. However, all three models illustrate significant wetting or drying trends not evident in NARR precipitation, with HadGEM2-ES showing significant precipitation increases in February (0.74 mm yr^{-1}) and decreases in August (0.97 mm yr^{-1}).

	NARR											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	-0.27	0.13	-0.52	0.22	-0.30	0.05	-0.09	0.00	-0.05	-0.42	-0.31	-0.17
P-Value	0.61	0.79	0.23	0.40	0.41	0.88	0.80	0.99	0.87	0.39	0.42	0.74

	CNRM-CM5											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	-0.76	-0.44	-0.16	0.53	0.15	0.23	-0.15	0.70	0.25	-0.87	0.42	1.05
P-Value	0.17	0.27	0.78	0.19	0.70	0.65	0.81	0.31	0.63	0.15	0.31	0.04

	HadGEM2-CC											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	-0.43	0.72	0.67	0.11	0.36	0.07	-1.21	-0.19	-0.63	-0.05	-0.60	0.37
P-Value	0.52	0.24	0.16	0.77	0.38	0.91	0.02	0.75	0.35	0.94	0.13	0.45

	HadGEM2-ES											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Trend (mm/yr)	0.16	0.74	0.59	-0.65	0.12	-0.45	-0.74	-0.97	0.18	-0.27	-0.37	0.04
P-Value	0.80	0.09	0.19	0.12	0.81	0.41	0.23	0.08	0.69	0.66	0.44	0.95

Table 3.4. Monthly precipitation trends (mm/month), from 1979 to 1999, for NARR, CNRM-CM5, HadGEM2-CC, and HadGEM2-ES over the extended domain. Statistically significant trends (90% level) are shown in salmon shading.

Monthly linear precipitation trends for NARR, CNRM-CM5, HadGEM2-CC, and HadGEM2-ES are compared (Figure 3.19). During the monsoon season, NARR

illustrates decreasing precipitation over the Sierra Madre Occidental and Chihuahua Desert, while significant increases in precipitation are established during July over portions of New Mexico. Similarly, significant drying is demonstrated over the Sierra Madre Occidental during October and January (Appendix 2). CNRM-CM5 does capture some of the interannual drying in January, but fails to reproduce any of the significant NARR trend patterns (Appendix 2). During the monsoon season, CNRM-CM5 shows no significant trend until August and September, at which time positive trends appear over Texas/Southern California and Chihuahua desert. Furthermore, positive trends over Mexico and the western U.S. are observed during November and December, a feature not illustrated in the reanalysis. A similar lack of consistent spatial trend patterns is observed for HadGEM2-CC monthly precipitation. No significant drying trend is evident during the monsoon season except for July, when the pattern stretches across the southwest U.S. (Figure 3.19). However, this pattern is not demonstrated during July for NARR precipitation. A significant wetting trend is also seen during February and March, but again is not illustrated in the reanalysis. Finally, HadGEM2-ES monthly precipitation trends establish a similar pattern as the other two downscaled model output (Figure 3.19, Appendix 2). Still, a drying trend over the Chihuahua Desert is revealed in the downscaled model during July, comparable to NARR. However, HadGEM2-ES fails to capture many of the drying patterns throughout the months.

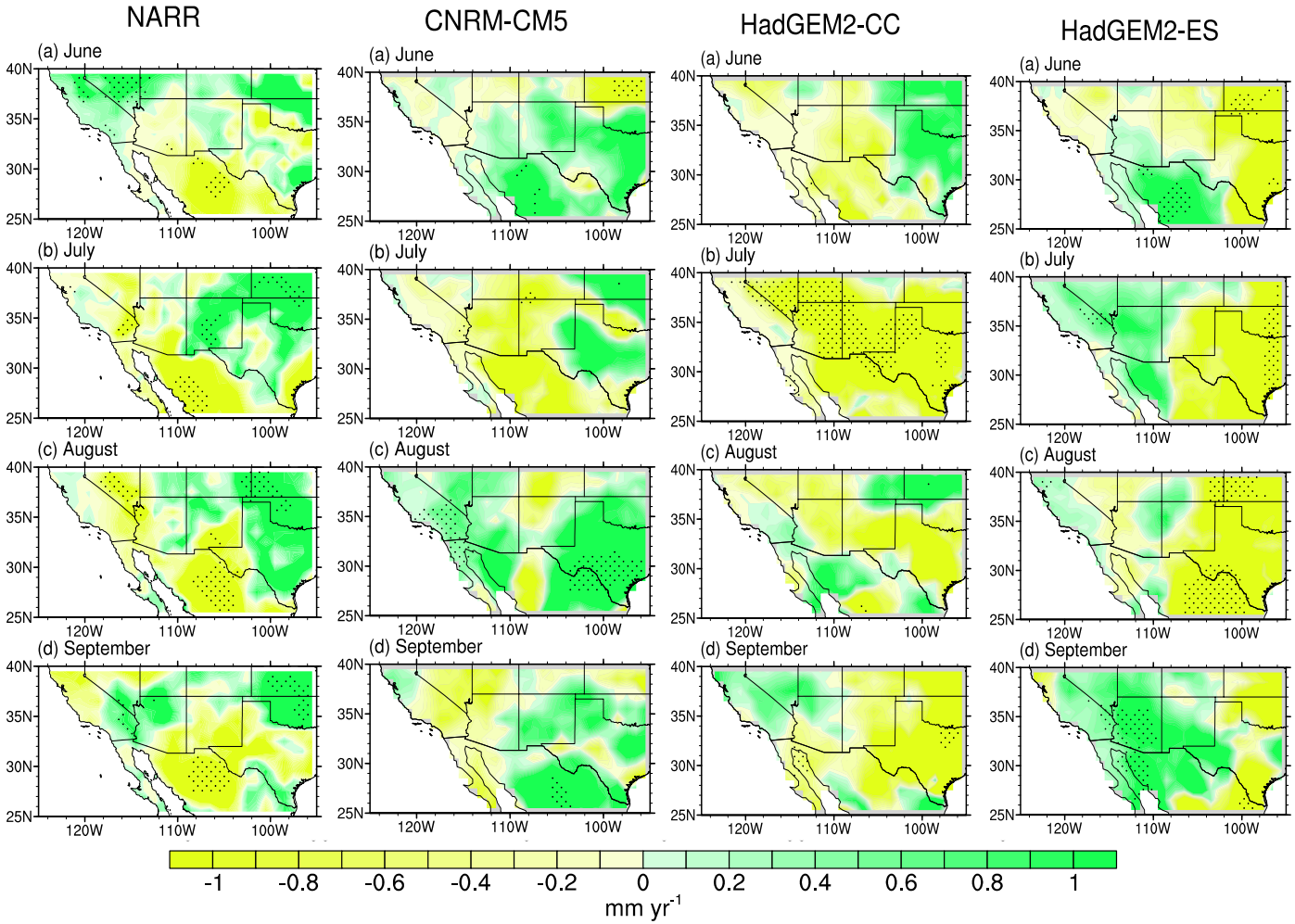


Figure 3.19. Monthly precipitation trends for (a) June, (b) July, (c) August, and (d) September for NARR and downscaled model output. Stippled regions indicate statistically significance (90% level).

III.4.5 Model Evaluation

Finally, to further quantify the downscaled GCMs' ability to accurately simulate seasonal and monthly temperature and precipitation characteristics, skill scores were created and descriptive statistics summarized via box-and-whisker plots. The skill of CNRM-CM5 for temperature shows good agreement with NARR during MAM and SON (0.93, 0.93), as well as for the annual average (Figure 3.20). In comparison, JJA and JJAS show lower skill scores, but still illustrate good agreement. July temperature depicts the lowest skill score of the three models (0.77), also contributing to the lowest seasonal (JJA) skill score. Furthermore, the monsoon-season temperature skill scores illustrate lower values when compared to other seasons.

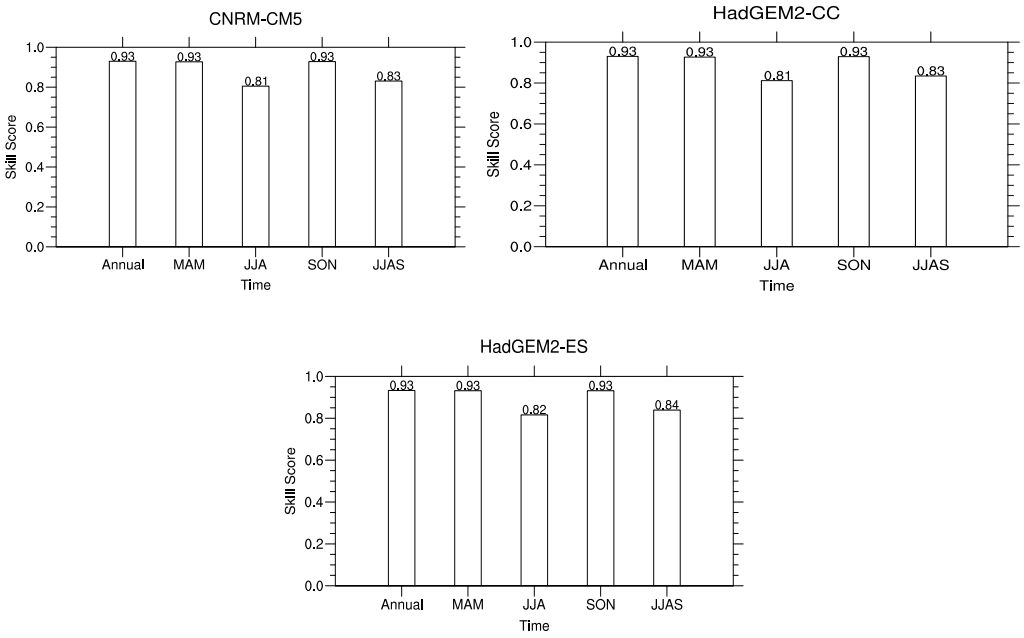


Figure 3.20. Annual and seasonal temperature skill scores for downscaled output from CNRM-CM5, HadGEM2-CC, and HadGEM2-ES.

The skill for precipitation depicts slightly greater agreement during MAM (~0.93) than temperature, but shows overall good agreement with NARR. This general pattern is demonstrated for all downscaled models, meaning good annual and seasonal performance is illustrated in the downscaled model simulations. Monthly precipitation skill scores show May to have the highest agreement with NARR, evident also in the seasonal MAM skill score (Appendix 2). Both CNRM-CM5 and HadGEM2-ES show the lowest agreement during January (0.62, 0.59), whereas HadGEM2-CC exhibits the lowest agreement during December (0.69). However, all three downscaled models depict good agreement with NARR.

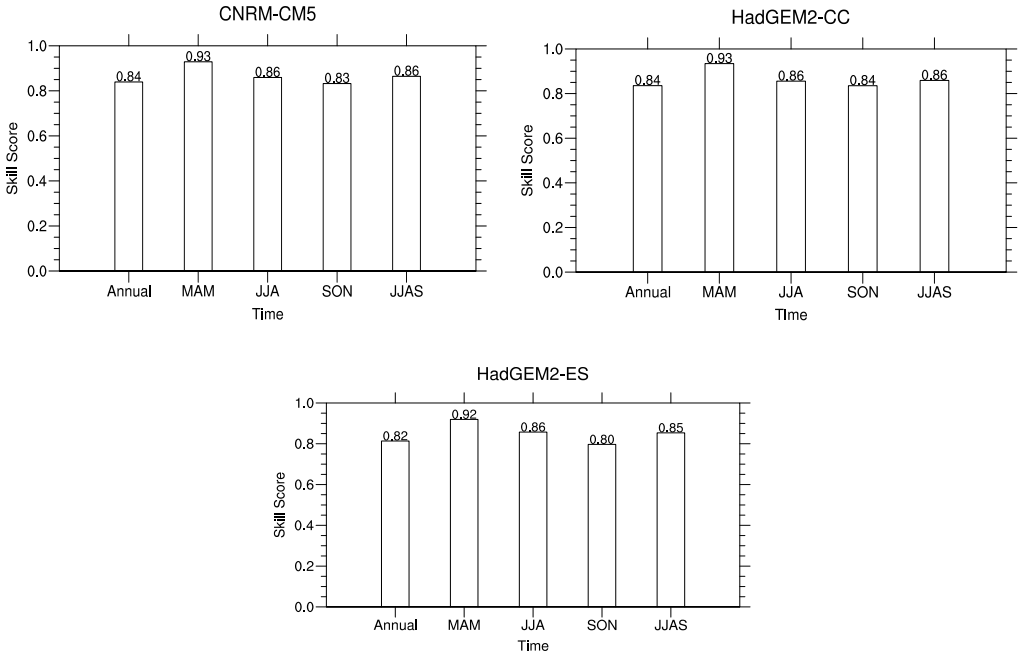


Figure 3.21. Annual and seasonal skill scores for precipitation for downscaled CNRM-CM5, HadGEM2-CC, and HadGEM2-ES.

III.5 Discussion and Conclusion

This objective assessed the ability of statistically downscaled model output (CNRM-CM5, HadGEM2-CC, and HadGEM2-ES) to capture observed NARR precipitation and temperature characteristics from 1979 to 1999, based on numerous spatial and temporal analyses. The results of the analysis illustrate:

1. Seasonal spatial climatologies (MAM, JJA, SON) demonstrate downscaled model output improves the representation of temperature and precipitation when compared to coarsely resolved models.
2. Monthly downscaled model climatologies of temperature and precipitation exhibit a close relationship with NARR temperature and precipitation over the core and extended monsoon domain, demonstrating improved model performance when statistical downscaling is implemented.
3. Interannual variability of downscaled temperature model output shows an improved performance during the monsoon season; however, precipitation remains problematic.
4. Statistical metrics reveal modeled temperatures to be in agreement with NARR, with individual downscaled models illustrating local-scale differences.
5. Model precipitation simulates a wet bias over high terrain for all downscaled models, with greater disagreement compared to temperature.
6. Seasonal skill scores and descriptive statistics for all downscaled model temperatures illustrate good agreement with NARR. Seasonal results for

precipitation show slightly worse agreement, revealing a precipitation bias over the core monsoon region.

Understanding the dynamical processes of the monsoon system and its response to climate change requires GCMs to have accurate model parameterizations and to realistically resolve fine-scale topography. CMIP5 original (coarse resolution) models have difficulty in simulating accurate monsoon characteristics, evident from many previous studies (Cook and Seager 2013, Geil et al., 2013, Liang et al., 2008). Castro et al. (2012) have shown the benefit of using higher resolution models to draw improved conclusions of regional climate variability and prediction. This study provides an assessment of statistically downscaled models for potential use in future climate projections. Based on our analysis, the three downscaled models can resolve historical temperature characteristics in the NAMS region. Different metrics implemented throughout our study to evaluate modeled temperature show good agreement with the reanalysis product; however, the models still overestimate some of the warming trends over the core monsoon domain. Although higher resolution has improved simulated temperatures, the remaining warm bias can partly be attributed to the steep terrain of the Sierra Madre Occidental. This can also be in combination with the timing of the monsoon system that has been shown to have a delayed onset and demise, as well higher precipitation than observations. Simulated terrain also has an impact on the spatial inconsistencies of simulated precipitation. This is evident in the modeled precipitation, but there are improvements over the coarsely resolved precipitation. Geil et al. (2013) have shown that the models that best represent the seasonal shifts of the monsoon ridge

and subtropical highs over the North Pacific and Atlantic tend to have the least issues during the end of the monsoon season. This, in turn, can provide an avenue to guide improvements in representing precipitation features and seasonality. Statistical downscaling provides improvements for understanding the dynamical processes that influence the NAMS. Although this study does not take into account dynamical processes, it is crucial to improve the representation of large-scale features to fully capture the monsoon system. The statistical downscaling technique uses coarse-scale CMIP5 output to create downscaled output, reflecting the biases within precipitation and temperature in our study. We have shown both improvements and limitations to statistically downscaled model output. There remains room for improvement, including even finer scale resolution to better simulate local scale processes; however, the downscaled output serves as a viable tool to investigate NAMS changes for both historical and future time periods.

CHAPTER IV

SUMMARY AND CONCLUSIONS

This research addresses the question, are statistically downscaled GCMs viable tools for assessments of the NAMS? The objectives for our study not only address our overarching question, but also investigate the validity of other widely used datasets. Both reanalyses and GCMs have been utilized extensively within the climate community as research tools to elucidate dynamical processes in a historical and future setting, including for processes such as the NAMS. Reanalyses vary in terms of their input variables, resolution, and model structure, making comparisons between products and studies difficult. Within our study, we focus on temperature and precipitation over the southwest U.S. and western Mexico as they represent key variables within the monsoon system. However, precipitation has been shown to be less of a reliable forecast parameter and is highly model dependent, whereas temperature contains fewer problems as it is assimilated directly into the reanalysis system (Janowiak et al., 2008). The objective for chapter II focuses on the ability of two highly resolved, gridded reanalyses (North American Regional Reanalysis; NARR, European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis; ERA-I) in simulating these two variables in comparison to in-situ observations (Global Historical Climatology Network; GHCN) through goodness-of-fit measurements. GCMs have been incorporated in many studies that focus on understanding historical and future climate, including the Intergovermental Panel on Climate Change Assessment Reports. This information is

then conveyed to a wide range of audiences including water managers and decision makers. The NAMS serves as a key component in the livelihoods of many communities, thus making future climate projections critical for adaptation strategies. A drawback that hinders the confidence in many climate projections is model resolution. Downscaling methodologies, including statistical downscaling, have been an avenue used to obtain fine-scale model results with the goal of capturing local to regional scale processes. Chapter III therefore centers on the evaluation of statistically downscaled model output in capturing temperature and precipitation characteristics related to the NAMS in comparison to NARR, and to their coarse counterparts.

IV.1 Assessment of Reanalysis Products

Two reanalysis products, the NARR and ERA-I, are compared with GHCN station based observations to assess their ability to simulate temperature and precipitation over the monsoon-affected region. Previous studies have utilized both products in historical and future simulations of the NAMS, including downscaling methodologies (Chan and Misra, 2011, Torres-Alavez et al., 2014). Both products are temporally complete and are initiated from 1979; however, NARR is on a finer resolution (0.3°) that spans North America. In comparison, ERA-I is on a larger grid resolution (0.75°) but spans the entire globe. Furthermore, both products differ in their precipitation estimates, as ERA-I uses microphysical processes in clouds, whereas NARR produces precipitation through assimilation of numerous precipitation products.

The ability of the each reanalysis to simulate precipitation and temperature from 1980 to 2010 was compared with in-situ observations an extended domain encompassing the southwest United States and west Mexico (125°W-96°W, 15°N-42°N). Different goodness-of-fit measures, such as Pearson's correlation coefficient (r), root mean square error (RMSE), and index of agreement (d), were used to investigate the association of each climate variable with observations at different spatiotemporal scales. In general, simulated temperatures for both reanalyses showed similar agreement throughout the study region. In comparison, precipitation agreement for the reanalyses varied, with NARR illustrating better agreement with observations than ERA-I. Error patterns over the Arizona, New Mexico, Colorado, and Utah demonstrate the largest biases relative to observations over this region, as well as large biases over steep terrain of the Sierra Madre Occidental. Best case precipitation agreement, based on R and RMSE, demonstrate NARR to generally better capture the interannual and seasonal variability in comparison to ERA-I, while worst case scenarios illustrate a large bias in precipitation and temperature for both reanalyses.

Overall, the results found of the first objective demonstrate a similar temperature agreement for ERA-I and NARR, while simulated precipitation is better captured by NARR. Precipitation demonstrates varying degrees of agreement in both NARR and ERA-I, signifying the difficulties and errors associated with simulating precipitation. However, monsoon season (JJAS) patterns show NARR to have better performance when compared to ERA-I. Therefore, NARR was used to represent spatially and temporally complete observations for the subsequent model evaluation analyses.

IV.2 Assessment of Downscaled CMIP5

Statistically downscaled models were compared with a complete reanalysis based on the results obtained in Chapter II. Statistical downscaling provides a means to interpolate coarse-scaled model output biases to a finer resolution. Based on previous literature describing simulations of the NAMS (Cook and Seager, 2013, Geil et al., 2013), three models were chosen: CNRM-CM5, HadGEM2-CC, and HadGEM2-ES. Two domains were identified for evaluation including a core (112°W - 102°W to 25°N-33°N) and extended region (125°W - 95°W to 25°N - 40°N) to examine temperature and precipitation characteristics that are associated with the NAMS. Various statistical metrics and reanalysis comparisons, including root-mean-square difference (RMSD), correlation coefficients (R), and linear regressions, were quantified over the historical time period (1979 – 1999).

Statistically downscaled model output demonstrates an improved representation of temperature and precipitation when compared to its coarsely resolved counterpart. Results for model climatologies over the core and extended domains capture the seasonality that is associated with the NAMS, including the rapid precipitation onset in June and the decay phase in September. Furthermore, the magnitudes of temperature during the summer months are well captured in the downscaled models; however, the interannual variability remains problematic for both simulated variables. Also, modeled temperature and precipitation illustrate a warm and wet bias over the Sierra Madre Occidental for all three models, indicating continued difficulties in simulating observations over complex terrain. Descriptive statistics and seasonal skill scores

showed the agreement between modeled temperatures and reanalysis was better captured than precipitation.

Overall, our objective established that statistically downscaled model output some of the limitations coarse scaled models illustrated. However, the errors over complex terrain still exist, leading to increased precipitation and temperature within the core monsoon domain. Model resolution has continued to be an issue in capturing regional to local scale processes that impact precipitation and temperature characteristics. Thus, our results reveal the need for improvements in simulating complex terrain to better capture precipitation characteristics. However, addressing the issue of high terrain is limited, as correct parameterization in simulating precipitation features is key within GCMs. Improved sub-grid scale processes, in combination with finer resolution, will allow for NAMS variability to be better simulated and thus provide greater confidence in future climate projections. Statistical downscaling provides an avenue to improve coarse-scaled model output to a finer resolution.

IV.3 Final Conclusions

Our study assessed two different sets of products that have been extensively utilized within the climate community. Two reanalysis datasets and three statistically downscaled models were assessed to evaluate their ability in capturing temperature and precipitation characteristics associated with the NAMS. Results suggest that temperature is far simpler to simulate in comparison to precipitation; however, complex terrain weakens both the reanalysis and downscaled models' abilities to simulate the climate

variables of interest. Nevertheless, statistically downscaled models provide an avenue for improved future climate projections as they have addressed various spatiotemporal biases in coarse scaled model output. GCM projections have shown a systematic delay in the onset, peak, and decay of the NAMS due to enhanced convective barriers (Cook and Seager, 2013; Grantz et al., 2007). Furthermore, Cook et al. (2015) have recently revealed a climatic shift with respect to the last millennium, using a combination of drought reconstruction and 17 CMIP5 GCMs. The mean state of drought in the late 21st century over the Southwest and Central Plains will likely exceed severe megadrought periods of the medieval era in both high and moderate emission scenarios (Cook et al., 2015). These severe droughts can have devastating consequences to the water resources throughout the Southwest, placing a substantial burden on water managers and decision makers. Thus, the use of statistically downscaled models can serve as a viable tool in expanding the uncertainty and knowledge gaps concerning the NAMS, including the need for improved climate projections to aid in future adaptation.

REFERENCES

Adams, D.K. and Comrie, A.C. 1997. The North American monsoon, *Bull. Am. Meteorol. Soc.,* **78**: 2197–2213.

Anderson, B.T., Wang, J., Salvucci, G., Gopal, S., and Islam, S. 2010. Observed trends in summertime precipitation over the southwestern United States. *J. Climate*, **23:** 1937–1944.

Bechtold, P., Chaboureau, J.P., Beljaars, A.C.M., Betts, A.K., Köhler, M., Miller, M., and Redelsperger, J-L. 2004. The simulation of the diurnal cycle of convective precipitation over land in a global model. *Q. J. R. Meteorol. Soc.,* **130**: 3119–3137.

Bulmer, M.G. 1979. *Principles of Statistics*. New York, NY: Dover, 252 pp.

Brunke, M.A., Wang, Z., Zeng, X., Bosilovich, M., and Shie, C. –L. 2011. An assessment of the uncertainties in ocean surface turbulent fluxes in 11 reanalysis, satellite-derived, and combined global data sets. *J. Climate*, **24:** 5469-5493.

Castro, C.L., McKee, T.B., and Pielke Sr., R.A. 2001. The Relationship of the North American Monsoon to Tropical and North Pacific Sea Surface Temperatures as Revealed by Observational Analyses. *J. Climate*, **14**: 4449–4473.

Castro, C.L., Chang, H., Dominguez, F., Carillo, C., Schemm, J., and Juang, H.H. 2012. Can a regional climate model improve the ability to forecast the North American monsoon? *J. Climate*, **25:** 8212–8237.

Cavazos, T. and Arriaga-Ramirez, S. 2012. Downscaled Climate Change Scenarios for Baja California and the North American Monsoon during the Twenty-First Century. *J. Climate*, **25:** 5904–5915.

Chai, T. and Draxler, R.R. 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, *Geoscientific Model Development*, **7:** 1247-1250.

Chen, L. and Frauenfeld, O.W. 2014. Surface Air Temperature Changes over the 20th and 21st Centuries in China Simulated by 20 CMIP5 Models. *J. Climate*, **27:** 3920–3937.

Christensen, J.H., Krishna Kumar, K., Aldrian, E., An, S. –I., Cavalcanti, I.F.A., de Castro, M., Dong, W., Goswami, P., Hall, A., Kanyanga, J.K., Kitoh, A., Kossin, J., Lau, N. –C., Renwick, J., Stephenson, D.B., Xie, S. –P., and Zhou, T. 2013. Climate Phenomena and their Relevance for Future Regional Climate Change. *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, 1217–1308.

Chuang, H.Y., Manikin, G., and Treadon, R. E. 2001. The NCEP Meso ETA Model Post Processor: A Documentation. *NCEP Office Note*, **438:** 52.

Cook, B. I. and Seager, R. 2013. The response of the North American Monsoon to increased greenhouse gas forcing. *J. Geophys. Res.*, **118:** 1690–1699.

de Leeuw, J., Methven, J., and Blackburn, M. 2014. Evaluation of ERA-Interim reanalysis precipitation products using England and Wales observations. *Q.J.R. Meteorol. Soc.*, **141:** 798–806.

Dee, D.P., Balmaseda, M., Balsamo, G., Engelen, R., Simmons, A.J., and Thépaut, J. N. 2014. Toward a Consistent Reanalysis of the Climate System. *Bull. Amer. Meteor. Soc.*, **95**: 1235–1248.

Fall, S., Niyogi, D., Gluhovsky, A., Pielke Sr., R.A., Kalnay, E., and Rochon, G. 2010. Impacts of land use land cover on temperature trends over the continental United States: Assessment using the North American Regional Reanalysis. *Int. J. Climatol.*, **30:** 1980–1993.

Favors, J. E. and Abatzoglou, J.T. 2012. Regional Surges of Monsoonal Moisture into the Southwestern United States. *Mon. Wea. Rev.*, **141:** 182–191.

Flato, G., Marotzke, J., Abioudun, B., Braconnot, P., Chou, S.C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilvardi, E., Jakob,

C., Kattsov, V., Reason, C., and Rummukainen, M. 2013. Evaluation of climate models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Eds. Cambridge University Press, 741-882.

Gutzler, D.S., Long, L.N., Schemm, J., Baidya Roy, S., Bosilovich, M., Collier, J.C., Kanamtsu, M., Kelly, P., Lawrence, D., Lee, M.I., Lobato Sanchez, R., Mapes, B., Mo, K., Nunes, A., Ritchie, E.A., Roads, J., Schubert, S., Wei, H., and Zhang, G.J. 2009. Simulations of the 2004 North American Monsoon: NAMAP 2. *J. Climate*, **22**: 6716–6740.

Geil, K., Serra, Y. L., and Zeng, X. 2013. Assessment of CMIP5 Model Simulations of the North American Monsoon System. *J. Climate,* **26:** 8787–8801.

Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona, 2007: Seasonal Shifts in the North American Monsoon, *J. Climate*, **20:** 1923–1935.

Higgins, R.W. 2000. Overview of the North American Monsoon Experiment. *American Meteorological Society Extended Abstract*, 1–9.

Higgins, R.W and Shi, W. 2001. Intercomparison of the Principal Modes of Interannual and Intraseasonal Variability of the North American Monsoon System. *J. Climate*, **14:** 403–417.

Janowiak, J.E., Gruber, A., Kondragunta, C.R., Livezey, R.E., and Huff-man, G.J. 1998. A comparison of NCEP–NCAR reanalysis precipi- tation and the GPCP rain gauge–satellite combined dataset with observational error considerations. *J. Climate,* **11:** 2960–2979.

Ji, L. and Gallo, K. 2006. An agreement coefficient for image comparison. *Photogrammetric Engineering and Remote Sensing,* **72:** 823–833.

Kumar, S. and Merwade, V. 2011. Evaluation of NARR and CLM3.5 outputs for surface water and energy budgets in the Mississippi River Basin. *J. Geophys. Res.,* **116:** 1–21.

Langford, S., Stevenson, S., and Noone, D. 2014. Analysis of Low-Frequency Precipitation Variability in CMIP5 Historical Simulations for Southwestern North America. *J. Climate*, **27:** 2735–2756.

Lawrimore, J.H., Menne, M. J., Gleason, B.E., Williams, C.N., Wuertz, D.B., Vose, R.S., and Renniel, J. 2011. An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *J. Geophys. Res.*, **116:** D19121.

Legates, D. R. and Davis, R.E. 1997. The continuing search for an anthropogenic climate change signal: Limitations of correlation-based approaches. *J. Geophys. Res.*, **24:** 2319–2322.

Legates, D. R and McCabe, G.J. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.*, **35:** 233–241.

Liang, X-Z., Zhu, J., Kunkel, K.E., Ting, M., and Wang, J.X.L. 2008. Do CGCMs simulate the North American monsoon precipitation seasonal-interannual variability? *J. Climate*, **21:** 4424–4448.

Lin, J., Mapes, B.E., Weickmann, K.M., and Kiladis, G.N. 2008. North American Monsoon and Convectively Coupled Equatorial Waves Simulated by IPCC AR4 Coupled GCMs. *J. Climate*, **21:** 2919–2937.

Lin, R., Zhou, T., and Qian, Y. 2014. Evaluation of global monsoon precipitation changes based on five reanalysis datasets. *J Climate*, **27:** 1271–1289.

Ma, L., Zhang, T., Frauenfeld, O. W., Ye, B., Yang, D., and Qin, D. 2009. Evaluation of Precipitation from the ERA-40, NCEP-1, and NCEP-2 Reanalyses, and CMAP-1, CMAP-2, and GPCP-2 with Ground-based Measurements in China. *J. Geophys. Res.*, **114**: D09105.

Maloney, E., Camargo, S., Chang, E., Colle, B., Fu, R., Geil, K., Hu, Q., Jiang, X., Johnson, N.,, Karnauskas, K., Kinter, J., Kirtman, B.,, Kumar, S., Langenbrunner, B., Lombardo, K., Long, L., Mariotti, A., Meyerson, J., Mo, K.,

Neelin, D., Pan, Z., Seager, R., Serra, Y., Seth, A., Sheffield, J., Stroeve, J., Thibeault, J., Xie, S., Wang, C., Wyman, B., and Zhao, M. 2013. North American Climate in CMIP5 experiments: part III: Assessment of 21st century projections. *J. Climate,* **27:** 2230–2270.

Martin, G.M. and Levine, R.C. 2012. The influence of dynamic vegetation on the present-day simulation and future projections of the South Asian summer monsoon in the HadGEM2 family. *Earth Syst. Dynam.,* **3:** 245–261.

Martin, G.M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J., Verhoef, A., Vidale, P. L., and Wiltshire, A. 2011. The HadGEM2 family of Met Office Unified Model climate configurations, *Geosci. Model Dev*., **4:** 723–757.

Maurer, E.P., Wood, A.W., Adam, A.C., Lettenmaier, D.P., and Nijssen, B. 2002. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States. *J. Climate*, **15**: 3237–3251.

Menne, M.J. and Williams Jr., C.N. 2005. Detection of undocumented change points using multiple test statistics and composite reference series. *J. Climate,* **18:** 4271–4286.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W. 2006. North American Regional Reanalysis. *Bull. Amer. Meteor. Soc*., **87:** 343–360.

NOAA/NWS, 2004: The North American Monsoon. Reports to the Nation on our Changing Planet. NOAA/National Weather Service.

Peterson, T.C. and Vose, R.S. 1997. An Overview of the Global Historical Climatology Network Temperature Database. *Bull. Amer. Meteor. Soc.,* **78:** 2837–2849.

Polade, S.D., Gershunov, A., Cayan, D.R., Dettinger, M.D., and Pierce, D.W. 2013. Natural Climate Variability and Teleconnections to Precipitation over the Pacific-North American Region in CMIP3 and CMIP5 Models. *J. Geophys. Res.,* **40:** 1–6.

Reclamation, 2013. Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with preceding Information, and Summary of User Needs. Prepared by the U.S. Department of the Interior, Bureau of Reclamation, 47.

Ropelewski, C.F. and Halpert, M.S. 1987. Global and Regional Scale Precipitation Associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.,* **115:** 1606–1626.

Ropelewski, C.F., Gutzler, D.S., Higgins, R.W., and Mechoso, C.R. 2005. The North American monsoon system. The Global Monsoon System: Research and Forecasts. WMO Tech Doc. WMO TD-1266. Tropical Meteorology Resaerch Program Rep. **70:** 207–218.

Seager, R., Ting, M., Held, I., Kushnir, Y., Lu, J., Vecchi, G., Huang, H., Harnik, N., Leetmaa, A., Lau, N., Li, C., Velez, J., and Naomi, N. 2007. Model Projections of an Imminent Transition to a More Arid Climate in Southwestern North America. *Science,* **316:** 1181–1184.

Seth, A., Rauscher, S.A., Rojas, M., Giannini, A., and Camargo, S.J. 2011. Enhanced Spring Convective Barrier for Monsoons in a Warmer World? *Clim. Chan.*, **104:** 403–414.

Sheffield, J., Barrett, A.P., Colle, B., Fernando, D.N., Fu, R., Geil, K.L., Hu, Q., Kinter, J., Kumar, S., Langenbrunner, B., Lombardo, K., Long, L.,N., Maloney, E., Mariotti, A., Meyerson, J.E., Mo, K.C., Nellin, J.D., Nigam, S., Pan, Z., Ren, T., Ruiz-Barradas, A., Serra, Y.L., Seth, A., Thibeault, J.M., Stroeve, J.C., Yang, Ze., and Yin, L. 2013. North American Climate in CMIP5 Experiments. Part I:

Evaluation of Historical Simulations of Continental and Regional Climatology. *J. Climate*, **26:** 9209–9245.

Song, F. and Zhou, T. 2014. Interannual variability of East Asian summer monsoon simulated by CMIP3 and CMIP5 AGCMs: Skill dependence on Indian Ocean—Western Pacific Anticyclone Teleconnection, *J. Climate*, **27:** 1679–1697.

Taylor, K. E. 2001. Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, **106:** 7183–7192.

Tompkins, A.M., Gierens, K., and Rädel, G. 2007. Ice supersaturation in the ECMWF integrated forecast system. *Q. J. R. Meteorol. Soc.*, **133**: 53–63.

Torres-Alavez, A., Cavazos, T., and Turrent, C. 2014: Land- Sea Thermal Contrast and Intensity of the North American Monsoon under Climate Change Conditions. *J. Climate*, **27:** 4566–4580.

Vera, C., Higgins, W., Amador, J., Ambrizzi, T., Garreaud, R., Gochis, D., Gutzler, D., Lettenmaier, D., Marengo, J., Mechoso, C.R., Nogues-Paegle, J., Silva Dias, P.L., and Zhang, C. 2006. Toward a unified view of the American monsoon systems. *J. Climate*, **19:** 4977–5000.

Voldaire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M. –P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F. 2012. The CNRM-CM5.1 global climate model: description and basic evaluation. *Clim. Dyn.*, **40:** 2091–2121.

Willmott, C.J. 1984. On the evaluation of model performance in physical geography. In *Spatial Statistics and Models*, 443–460.

Willmott, C.J and Matsuura, K. 2005. Advantage of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, **30:** 79-82.

Willmott, C.J, Robeson, S.M., and Matsuura, K. 2012. A refined index of model performance. *Int. J. Climatol.,* **32:** 2088–2094.

Zehnder, J.A. 2004. Dynamic Mechanisms of the Gulf Surge. *J. Geophys. Res.,* **109:** 1–14.

APPENDIX 1

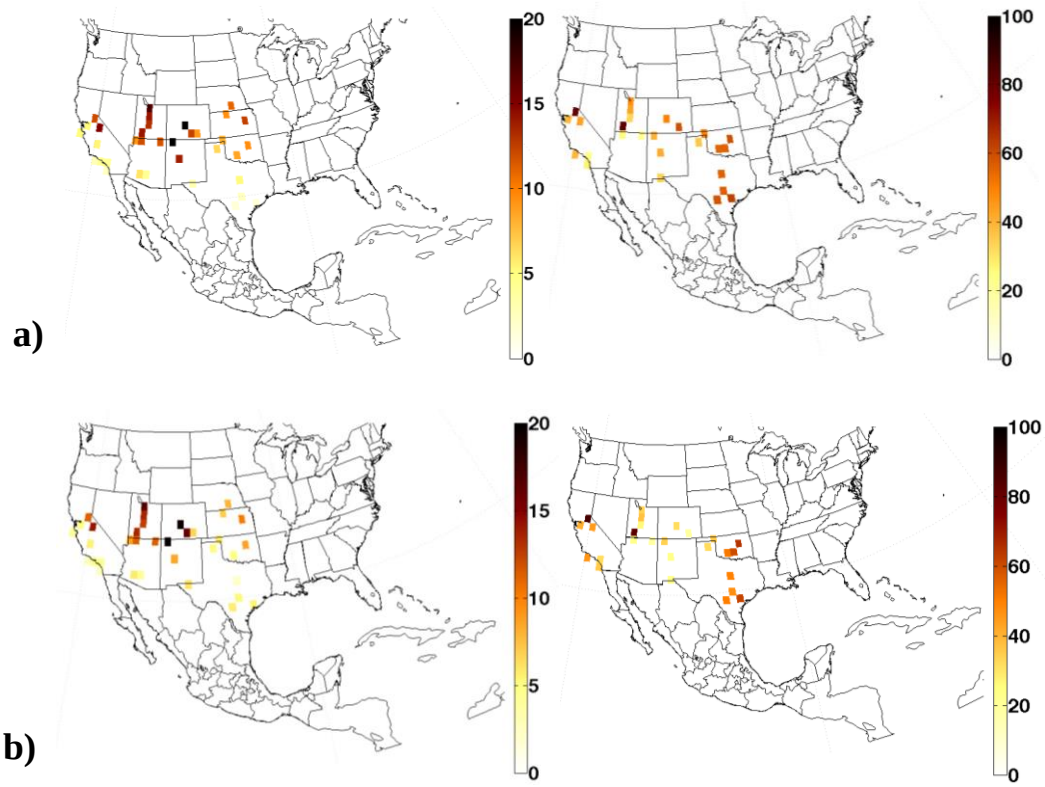


Figure A1-1. Annual temperature (left; °C) and precipitation (right; mm/month) MAE for NARR (a) and ERA-I (b).

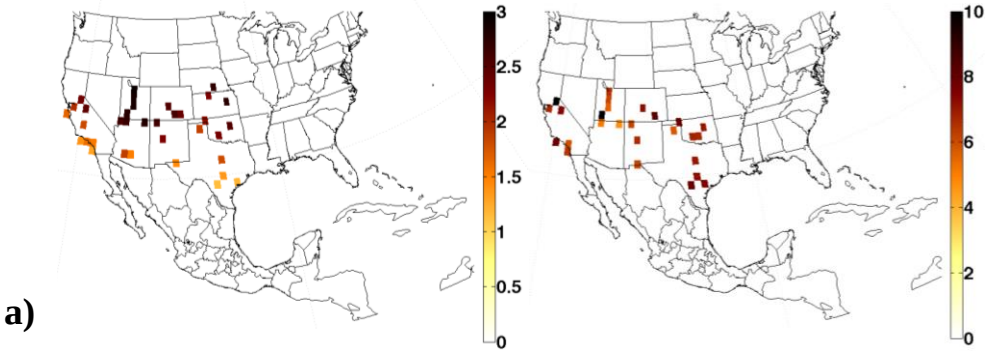


Figure A1-2. Annual temperature (left; °C) and precipitation (right; mm/month) SDB for NARR (a) and ERA-I (b).

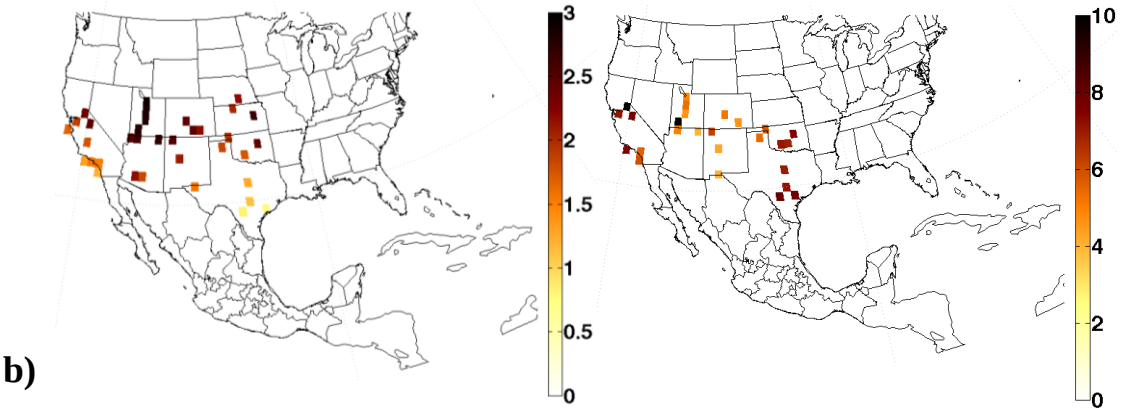


Figure A1-2 Continued. Annual temperature (left; °C) and precipitation (right; mm/month) SDB for NARR (a) and ERA-I (b).

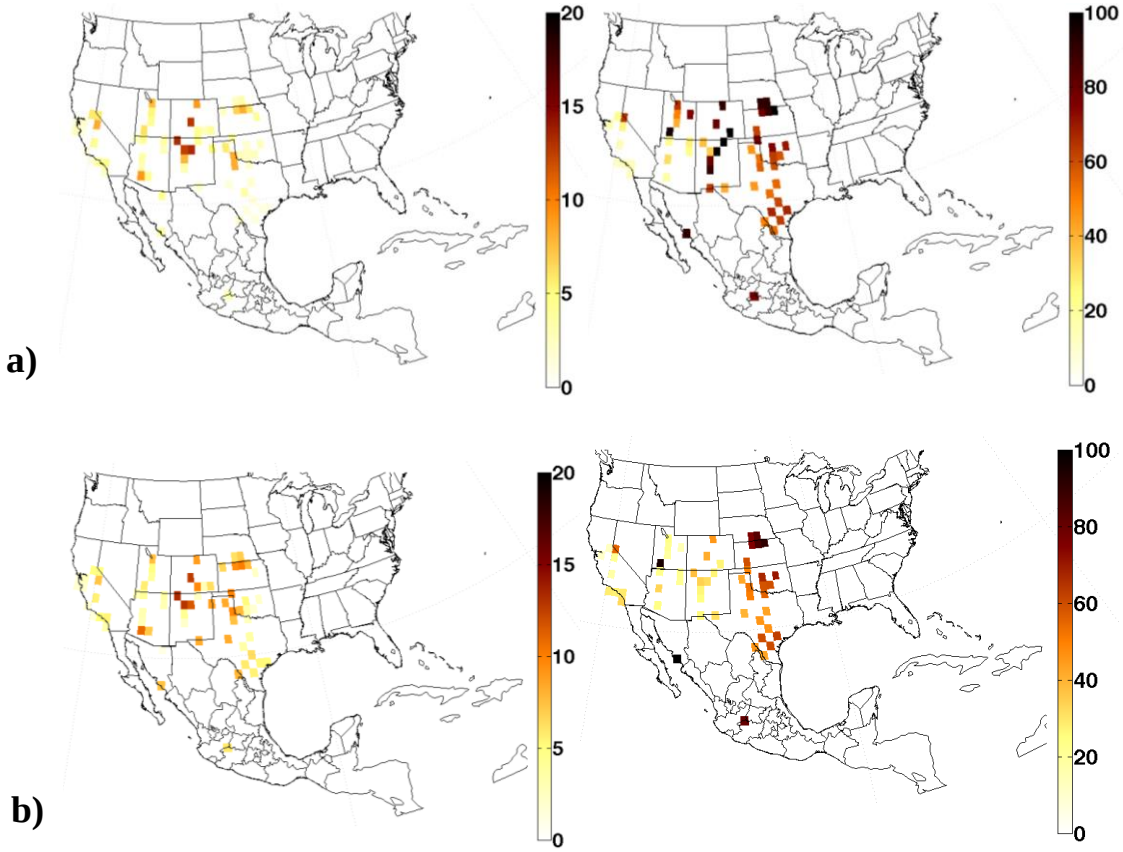


Figure A1-3. Seasonal temperature (left; °C) and precipitation (right; mm/month) MAE for NARR (a) and ERA-I (b).

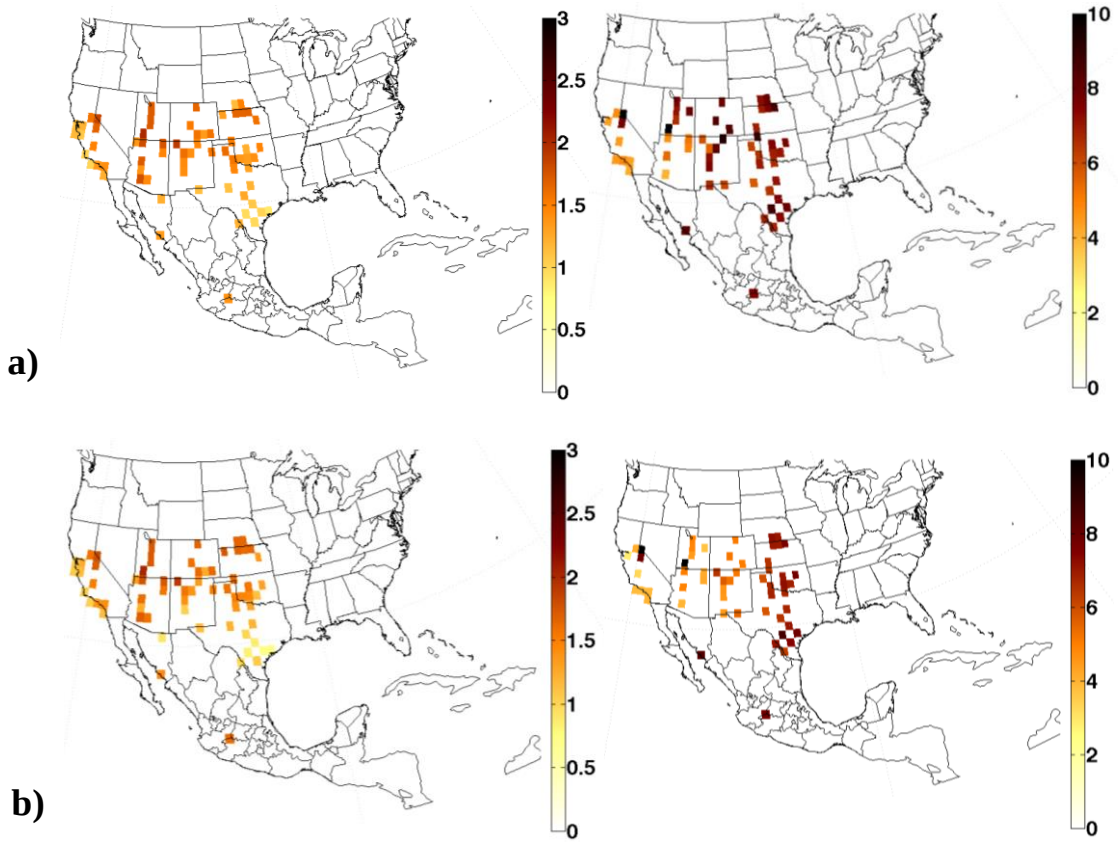


Figure A1-4. Seasonal temperature (left; °C) and precipitation (right; mm/month) SDB for NARR (a) and ERA-I (b).

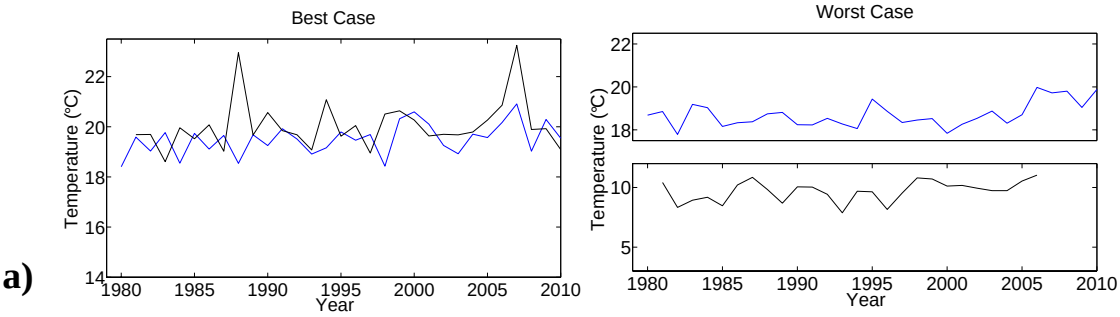


Figure A1-5. Best (left) and worst (right) case annual temperature grid cell time series for NARR (a) and ERA-I (b) using correlation coefficients as the determinant.

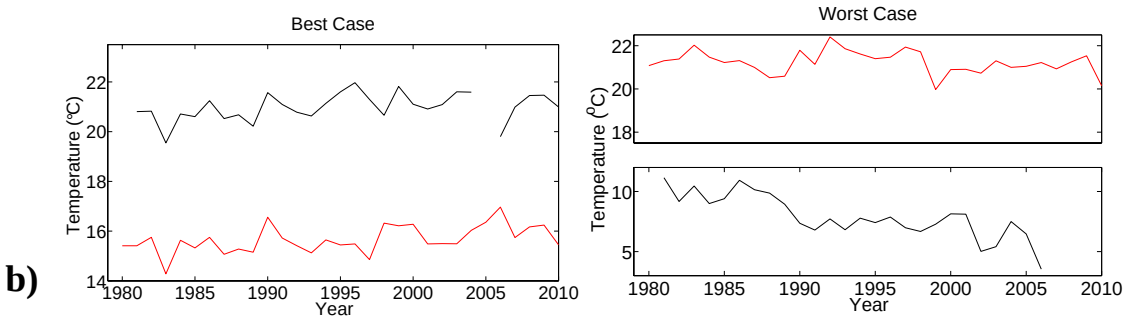


Figure A1-5 Continued. Best (left) and worst (right) case annual temperature grid cell time series for NARR (a) and ERA-I (b) using correlation coefficients as the determinant.

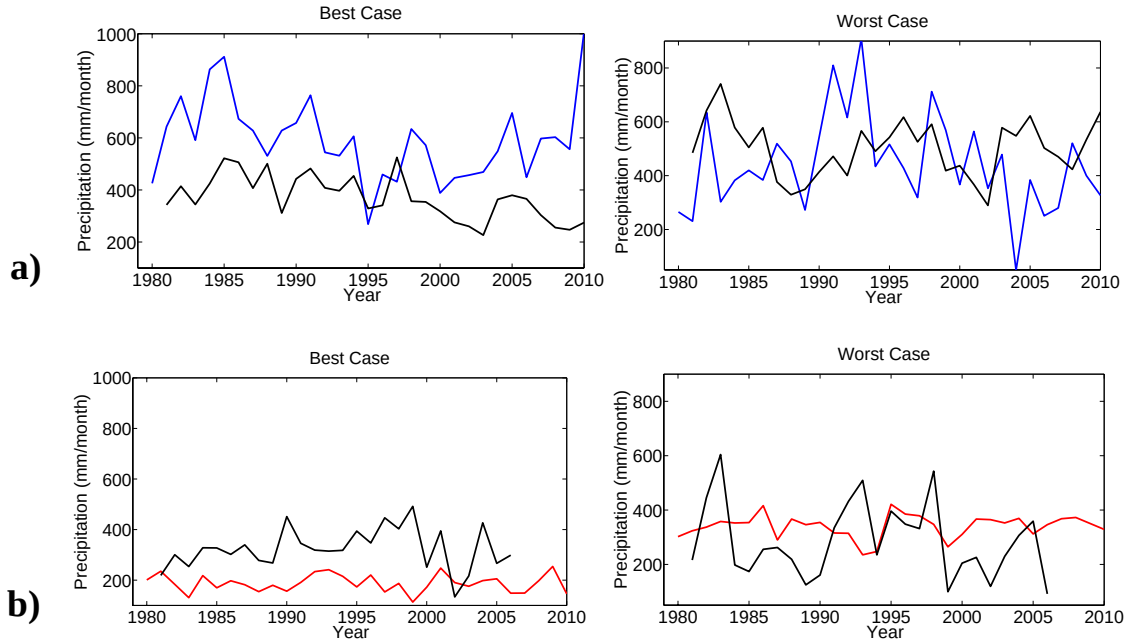


Figure A1-6. Best (left) and worst (right) case annual precipitation grid cell time seriess for NARR (a) and ERA-I (b) using correlation coefficients as the determinant.

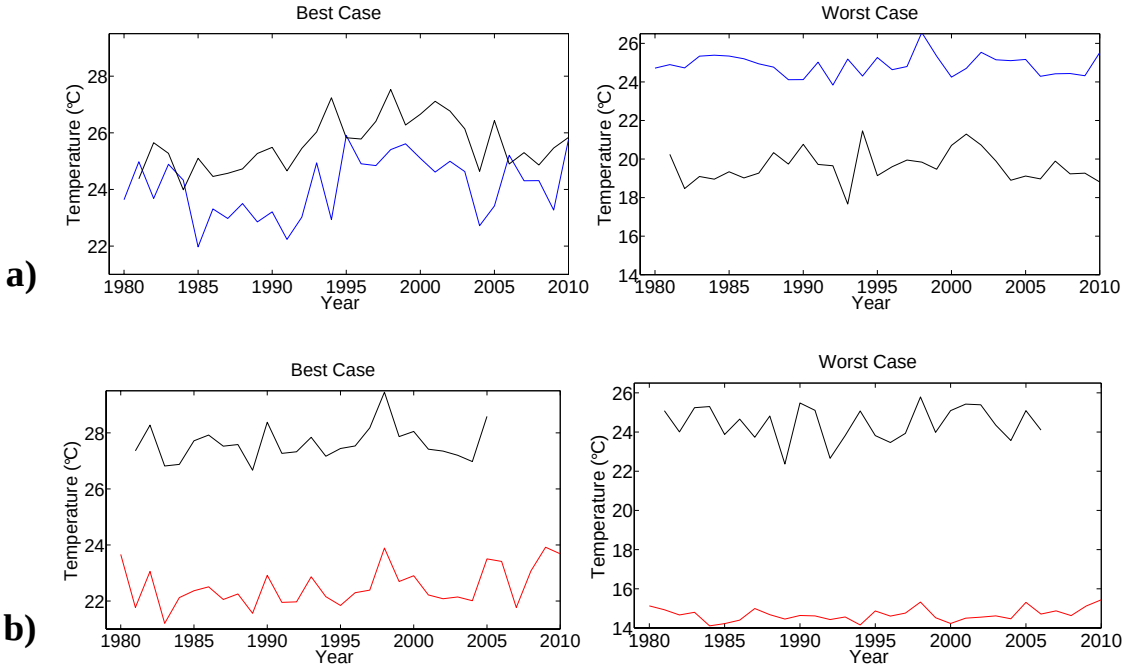


Figure A1-7. Best (left) and worst (right) case seasonal temperature grid cell time series for NARR (a) and ERA-I (b) using correlation coefficients as the determinant.

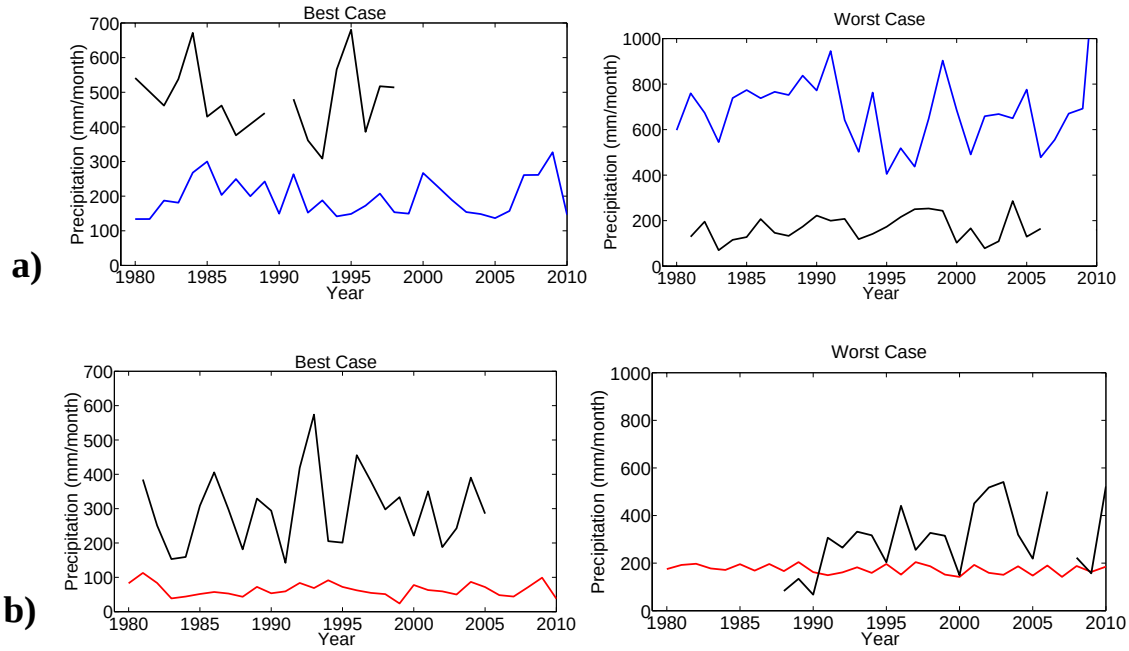


Figure A1-8. Best (left) and worst (right) case seasonal precipitation grid cell time series for NARR (a) and ERA-I (b) using correlation coefficients as the determinant.

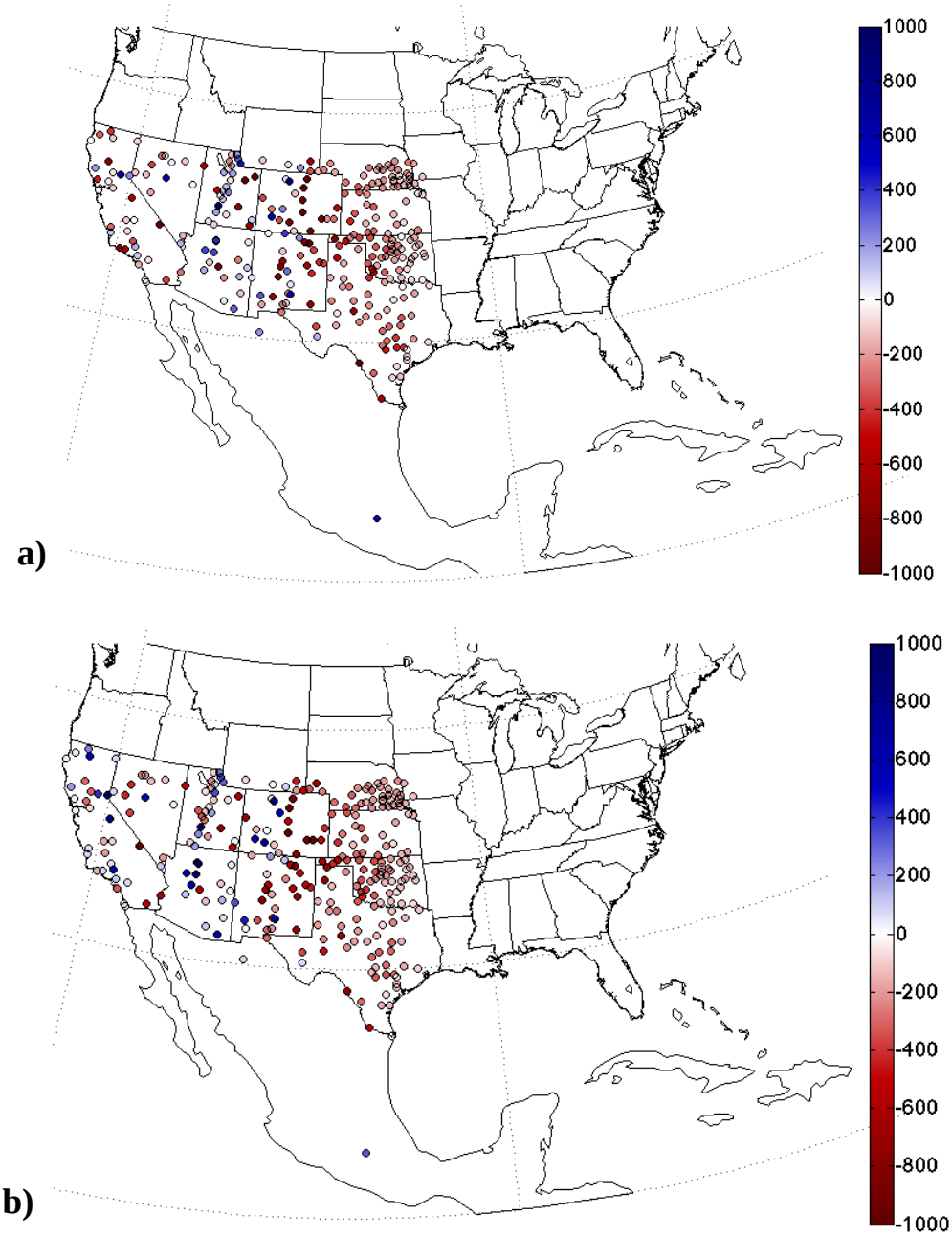


Figure A1-9. Elevation differences from NARR (a) and ERA-I (b) against GHCN observations in meters. Positive differences illustrate higher elevations while negative differences show lower elevations in the reanalyses.

APPENDIX 2

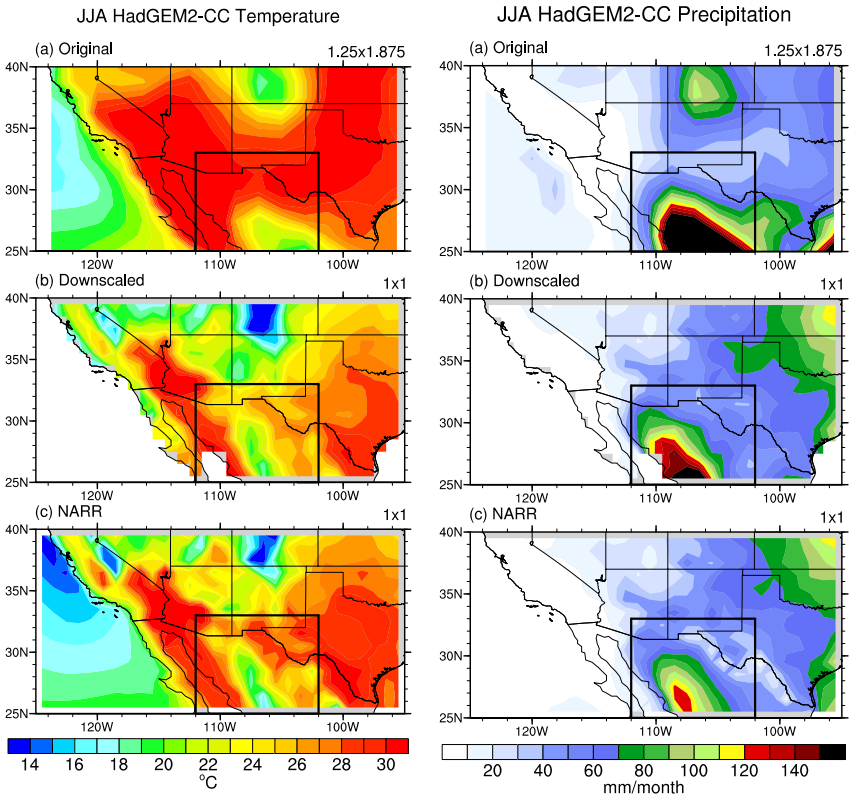


Figure A2-1. JJA temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled HadGEM2-CC output, (b) original HadGEM2-CC output, and (c) NARR observations from 1979 to 1999.

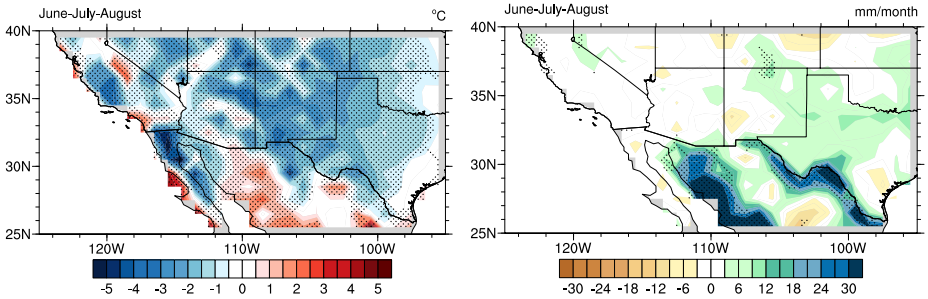


Figure A2-2. JJA temperature (left) and precipitation (right) spatial differences between downscaled HadGEM2-CC output and NARR with 90% statistical significance (stippled).

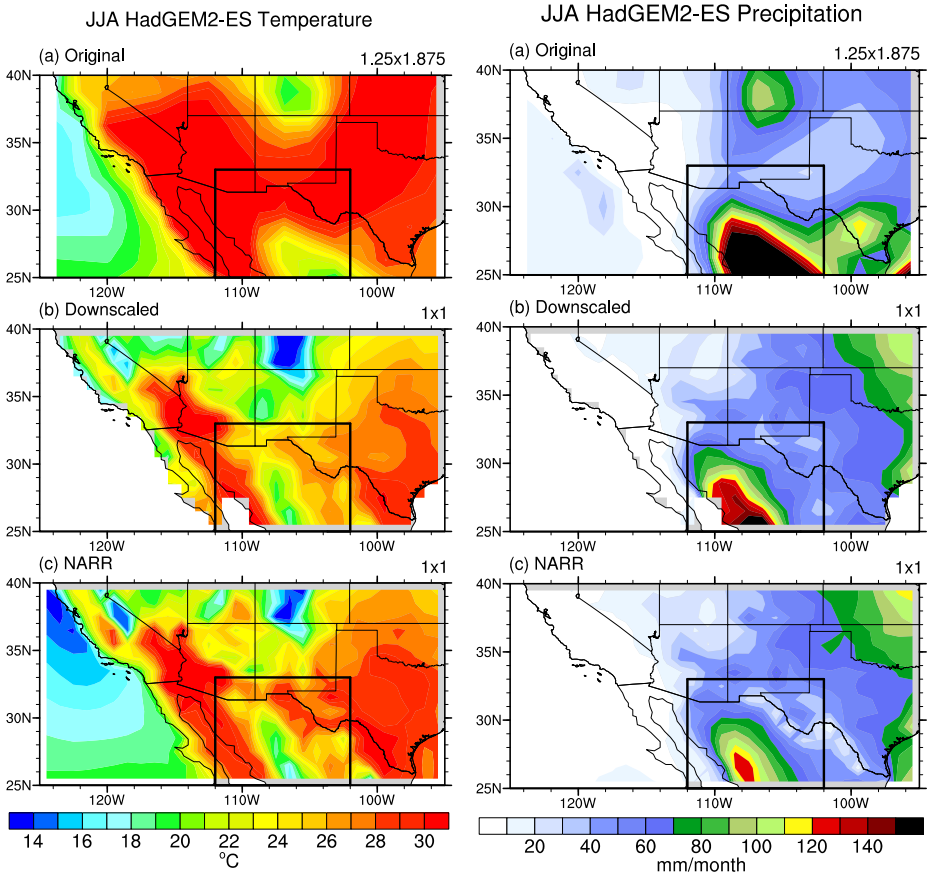


Figure A2-3. JJA temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled HadGEM2-ES output, (b) original HadGEM2-ES output, and (c) NARR observations from 1979 to 1999.

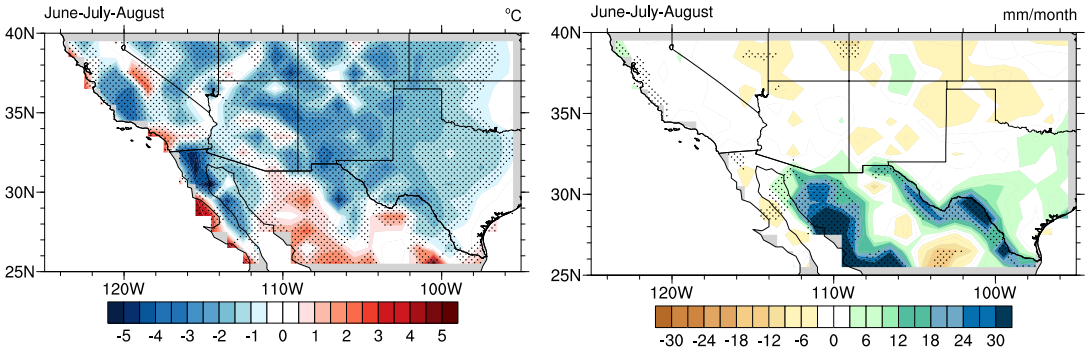


Figure A2-4. JJA temperature (left) and precipitation (right) spatial differences between downscaled HadGEM2-ES output and NARR with 90% statistical significance (stippled).

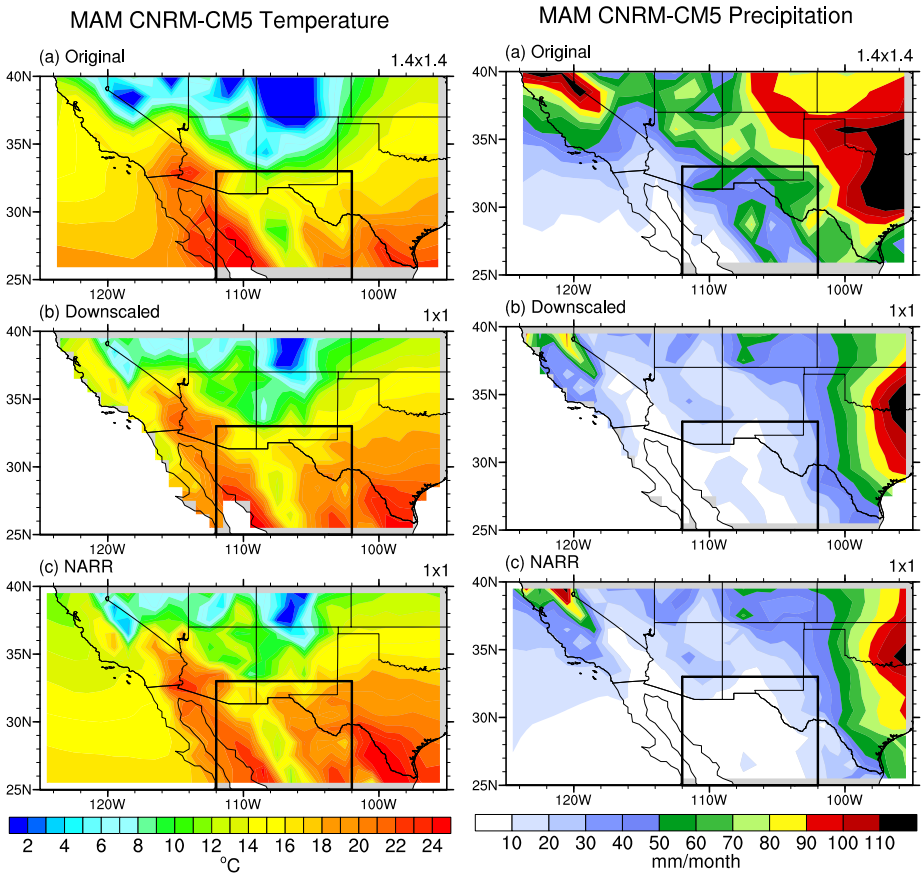


Figure A2-5. MAM temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled CNRM-CM5 output, (b) original CNRM-CM5 output, and (c) NARR observations from 1979 to 1999.

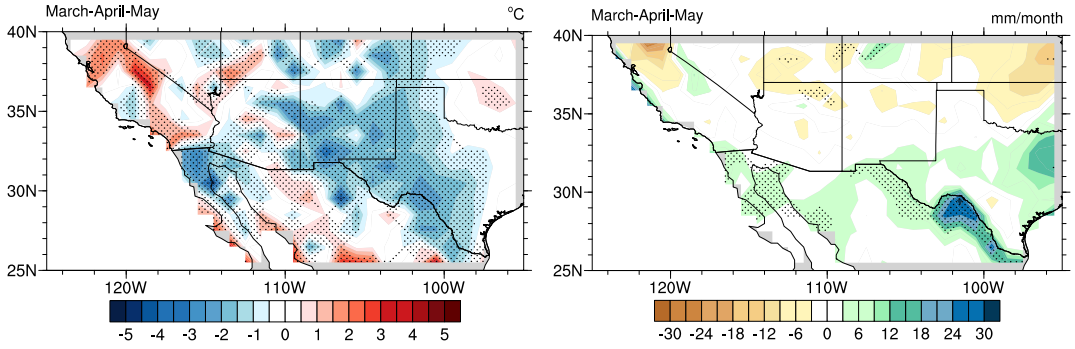


Figure A2-6. MAM temperature (left) and precipitation (right) spatial differences between downscaled CNRM-CM5 output and NARR with 90% statistical significance (stippled).

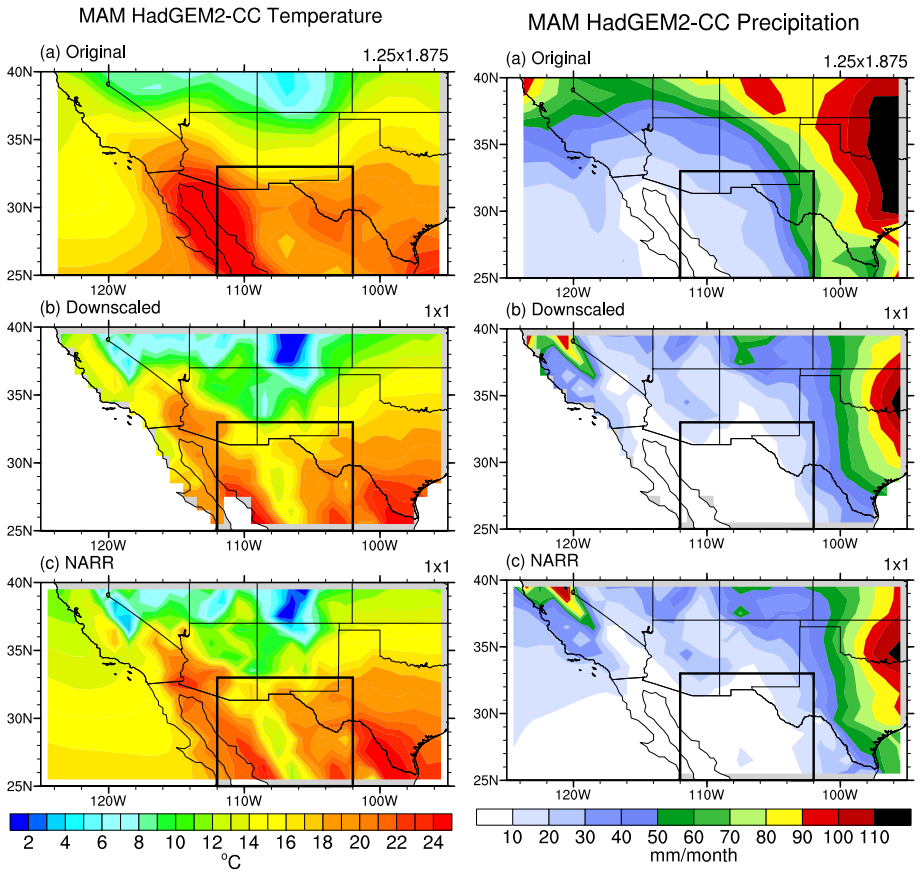


Figure A2-7. MAM temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled HadGEM2-CC output, (b) original HadGEM2-CC output, and (c) NARR observations from 1979 to 1999.

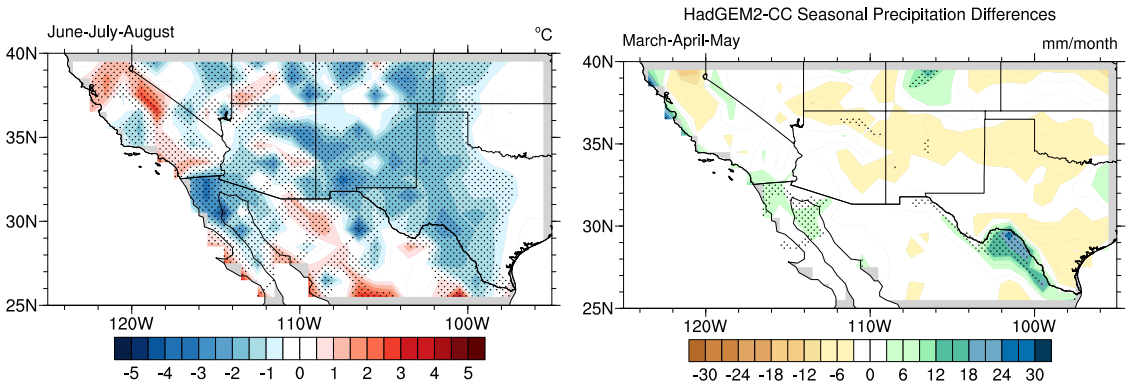


Figure A2-8. MAM temperature (left) and precipitation (right) spatial differences between downscaled HadGEM2-CC output and NARR with 90% statistical significance (stippled).

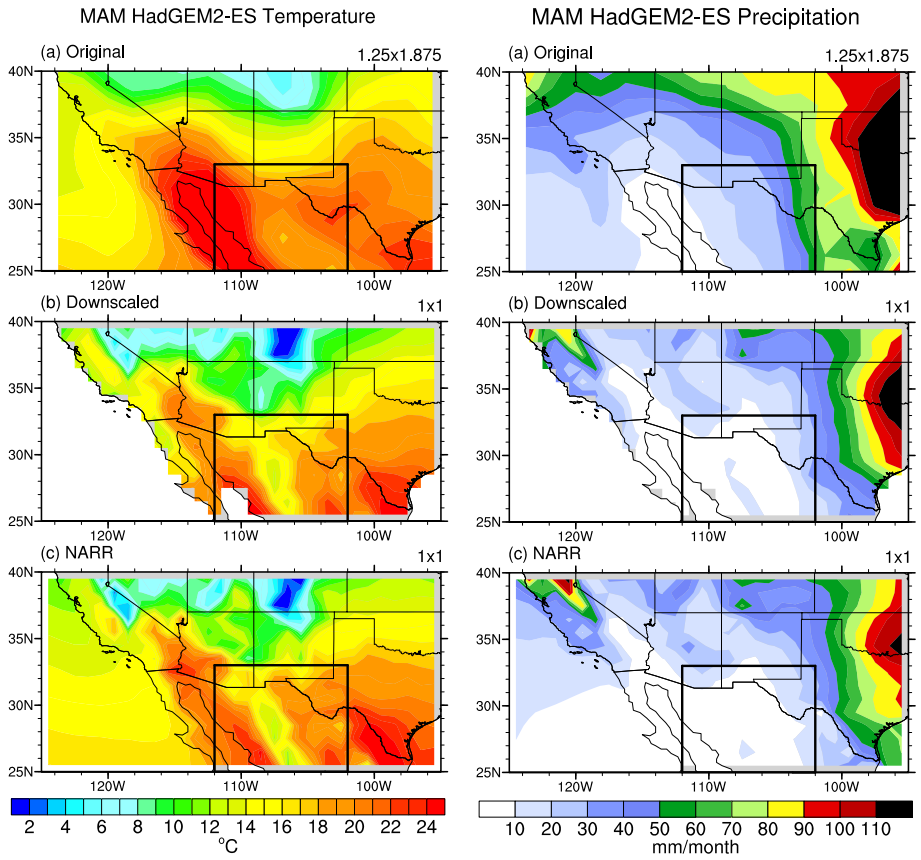


Figure A2-9. MAM temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled HadGEM2-ES output, (b) original HadGEM2-ES output, and (c) NARR observations from 1979 to 1999.

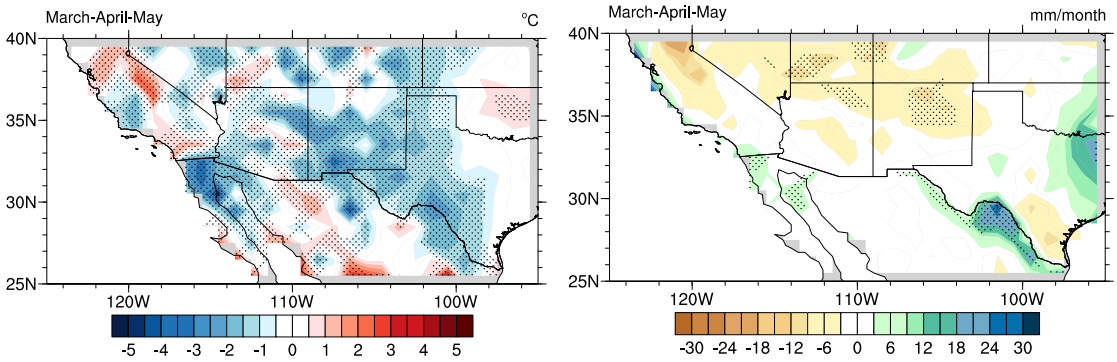


Figure A2-10. MAM temperature (left) and precipitation (right) spatial differences between downscaled HadGEM2-ES output and NARR with 90% statistical significance (stippled).

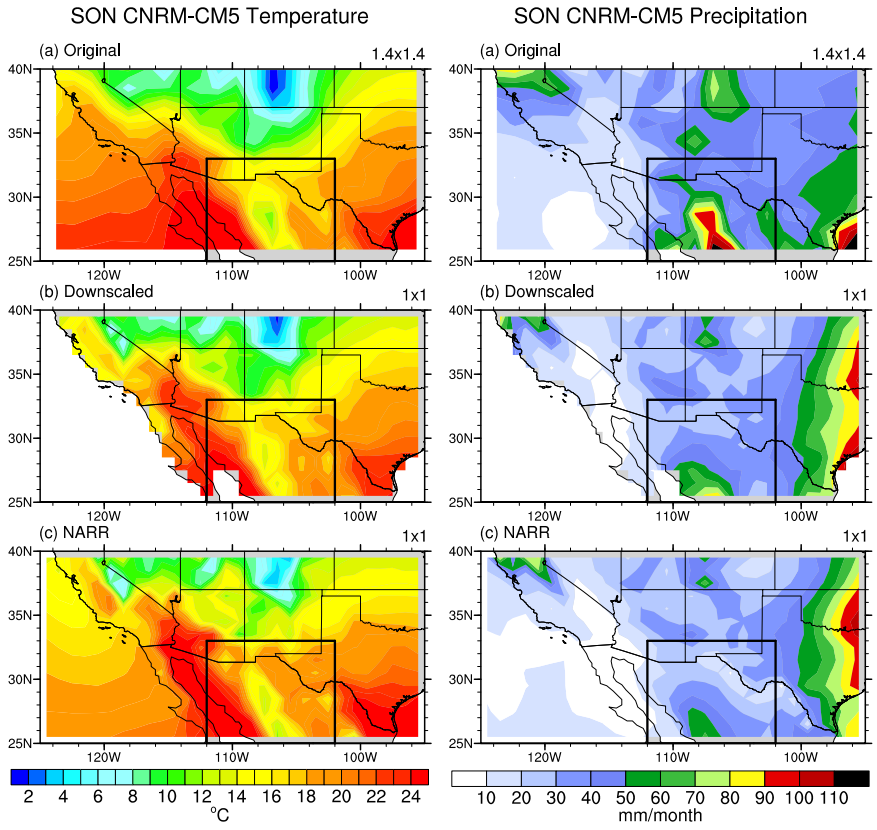


Figure A2-11. SON temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled CNRM-CM5 output, (b) original CNRM-CM5 output, and (c) NARR observations from 1979 to 1999.

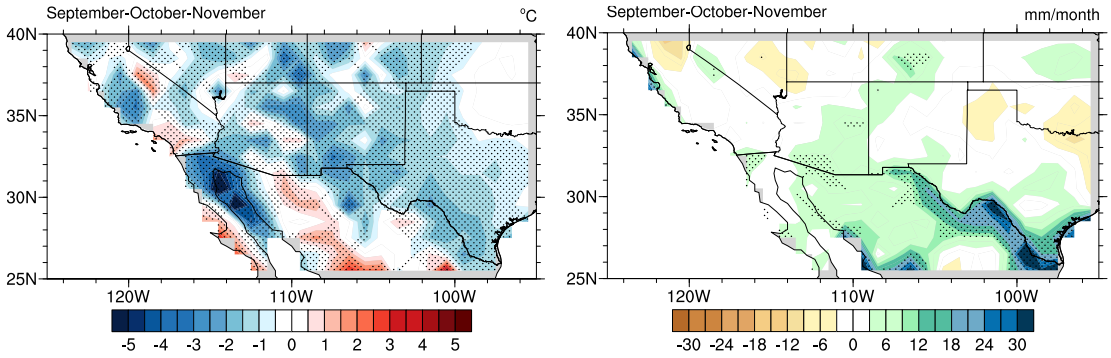


Figure A2-12. SON temperature (left) and precipitation (right) spatial differences between downscaled CNRM-CM5 output and NARR with 90% statistical significance (stippled).

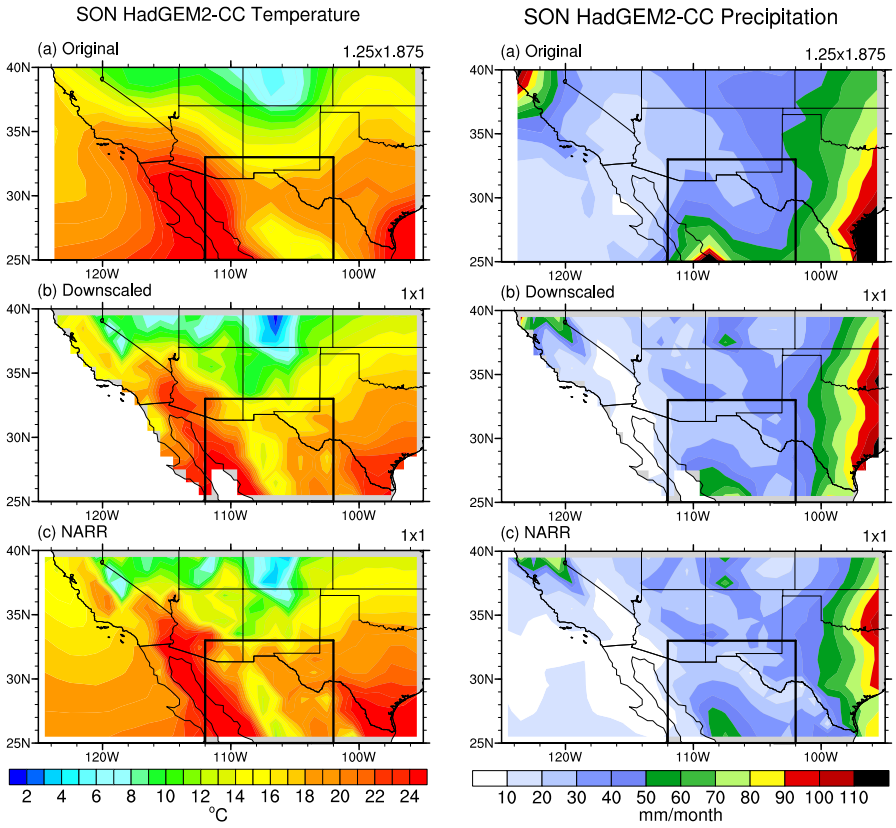


Figure A2-13. SON temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled HadGEM2-CC output, (b) original HadGEM2-CC output, and (c) NARR observations from 1979 to 1999.

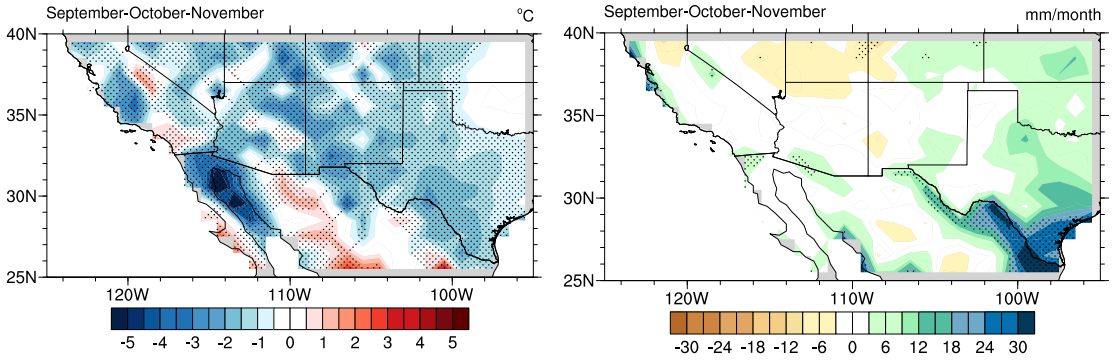


Figure A2-14. SON temperature (left) and precipitation (right) spatial differences between downscaled HadGEM2-CC output and NARR with 90% statistical significance (stippled).

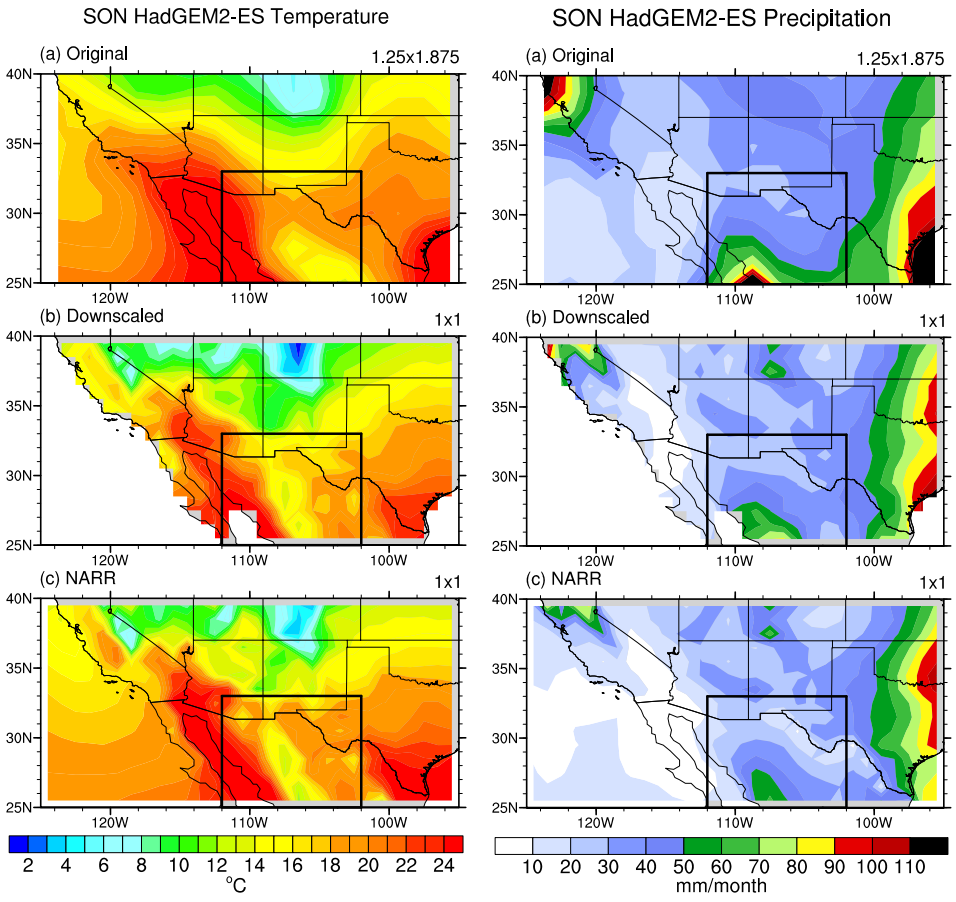


Figure A2-15. SON temperature (left) and precipitation (right) spatial patterns for (a) statistically downscaled HadGEM2-ES output, (b) original HadGEM2-ES output, and (c) NARR observations from 1979 to 1999.

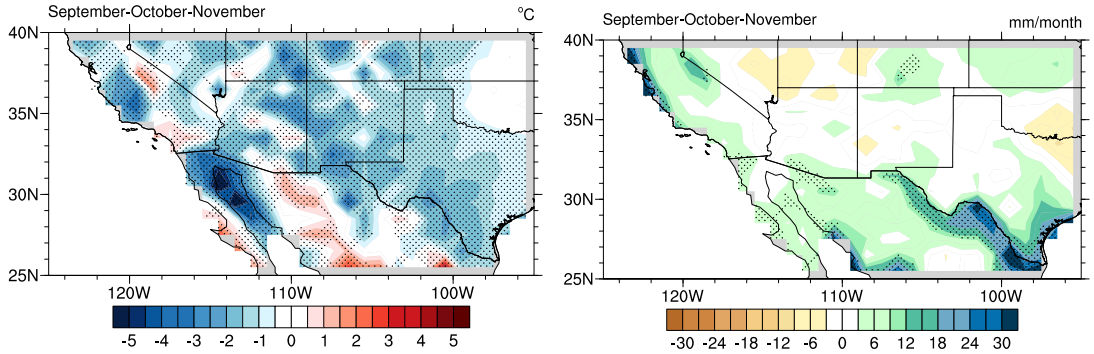


Figure A2-16. SON temperature (left) and precipitation (right) spatial differences between downscaled HadGEM2-ES output and NARR with 90% statistical significance (stippled).

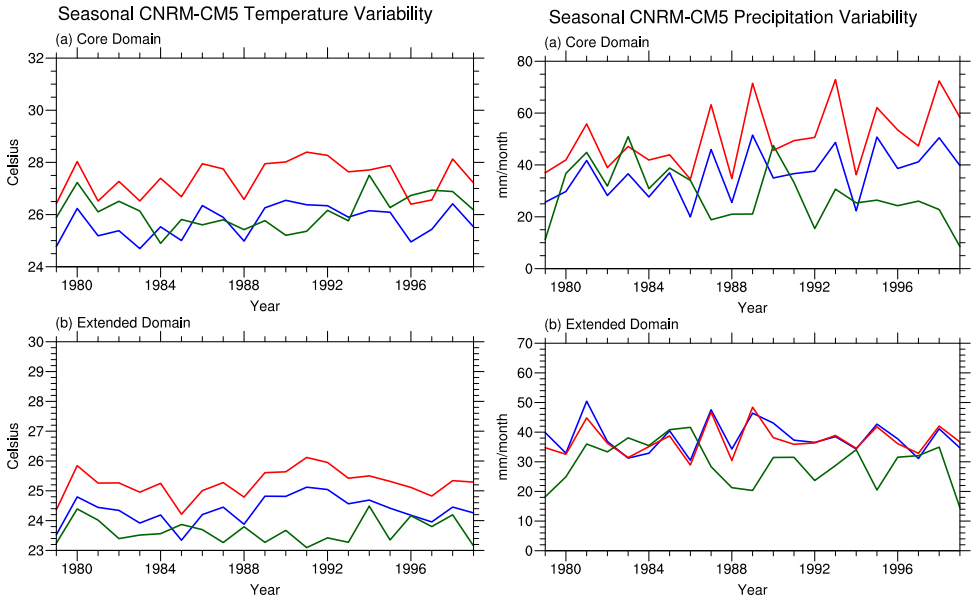


Figure A2-17. JJA temperature (left) and precipitation (right) interannual variability of downscaled CNRM-CM5 output (blue), CNRM-CM5 original output (red), and NARR (green) for the core (a) and extended domain (b).

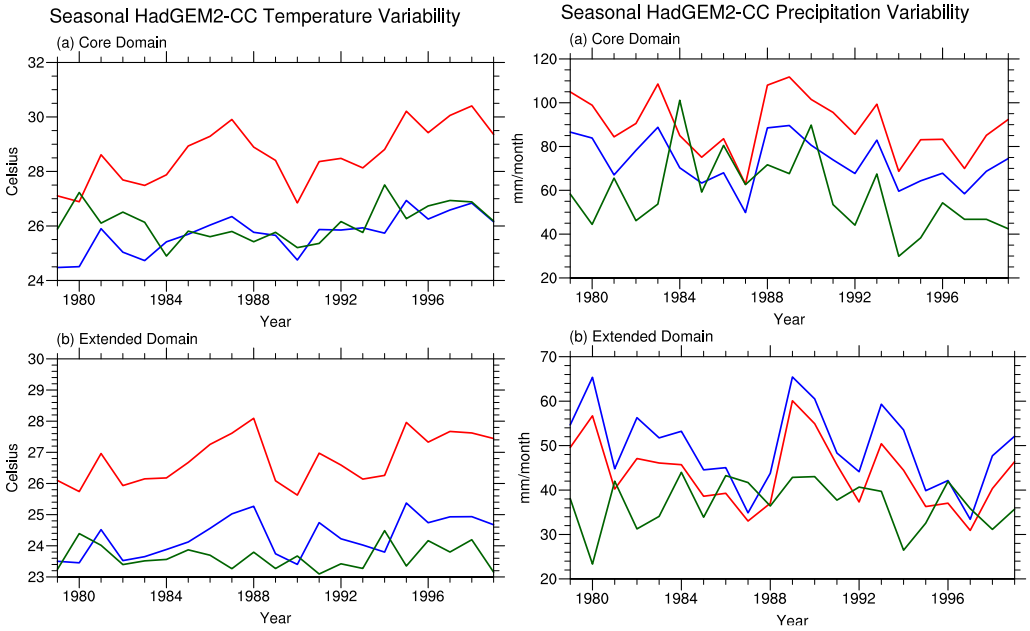


Figure A2-18. JJA temperature (left) and precipitation (right) interannual variability of downscaled HadGEM2-CC output (blue), HadGEM2-CC original output (red), and NARR (green) for the core (a) and extended domain (b).

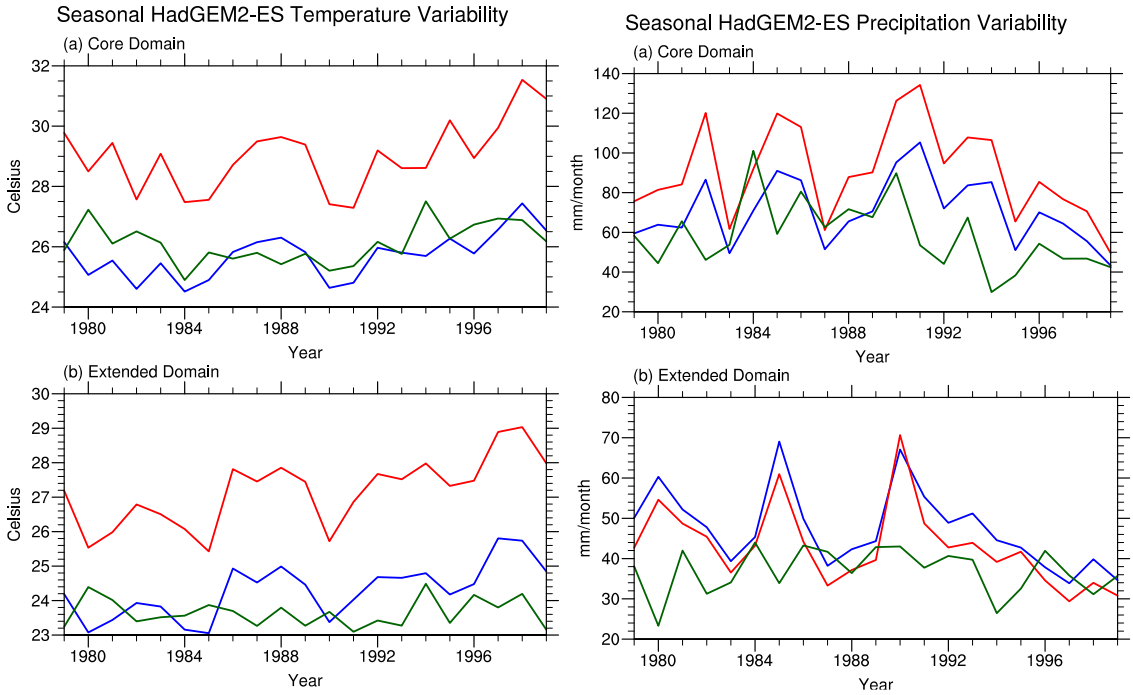


Figure A2-19. JJA temperature (left) and precipitation (right) interannual variability of downscaled HadGEM2-ES output (blue), HadGEM2-ES original output (red), and NARR (green) for the core (a) and extended domain (b).

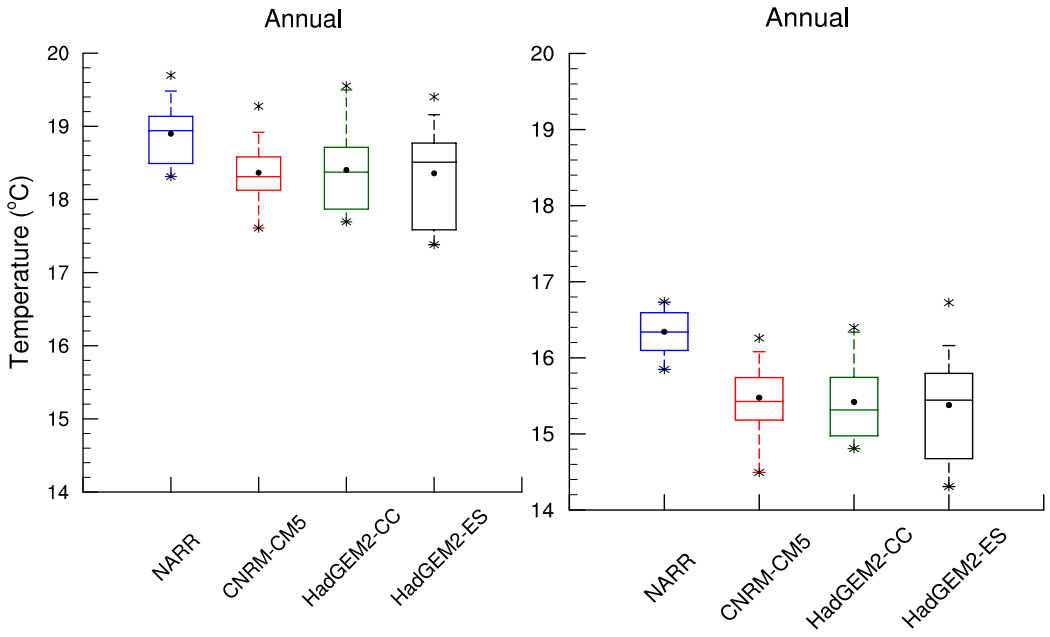


Figure A2-20. Annual temperature box-and-whisker plots for the core (left) and extended (right) domain. The 25th and 75th percentiles are shown by the edges of the boxes, the whiskers represent the 5th and 95th percentiles, and the medians are plotted as horizontal lines within the boxes. The asterisks illustrate the maximum and minimum temperature, average represents a black dot, while the yellow dots represent significant differences at the 90% level.

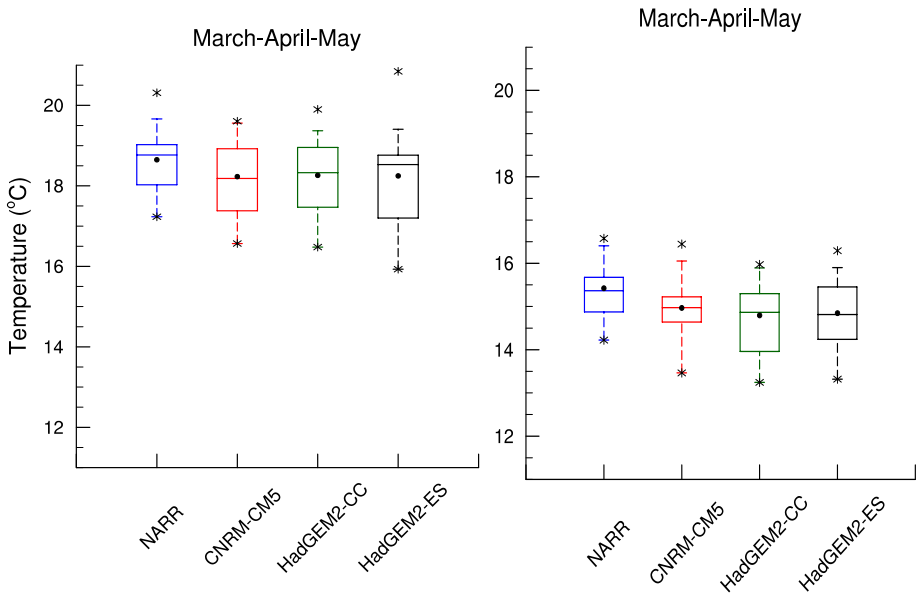


Figure A2-21. Similar to Figure 22, except for MAM.

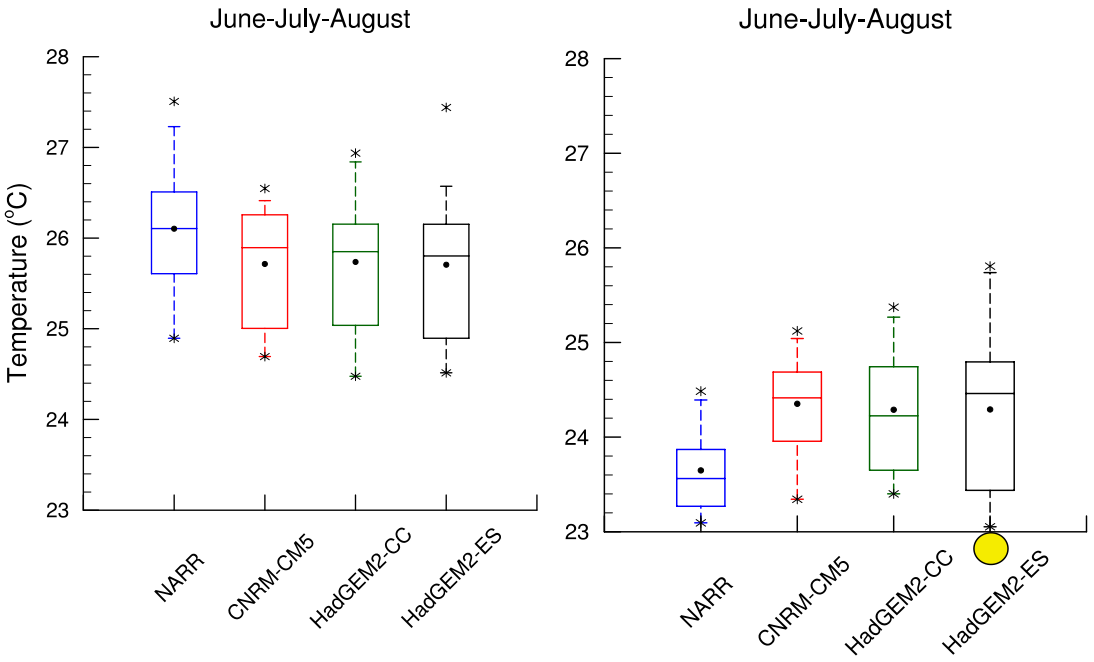


Figure A2-22. Similar to Figure 22, except for JJA.

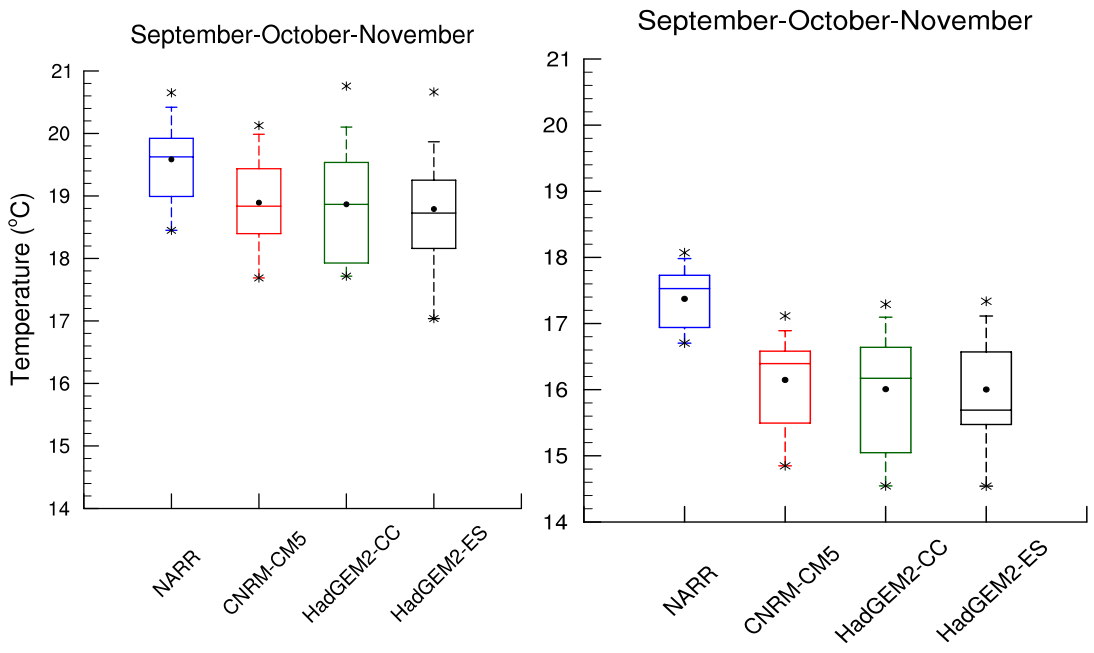


Figure A2-23. Similar to Figure 22, except for SON.

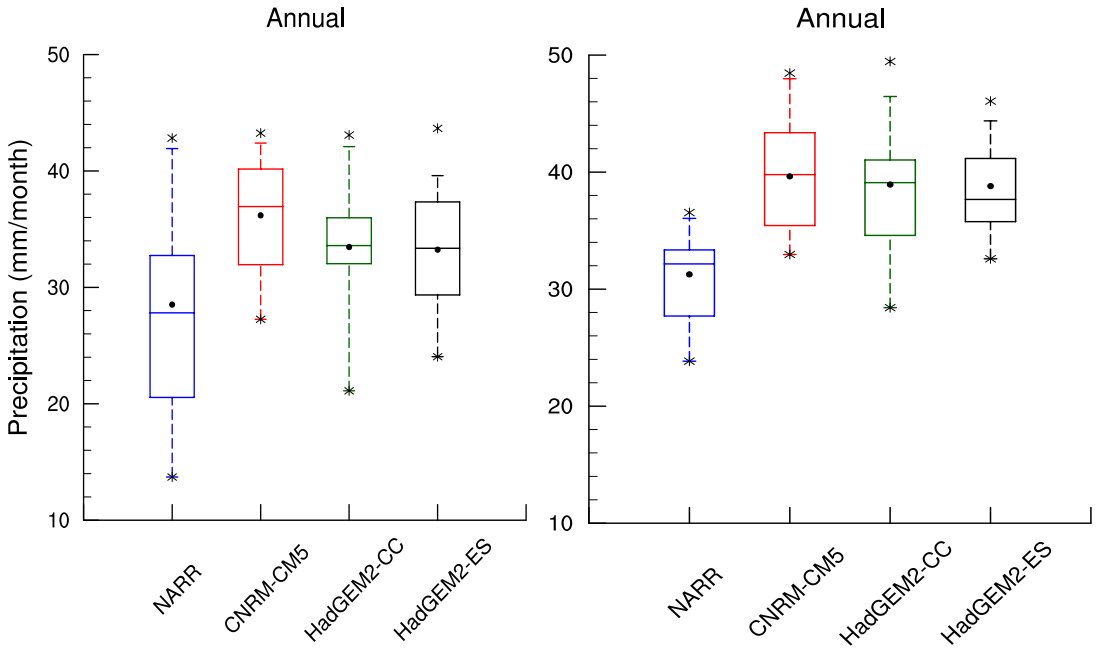


Figure A2-24. Annual precipitation box-and-whisker plots for the core (left) and extended (right) domain. The 25th and 75th percentiles are shown by the edges of the boxes, the whiskers represent the 5th and 95th percentiles, and the medians are plotted as horizontal lines within the boxes. The asterisks illustrate the maximum and minimum precipitation, average represents a black dot, while the yellow dots represent significant differences at the 90% level.

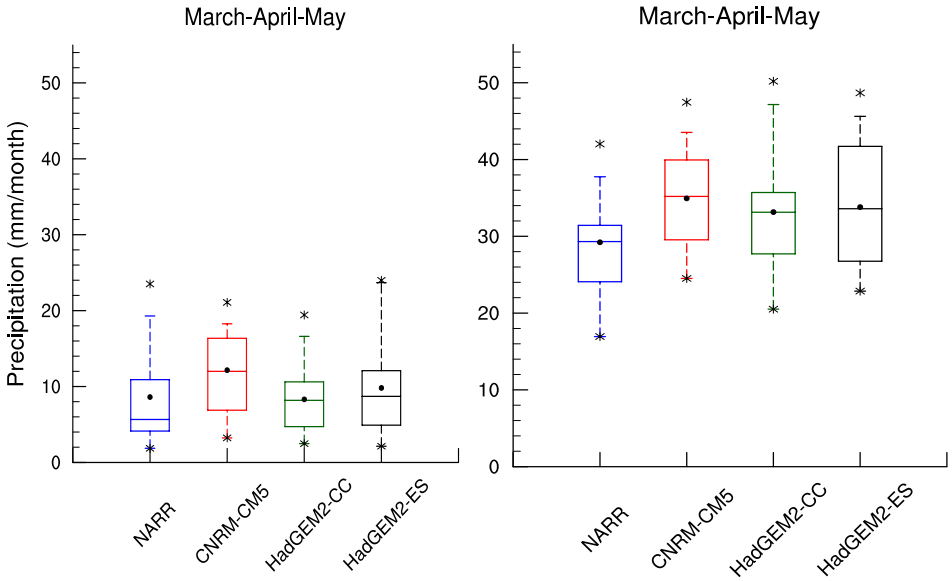


Figure A2-25. Similar to Figure 26, except for MAM.

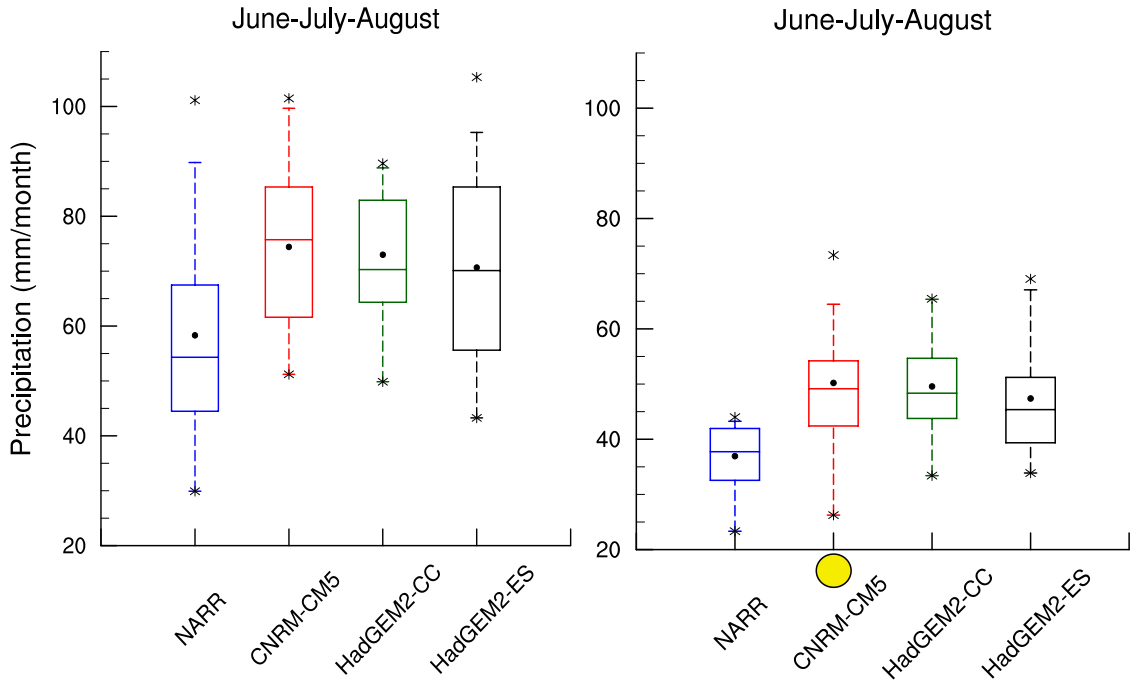


Figure A2-26. Similar to Figure 26, except for JJA.

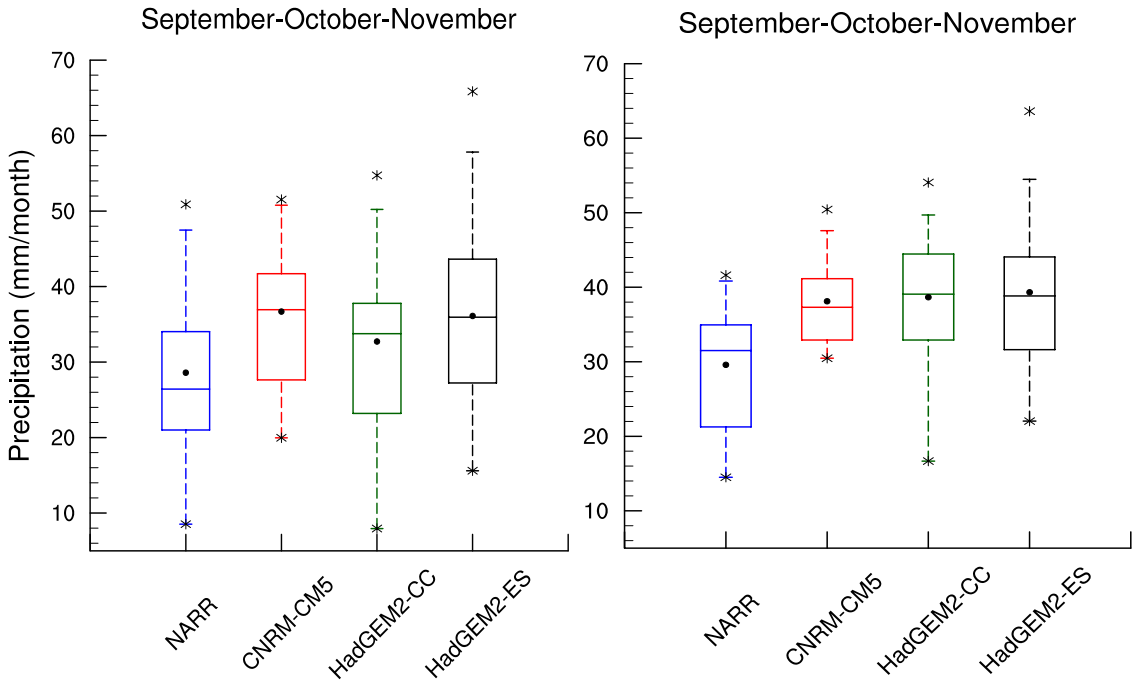


Figure A2-27. Similar to Figure 26, except for SON.

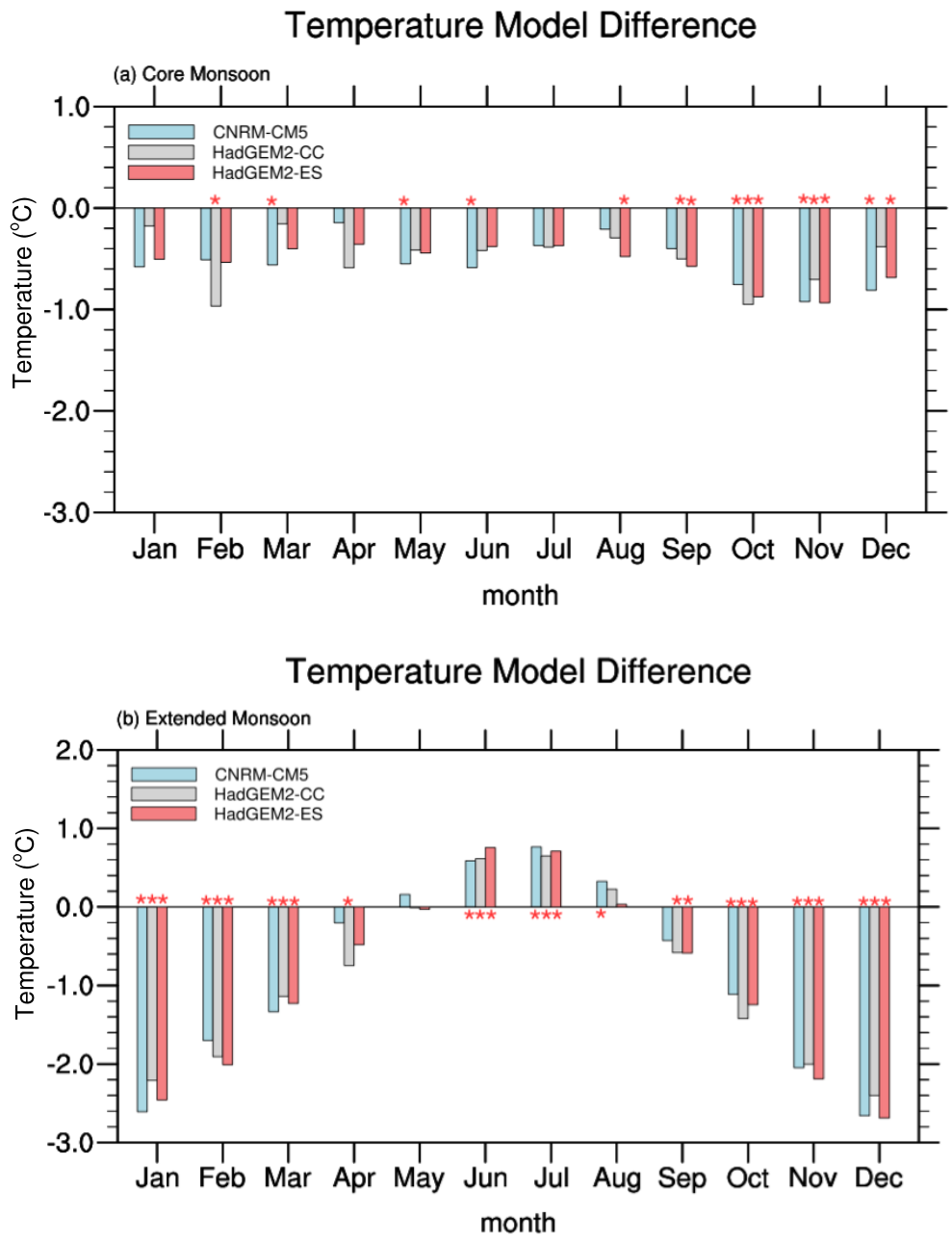


Figure A2-28. Monthly temperature climatology differences between downscaled GCMs' and NARR output for the (a) core and (b) extended monsoon domain from 1979 to 1999. Stars denote 90% significance.

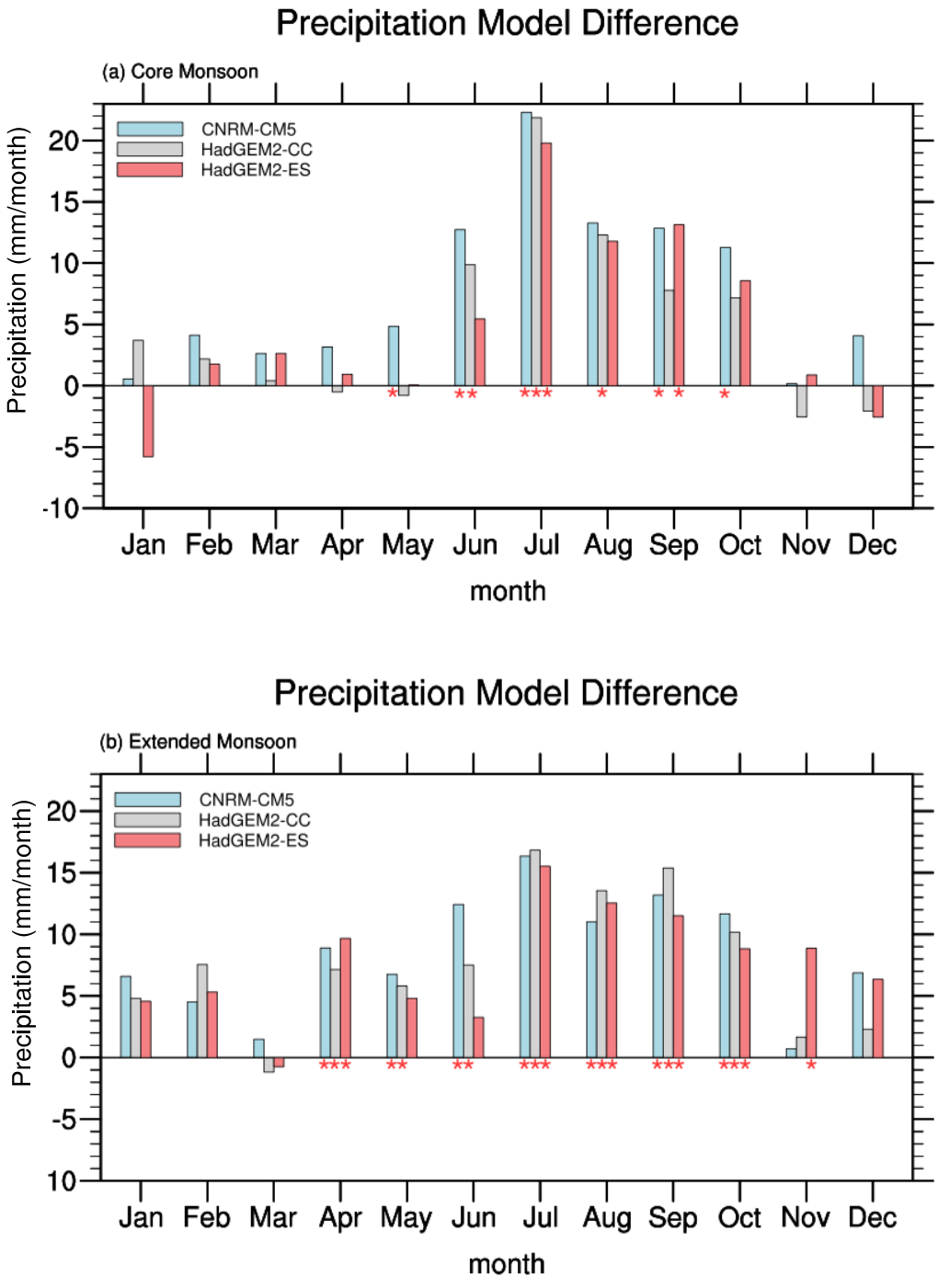


Figure A2-29. Monthly precipitation climatology differences between downscaled GCMs' and NARR output for the (a) core and (b) extended monsoon domain from 1979 to 1999. Stars denote 90% significance.

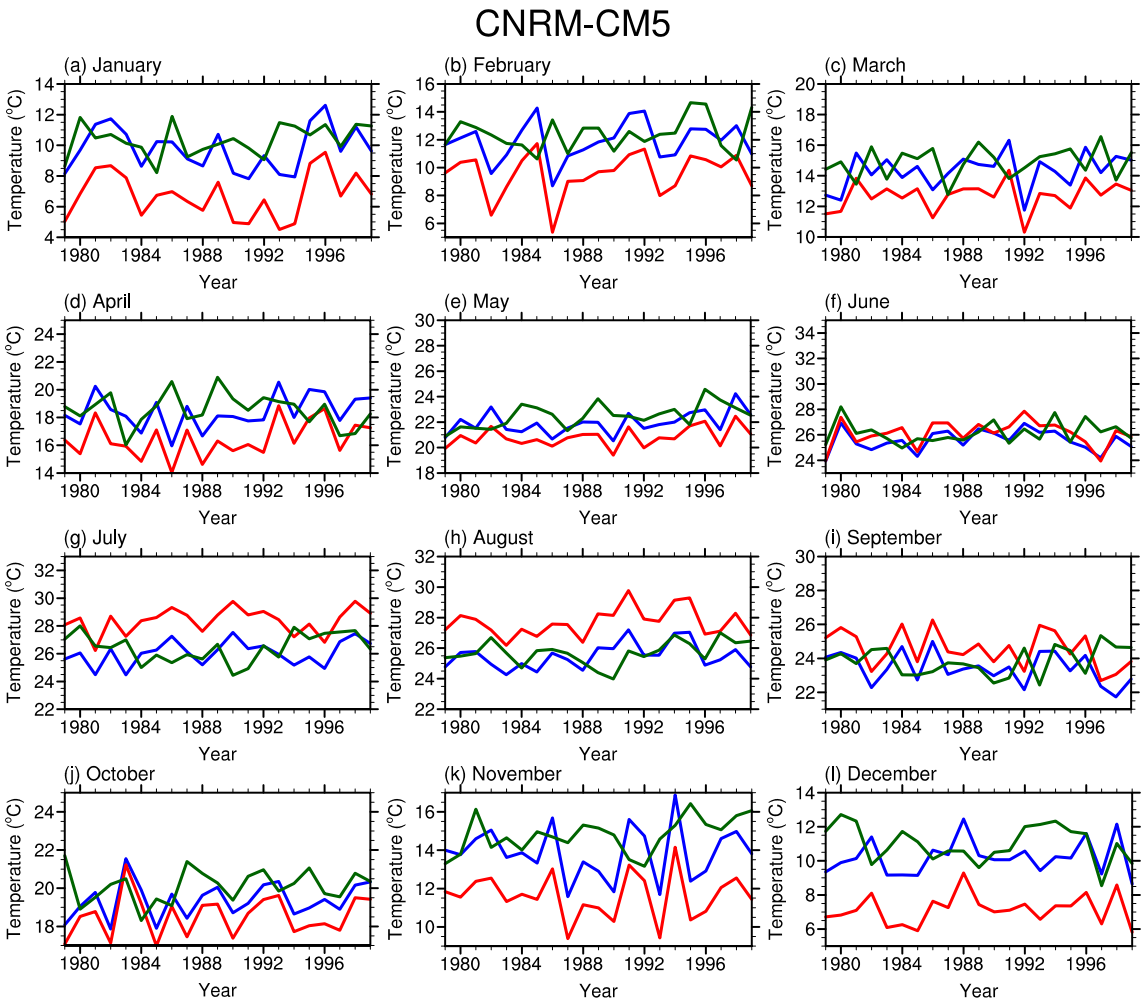


Figure A2-30. Interannual temperature variability at each month, from 1979 to 1999, for the core monsoon domain in the CNRM-CM5 model. NARR (green), downscaled CNRM-CM5 output (blue), and original CNRM-CM5 output (red) are shown.

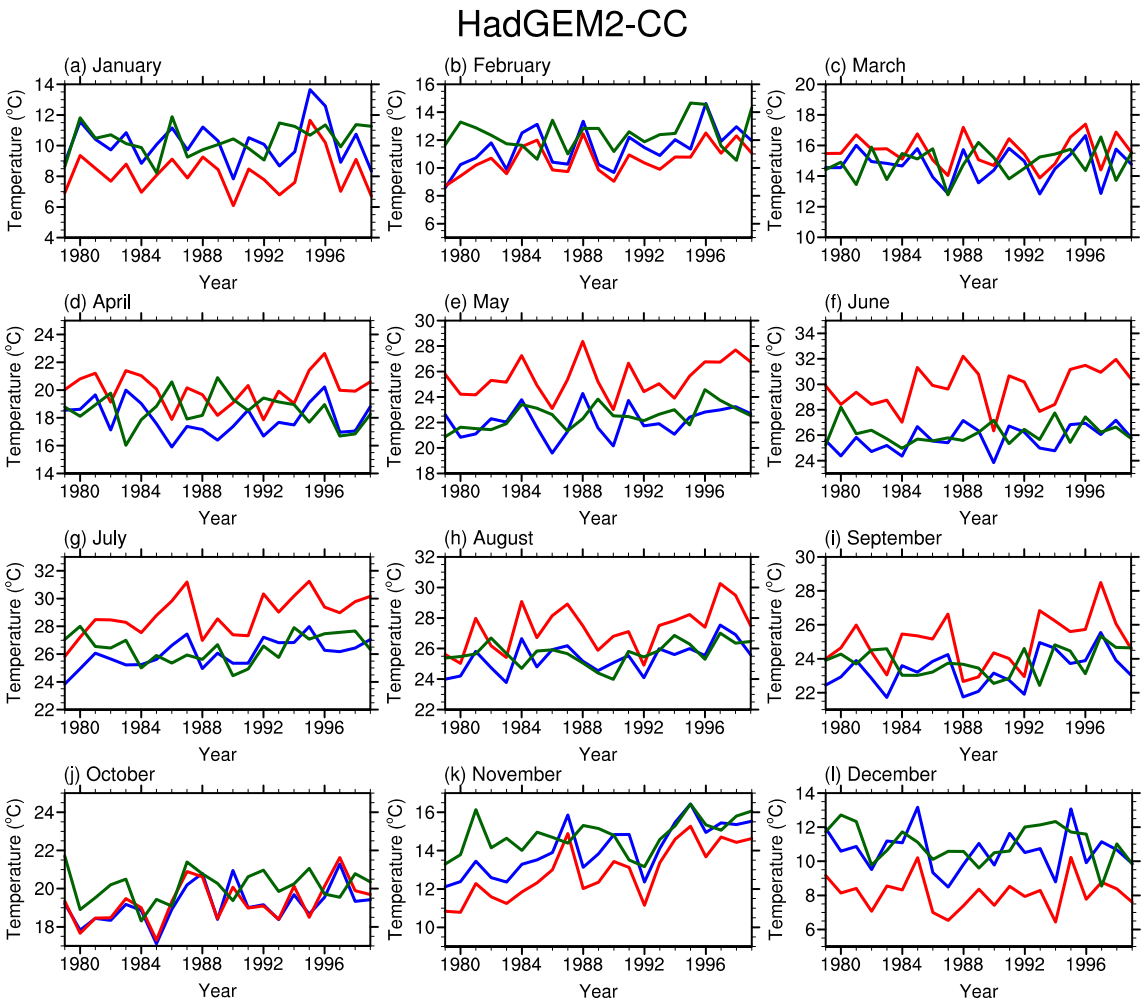


Figure A2-31. Interannual temperature variability at each month, from 1979 to 1999, for the core monsoon domain in the HadGEM2-CC model. NARR (green), downscaled HadGEM2-CC output (blue), and original HadGEM2-CC output (red) are shown.

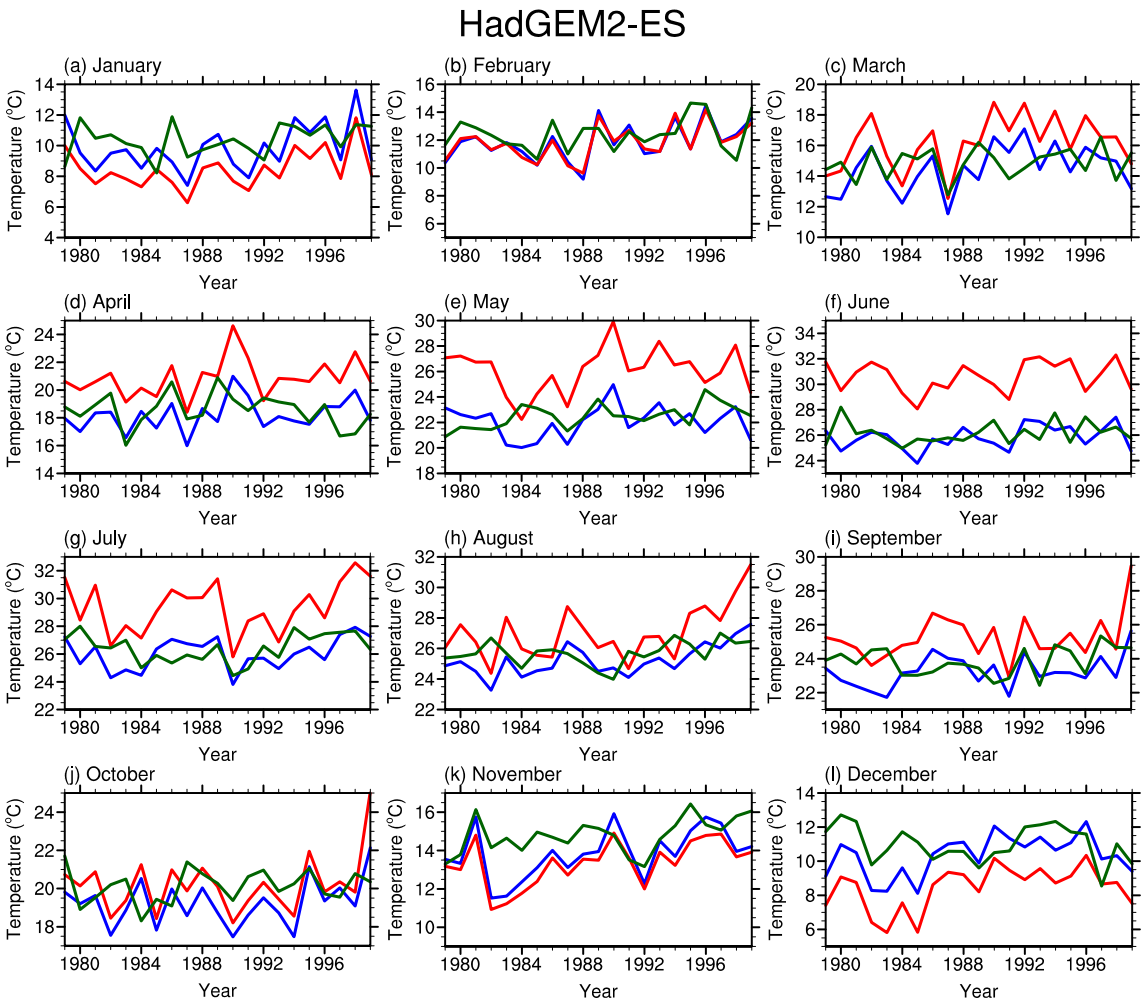


Figure A2-32. Interannual temperature variability at each month, from 1979 to 1999, for the core monsoon domain in the HadGEM2-ES model. NARR (green), downscaled HadGEM2-ES output (blue), and original HadGEM2-ES output (red) are shown.

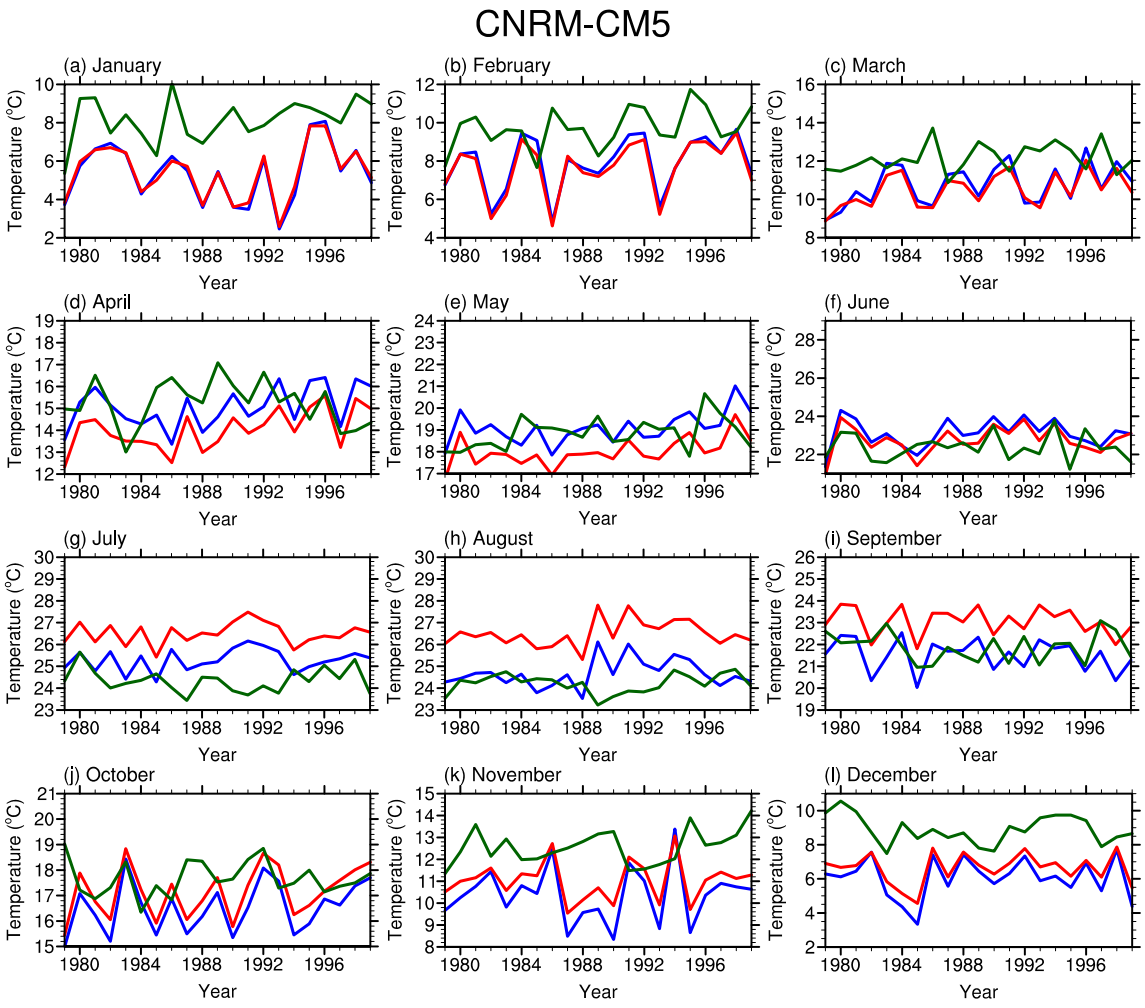


Figure A2-33. Similar to Figure 30, except for the extended domain.

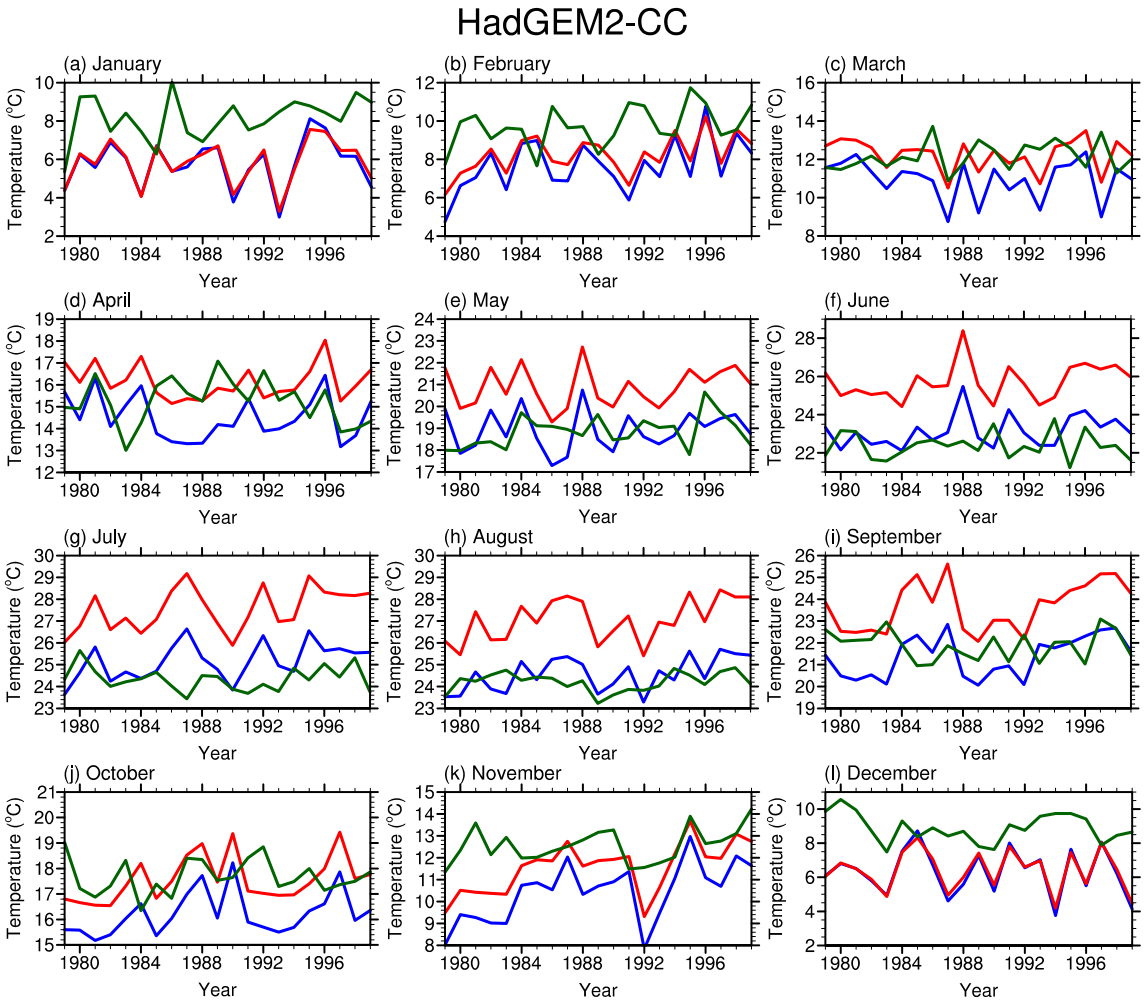


Figure A2-34. Similar to Figure 31, except for the extended domain.

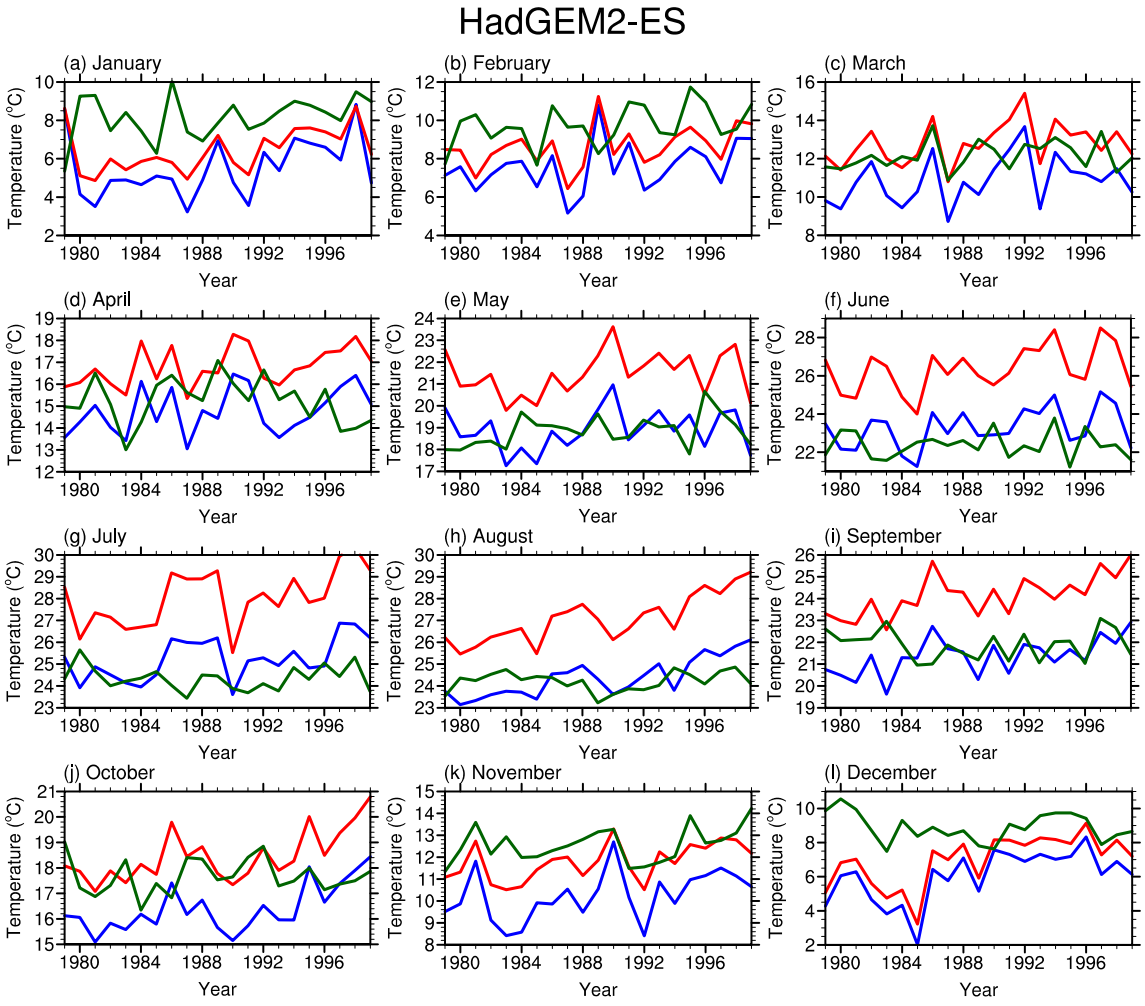


Figure A2-35. Similar to Figure 32, except for the extended domain.

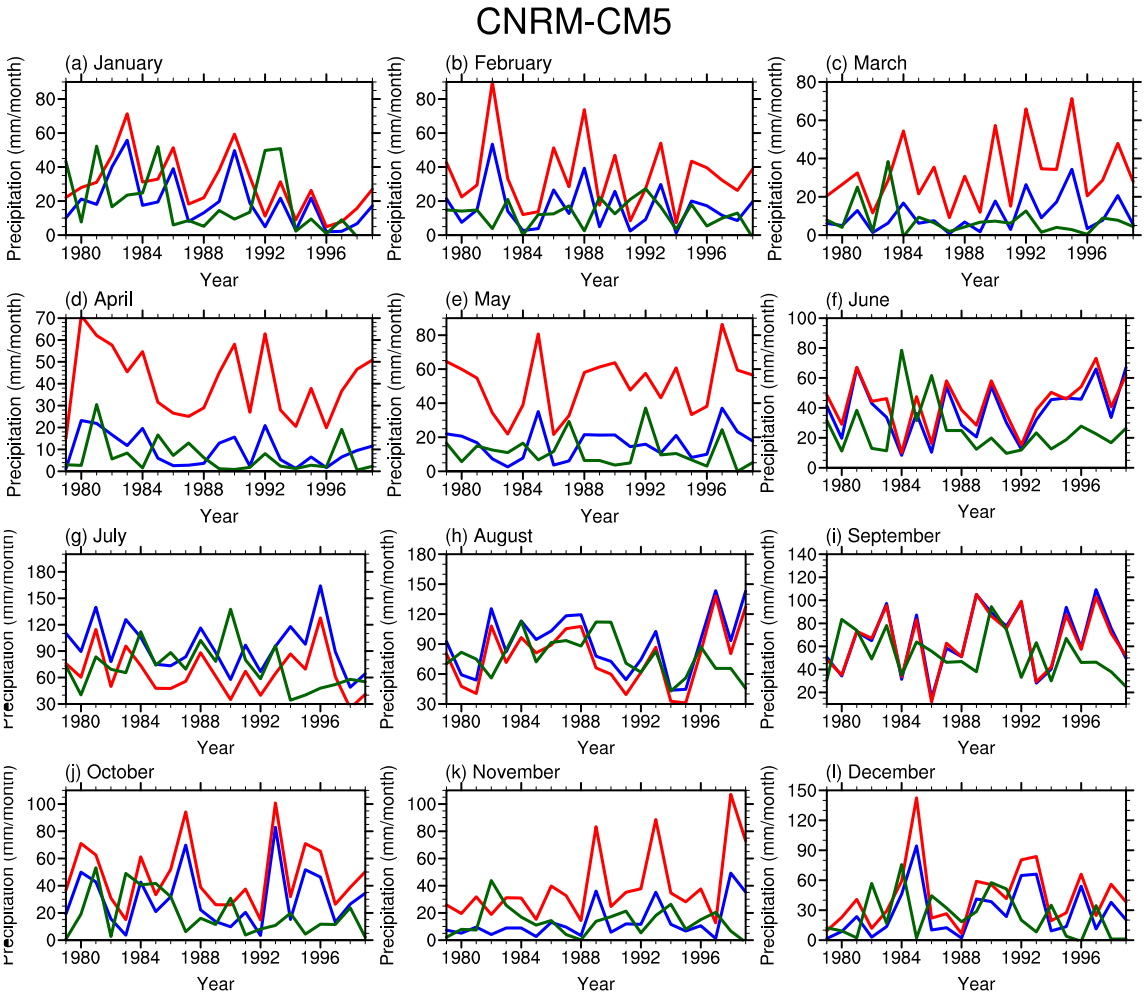


Figure A2-36. Interannual precipitation variability at each month, from 1979 to 1999, for the core monsoon domain in the CNRM-CM5 model. NARR (green), downscaled CNRM-CM5 output (blue), and original CNRM-CM5 output (red) are shown.

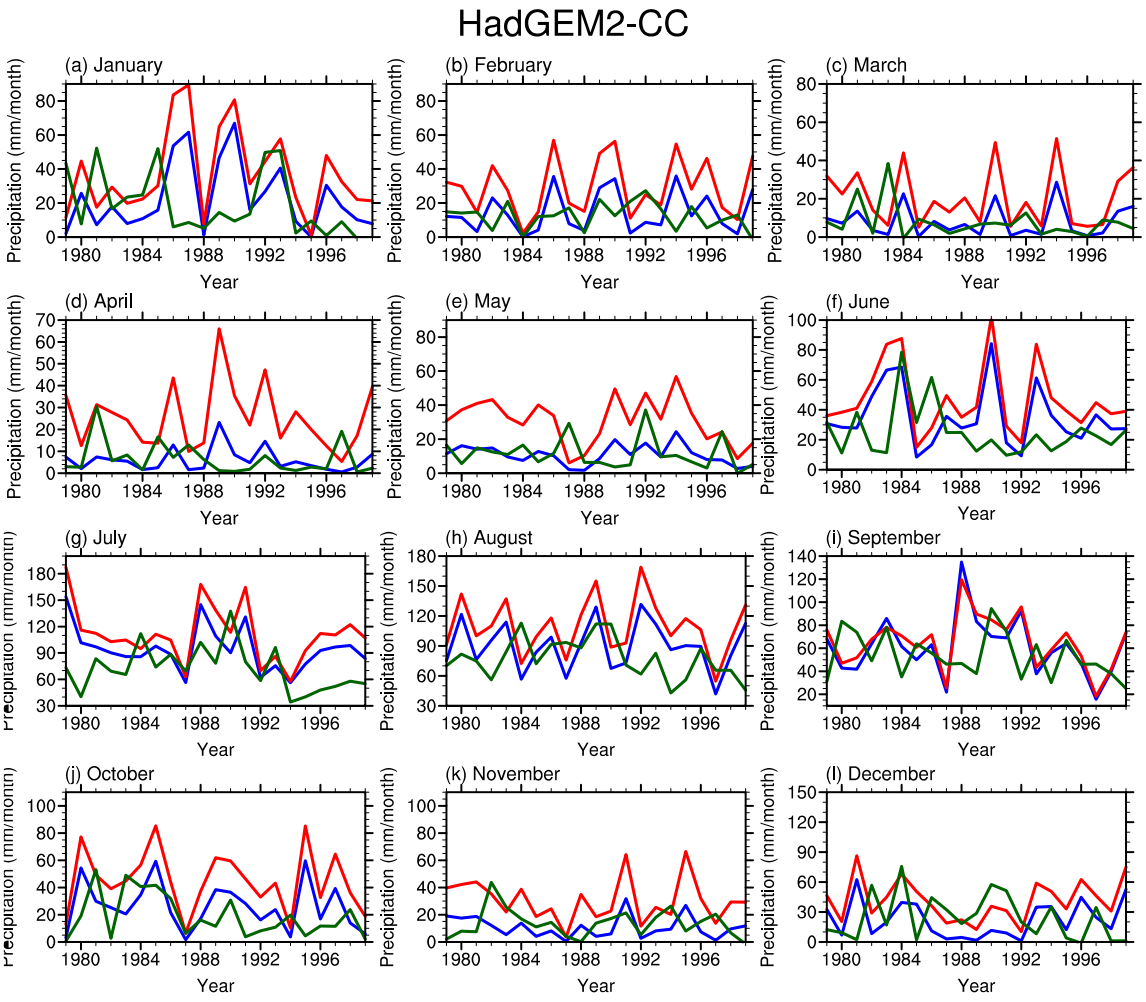


Figure A2-37. Interannual precipitation variability at each month, from 1979 to 1999, for the core monsoon domain in the HadGEM2-CC model. NARR (green), downscaled HadGEM2-CC output (blue), and original HadGEM2-CC output (red) are shown.

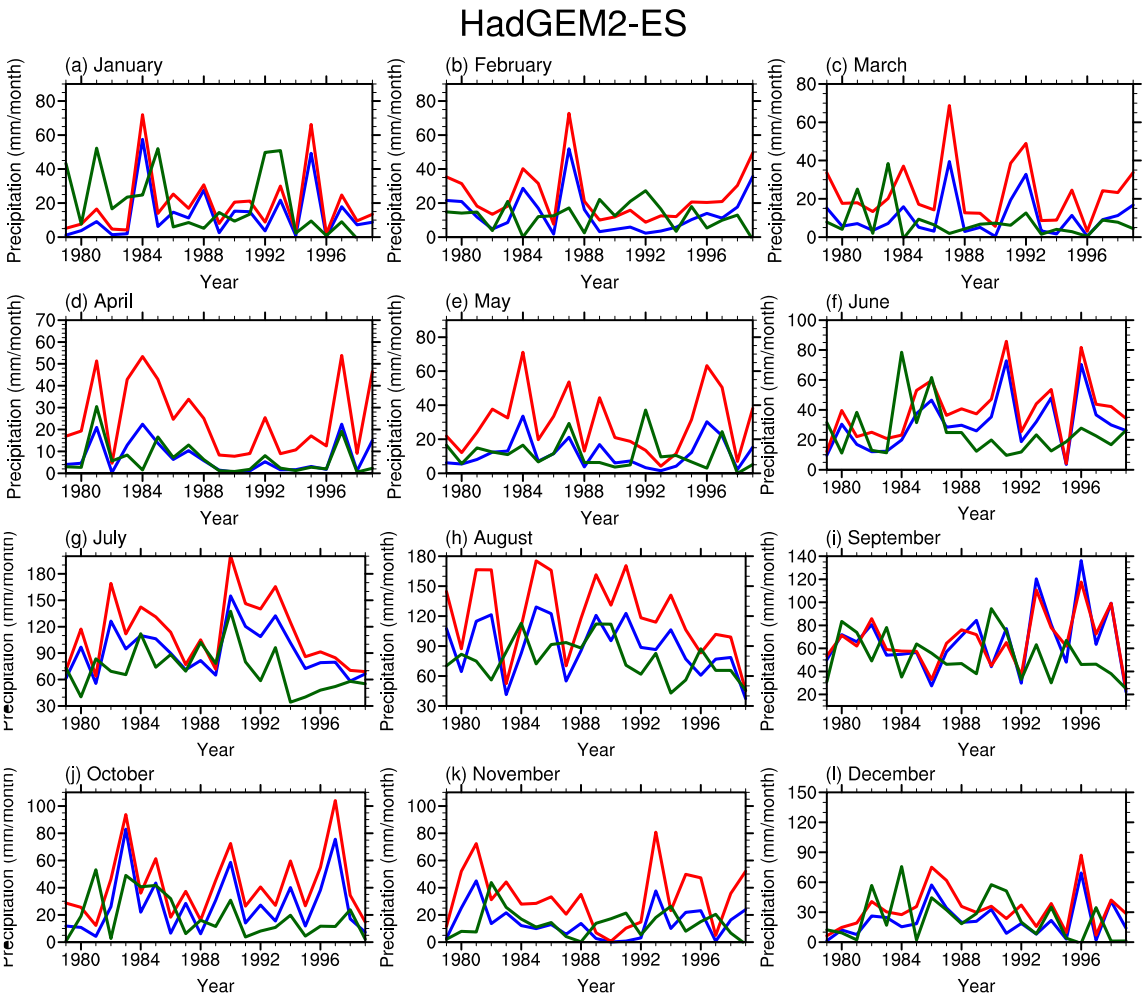


Figure A2-38. Interannual precipitation variability at each month, from 1979 to 1999, for the core monsoon domain in the HadGEM2-ES model. NARR (green), downscaled HadGEM2-ES output (blue), and original HadGEM2-ES output (red) are shown.

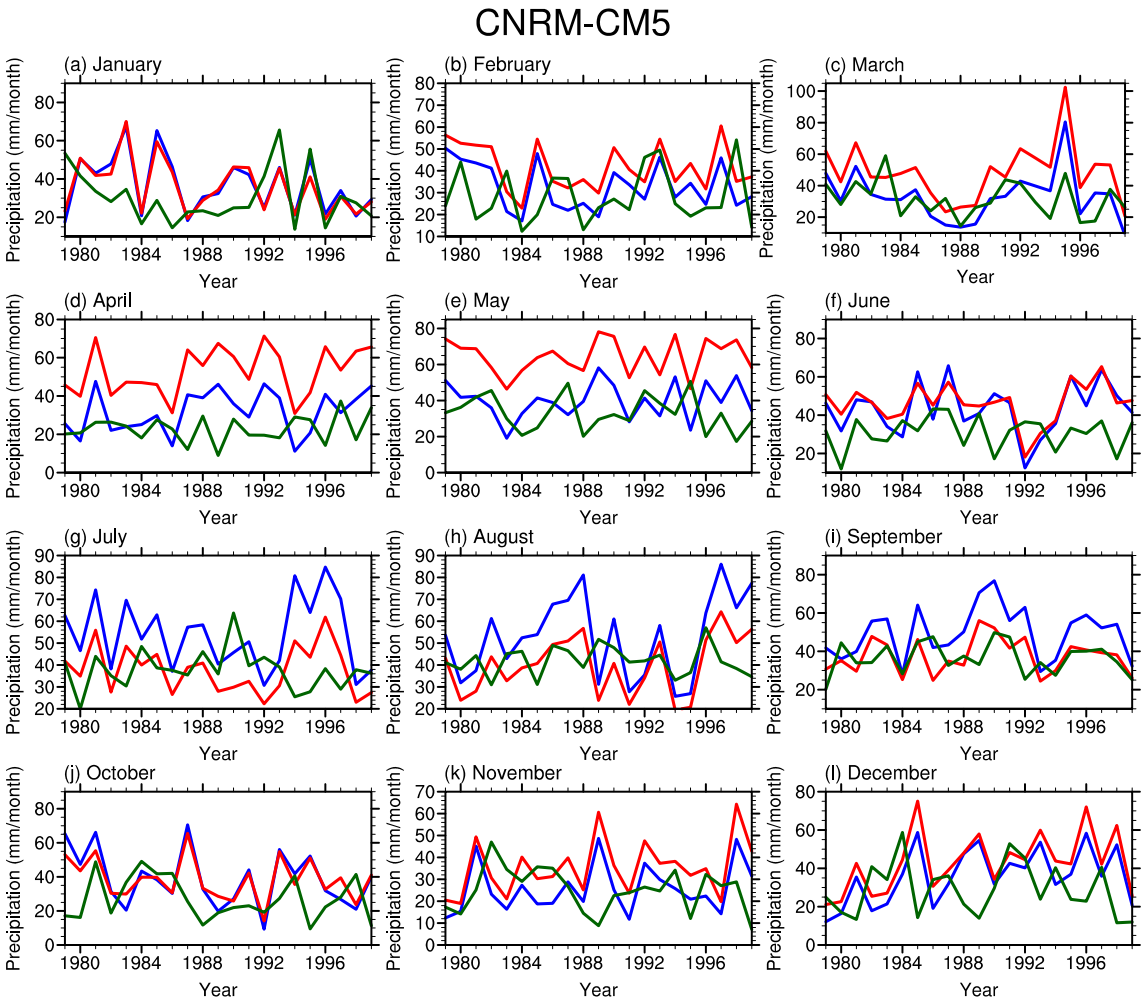


Figure A2-39. Similar to Figure 36, except for the extended domain.

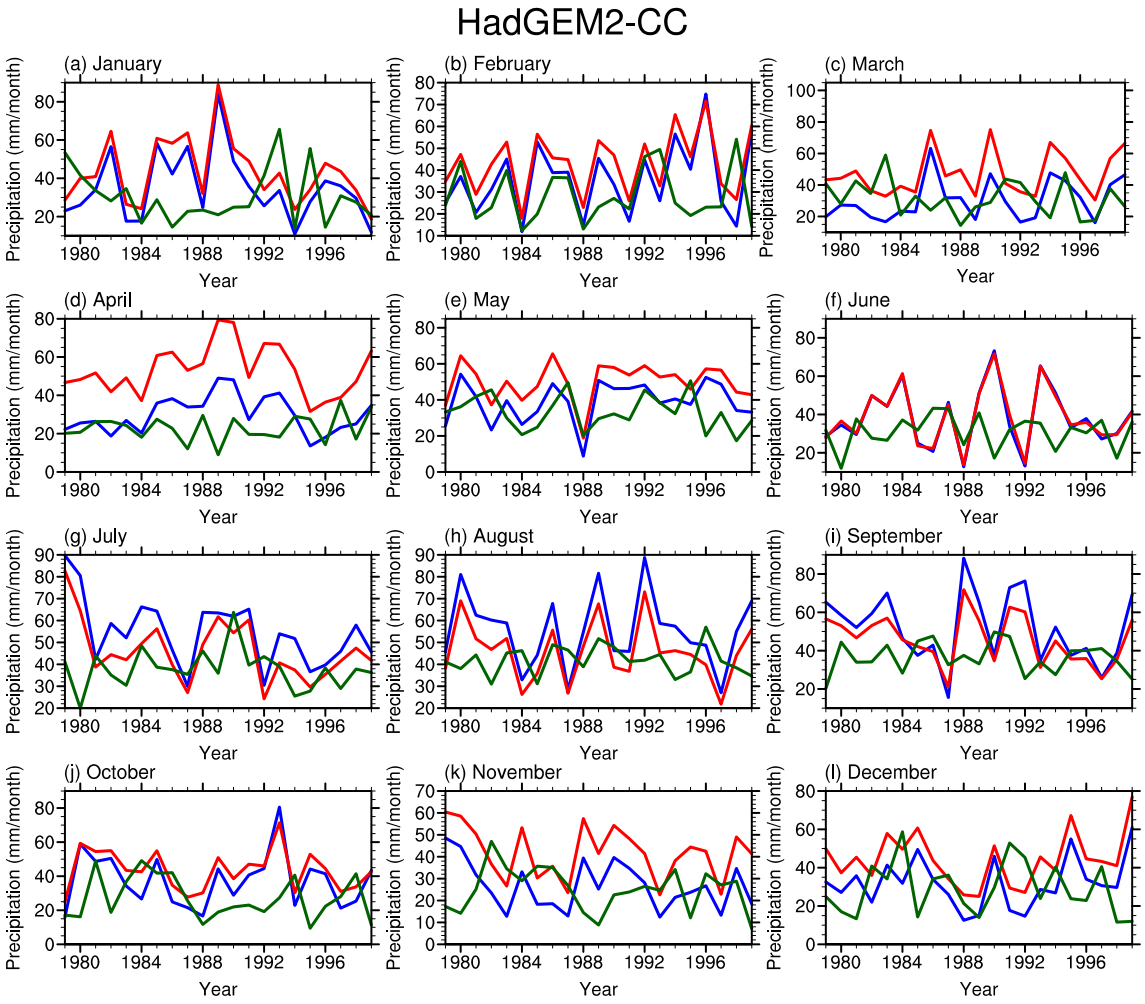


Figure A2-40. Similar to Figure 37, except for the extended domain.

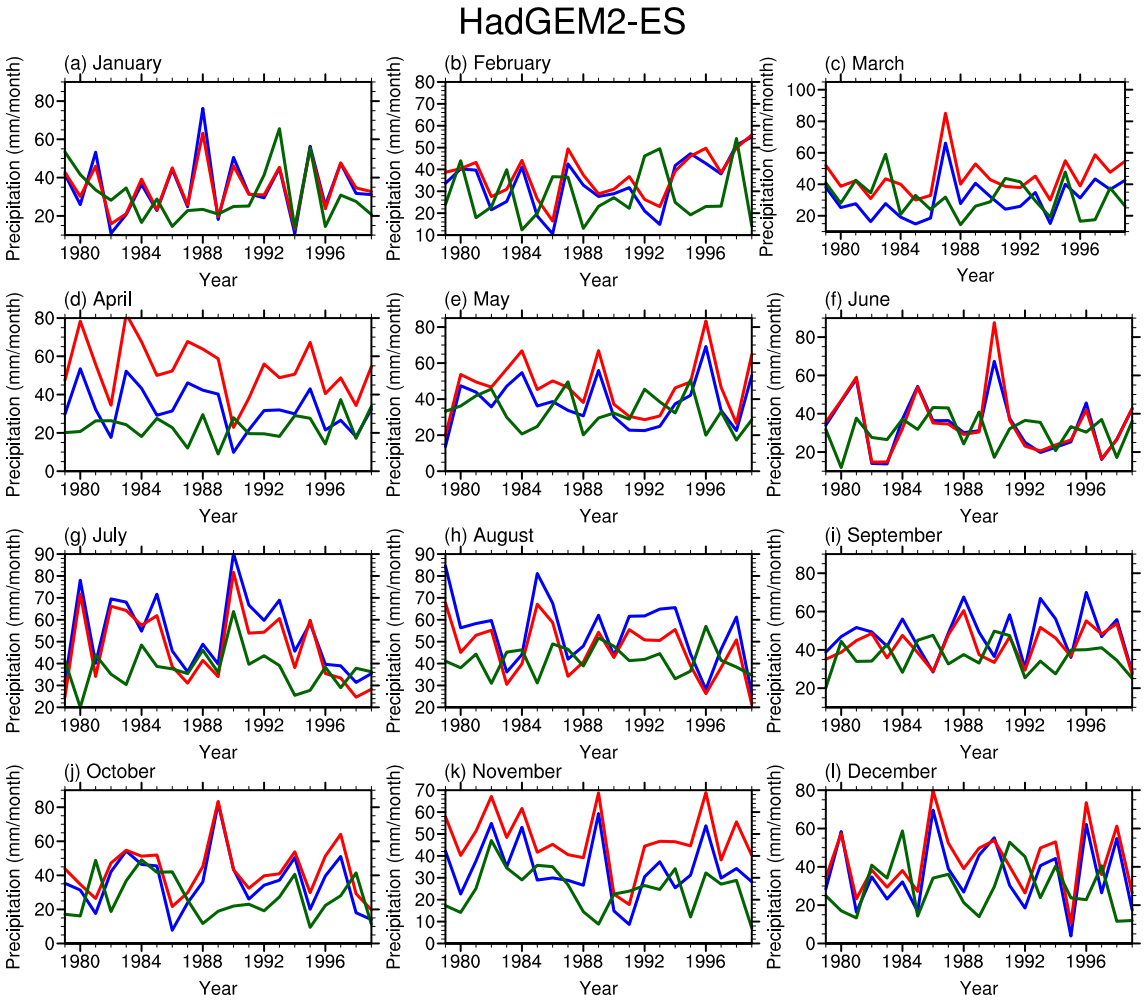


Figure A2-41. Similar to Figure 38, except for the extended domain.

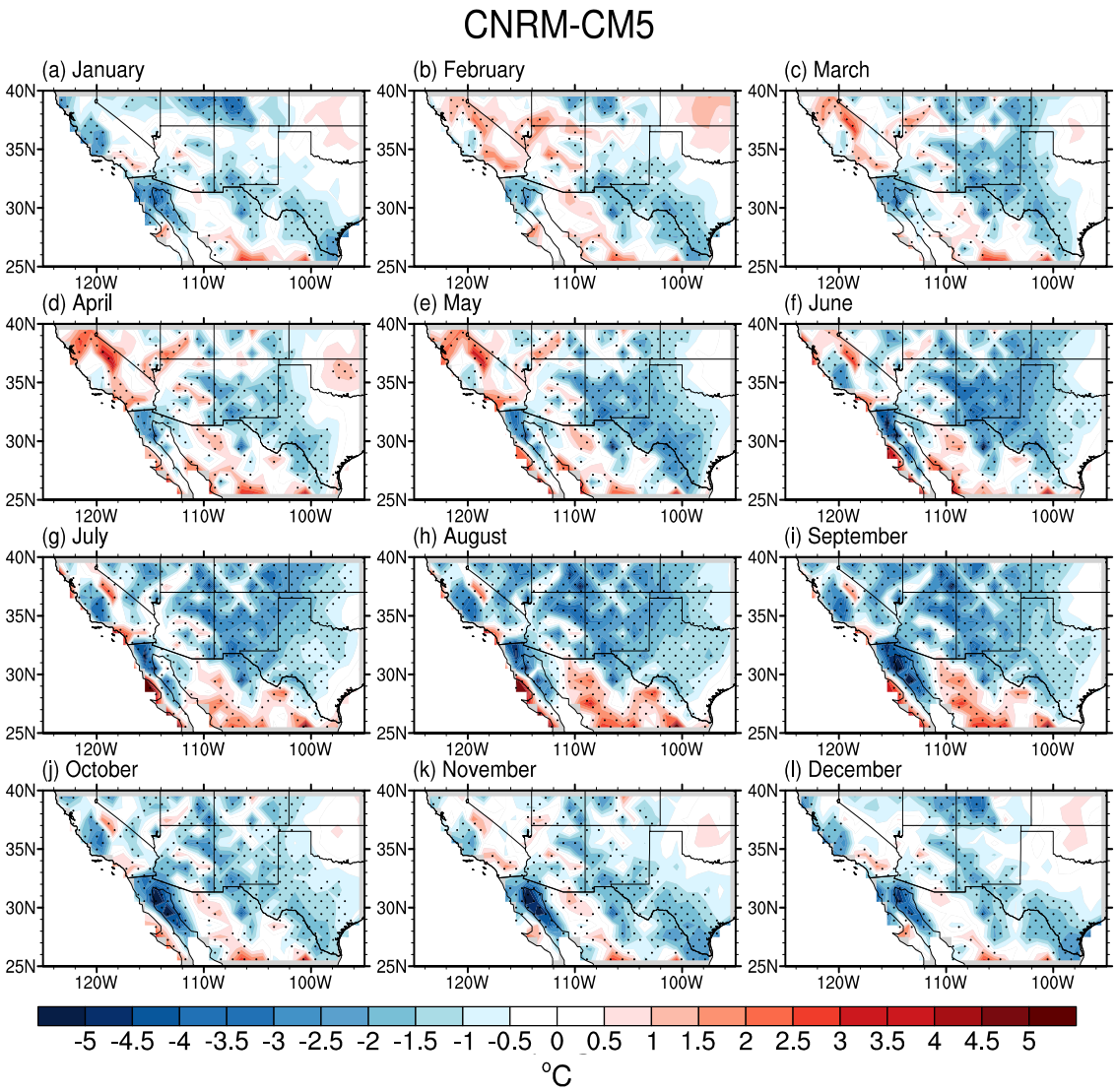


Figure A2-42. Monthly temperature differences between CNRM-CM5 output and NARR with a statistical significance of 90% (stippled).

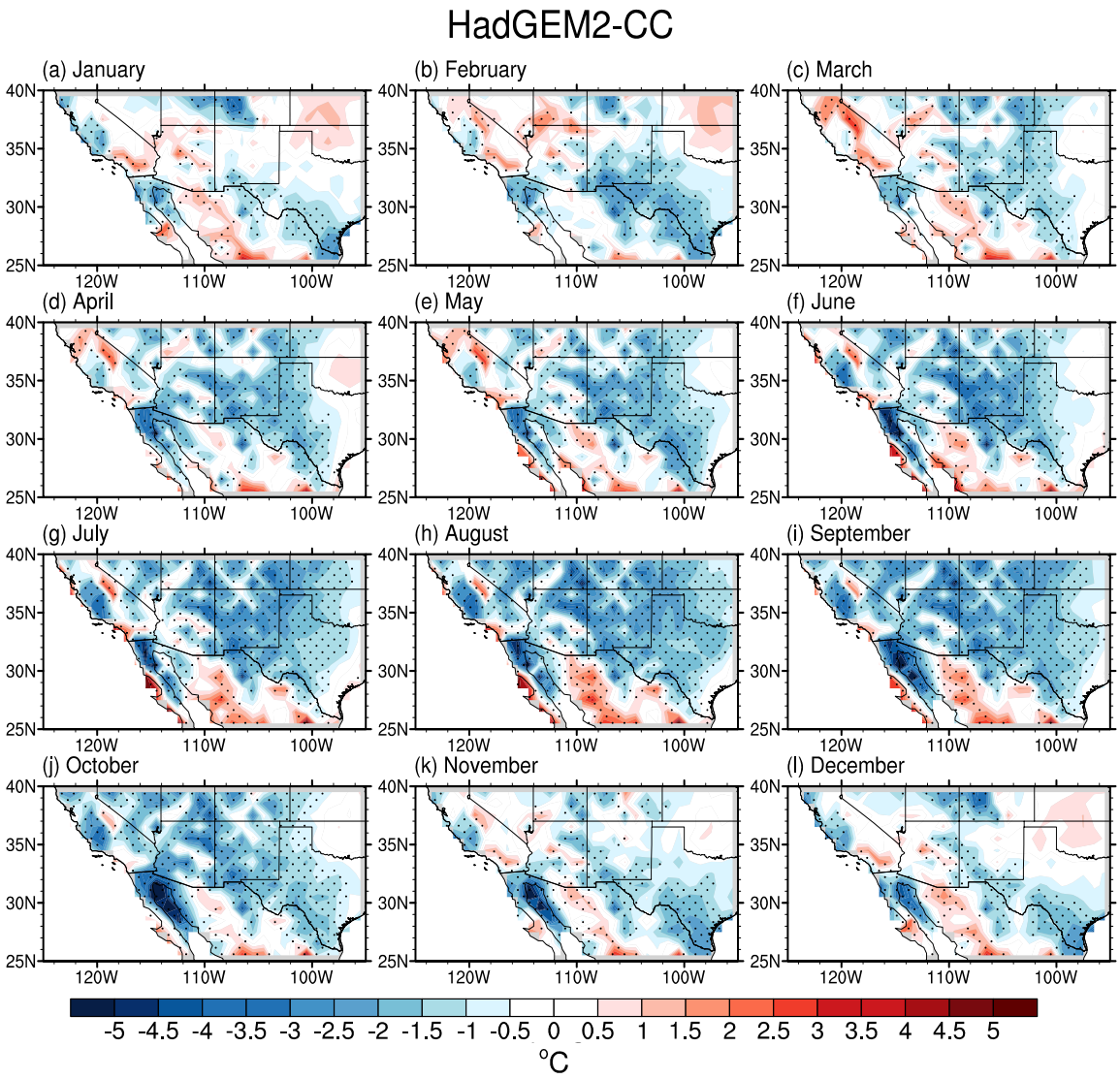


Figure A2-43. Monthly temperature differences between HadGEM2-CC output and NARR with a statistical significance of 90% (stippled).

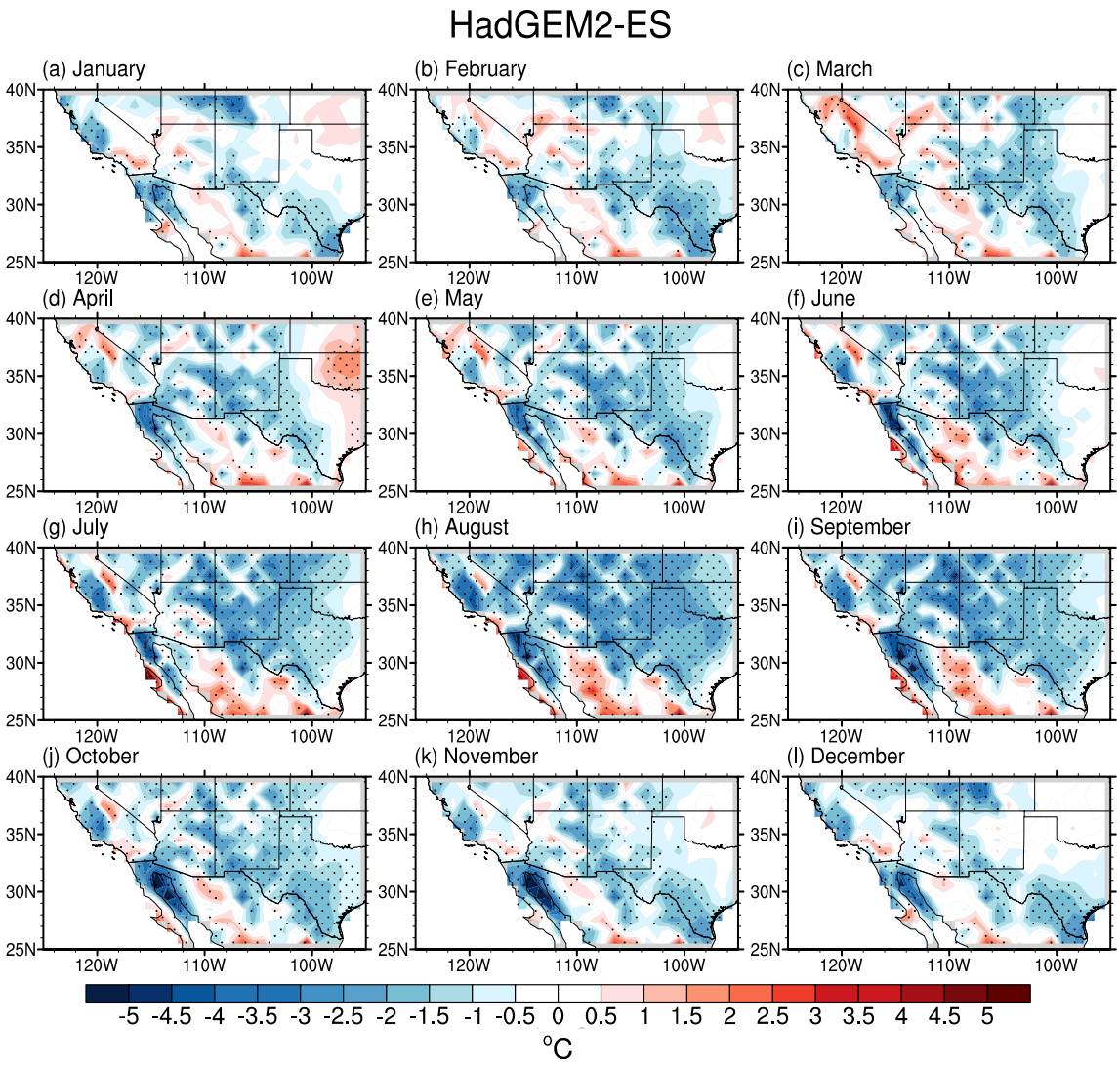


Figure A2-44. Monthly temperature differences between HadGEM2-ES output and NARR with a statistical significance of 90% (stippled).

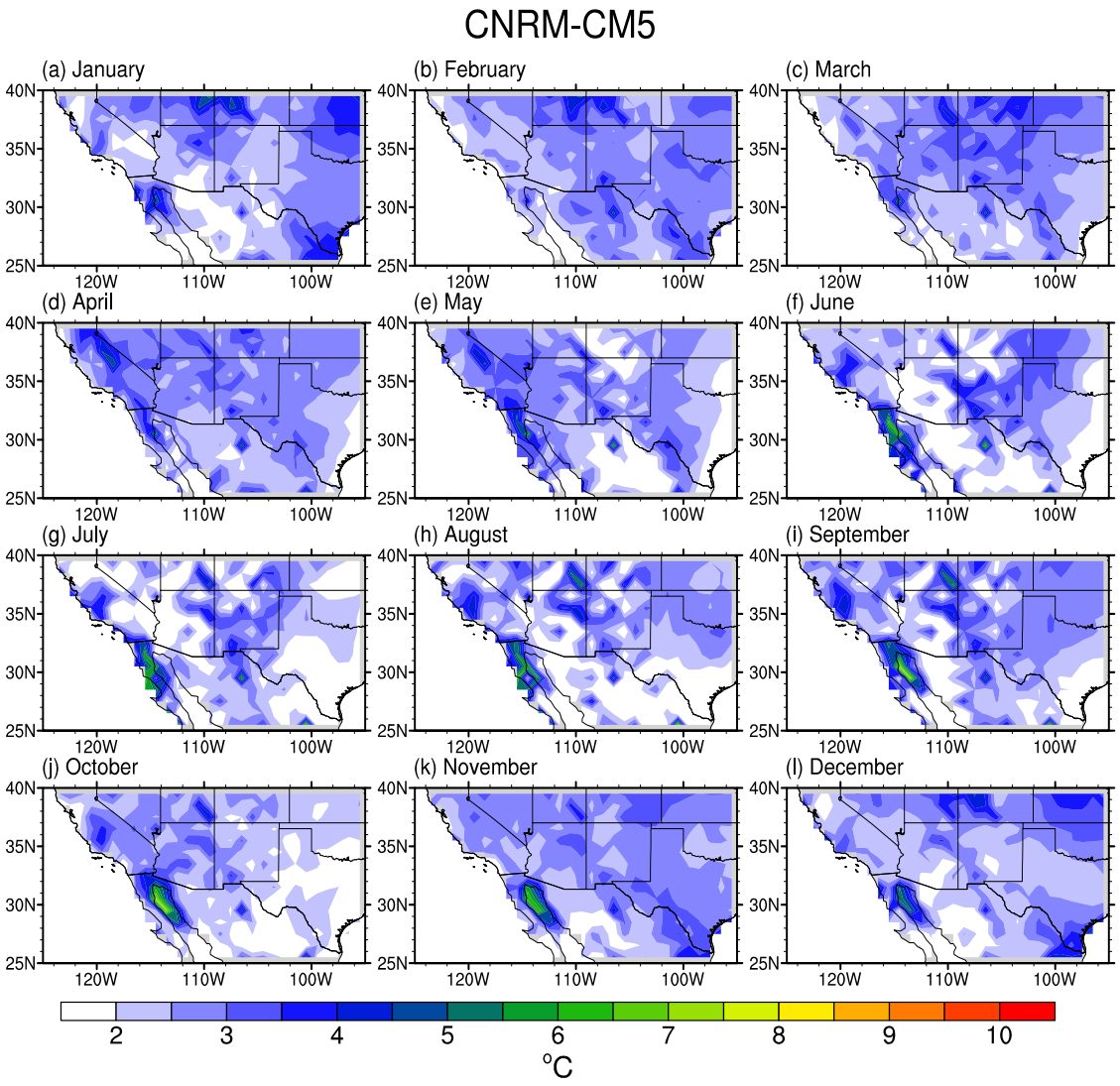


Figure A2-45. Monthly temperature root mean square difference (deg C) between the CNRM-CM5 downscaled output and NARR.

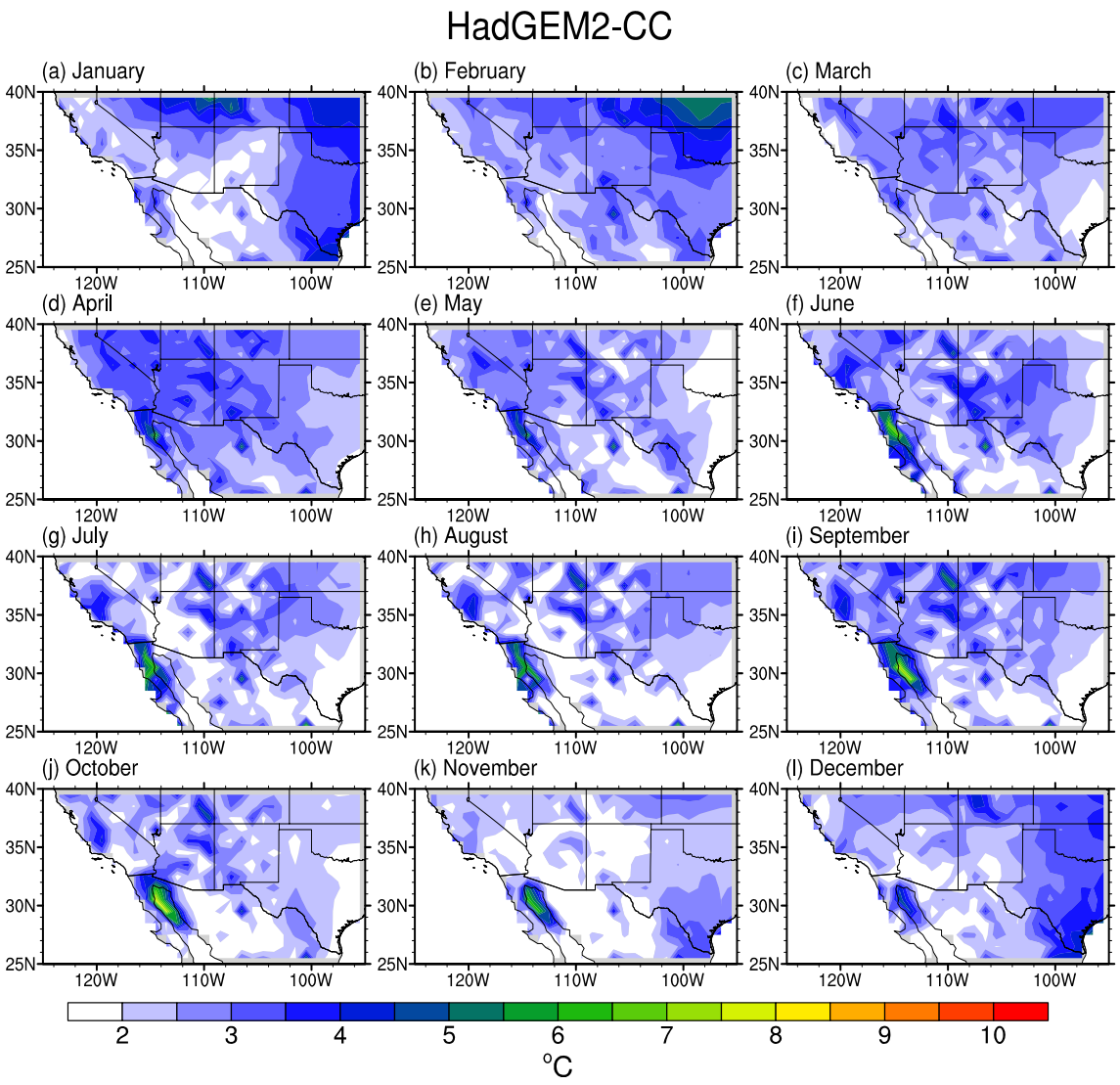


Figure A2-46. Monthly temperature root mean square difference (deg C) between the HadGEM2-CC downscaled output and NARR.

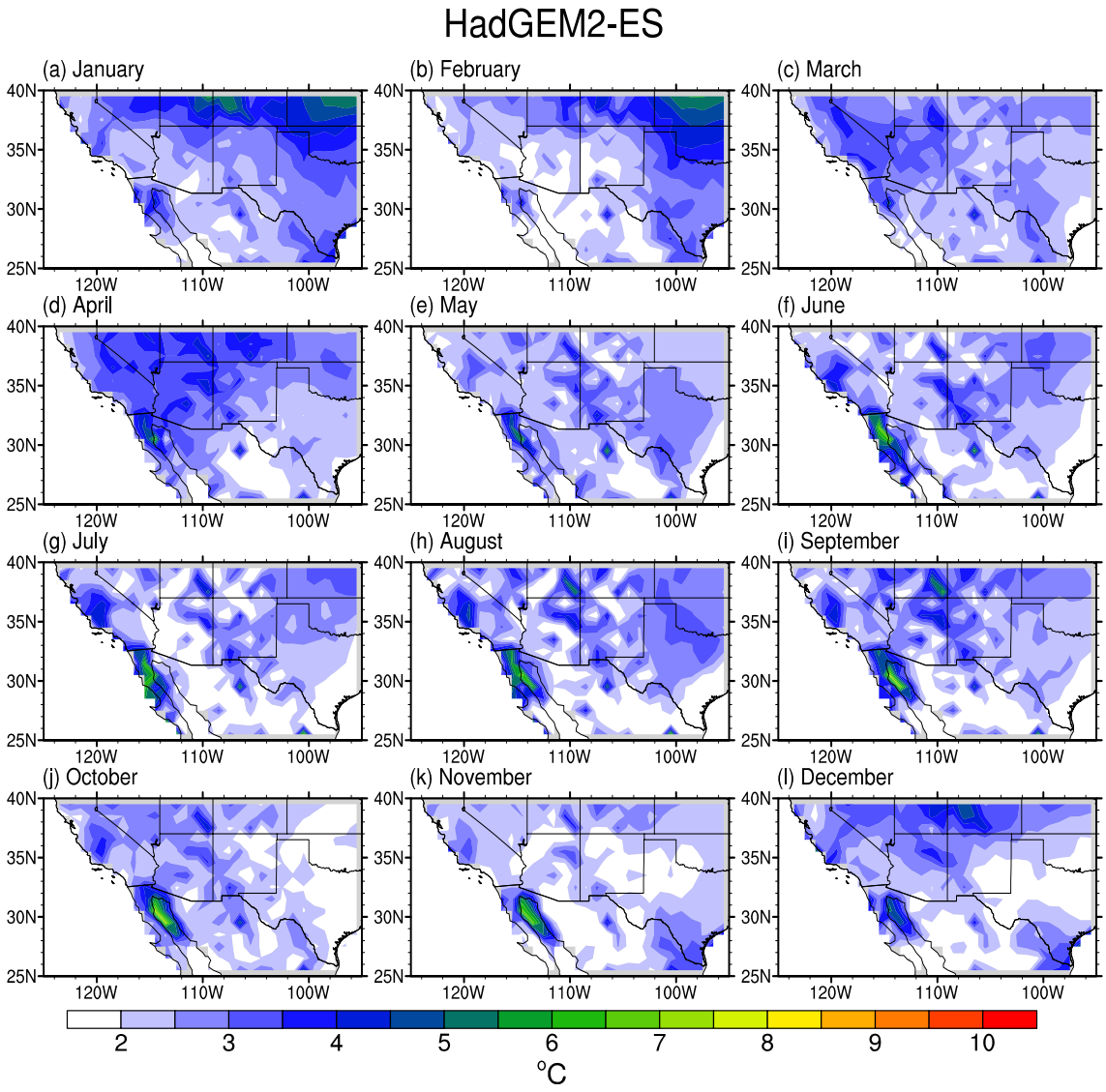


Figure A2-47. Monthly temperature root mean square difference (deg C) between the HadGEM2-ES downscaled output and NARR.

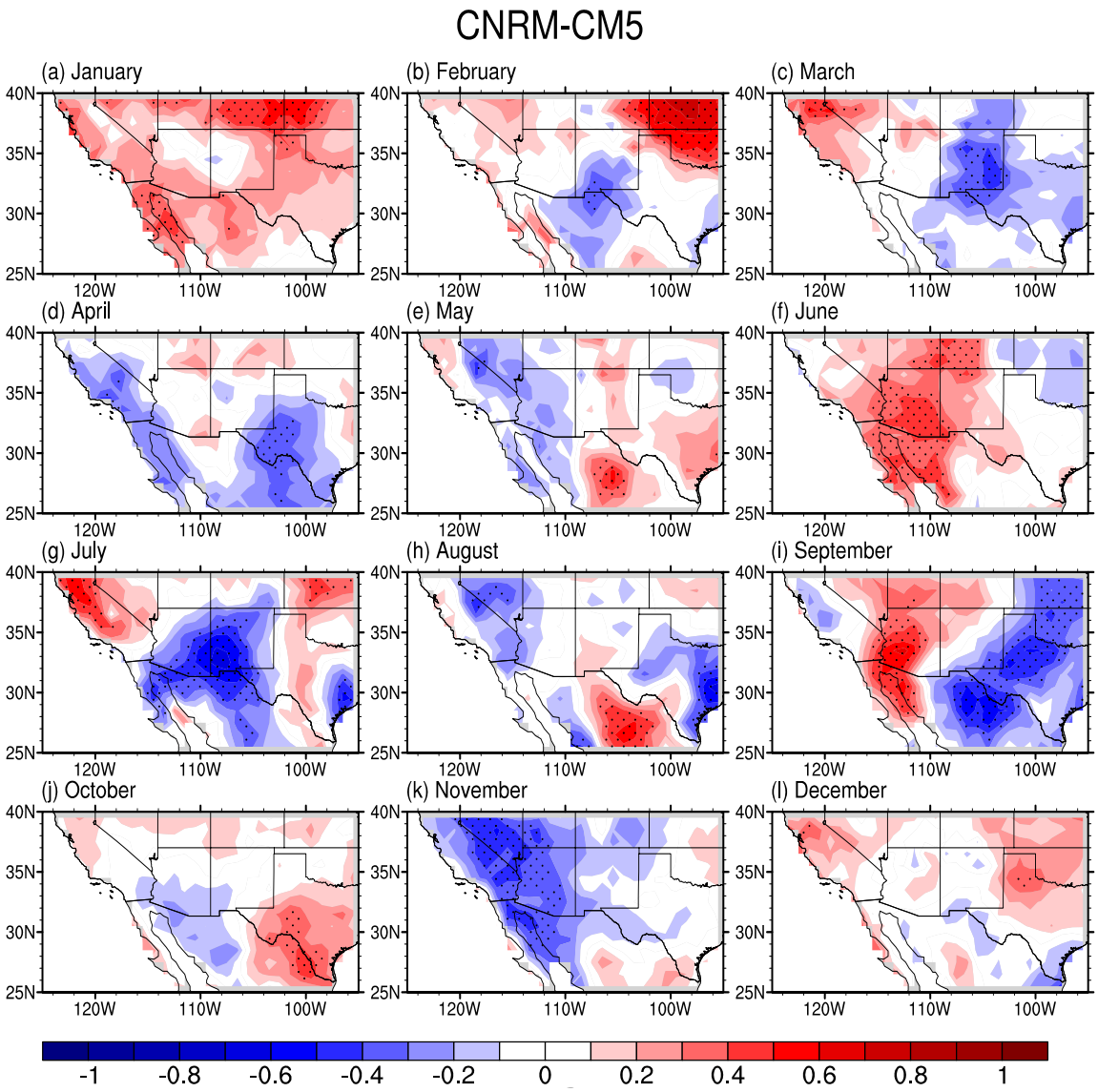


Figure A2-48. Monthly temperature correlations between CNRM-CM5 output and NARR with a statistical significance of 90% (stippled).

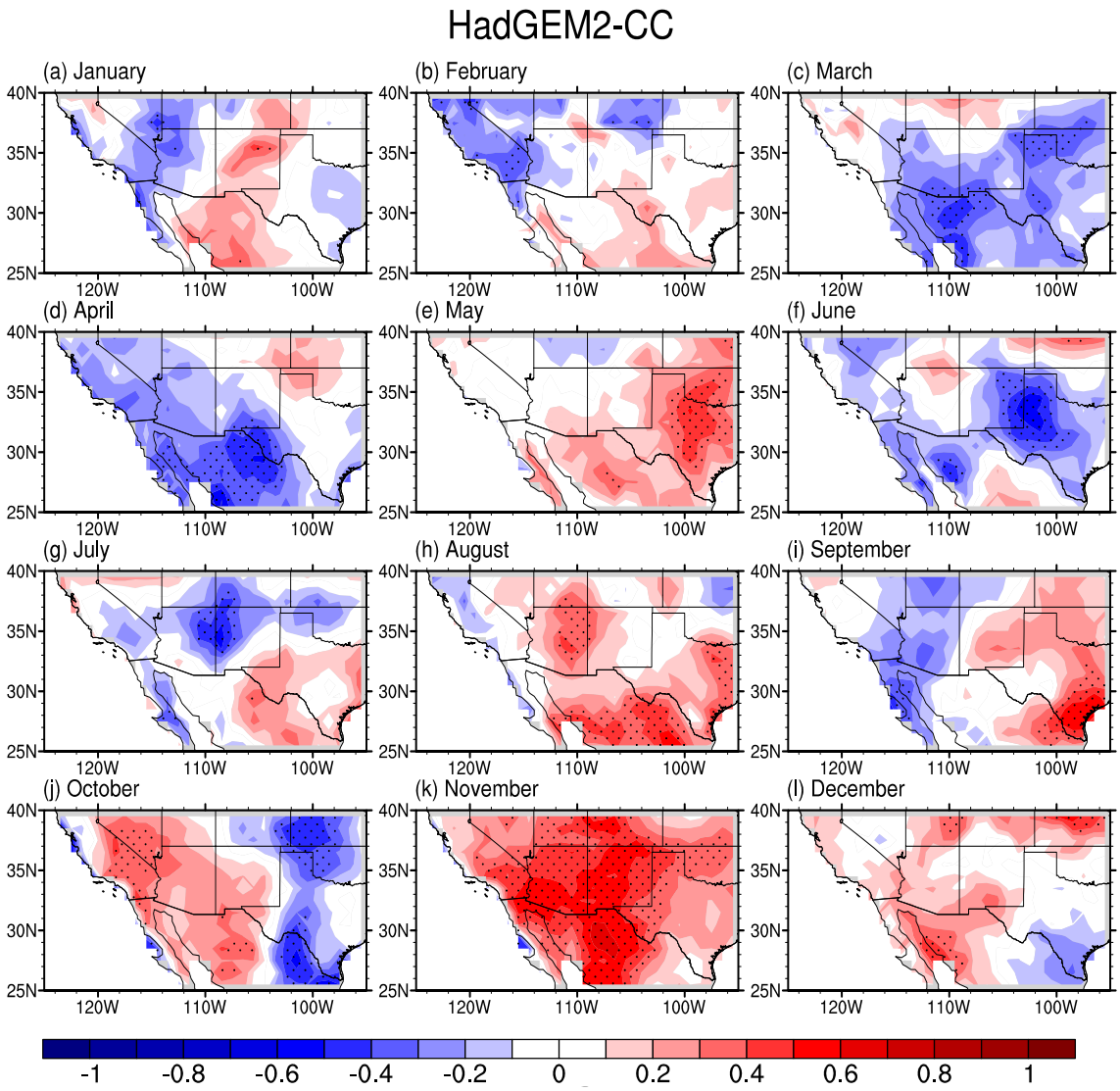


Figure A2-49. Monthly temperature correlations between HadGEM2-CC output and NARR with a statistical significance of 90% (stippled).

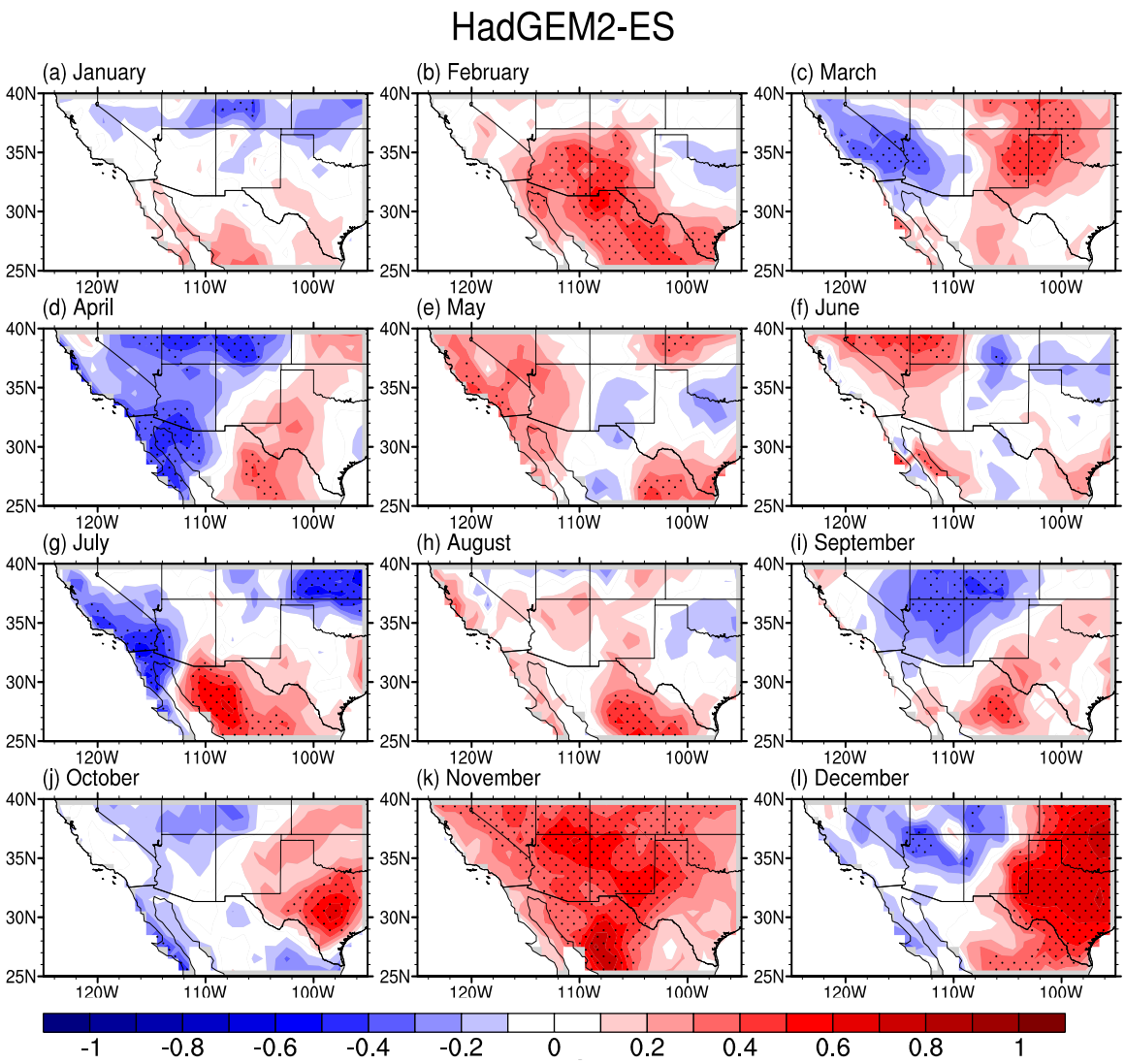


Figure A2-50. Monthly temperature correlations between HadGEM2-ES output and NARR with a statistical significance of 90% (stippled).

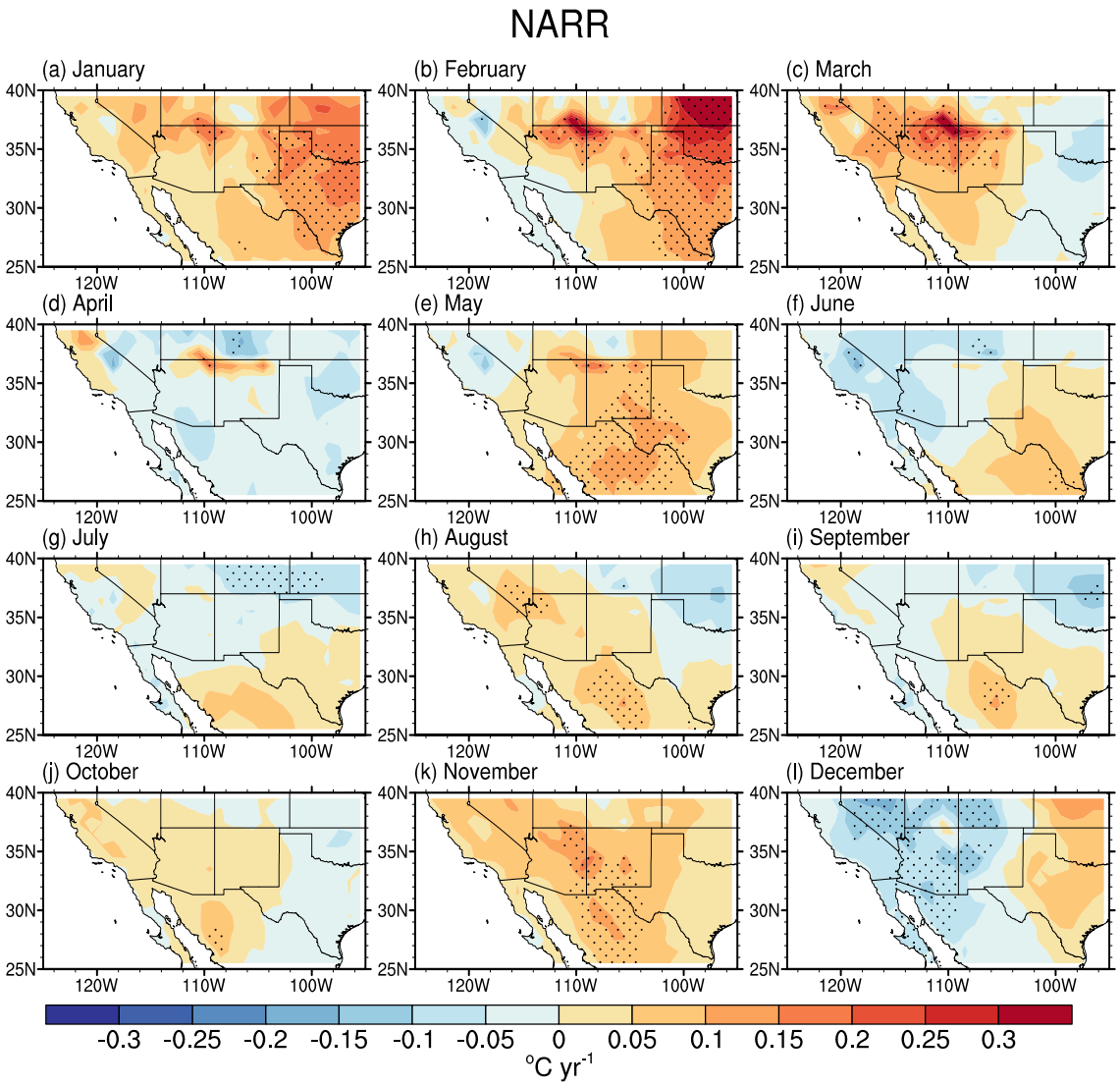


Figure A2-51. Monthly temperature trends for NARR. Stippled regions indicate statistically significance (90% level).

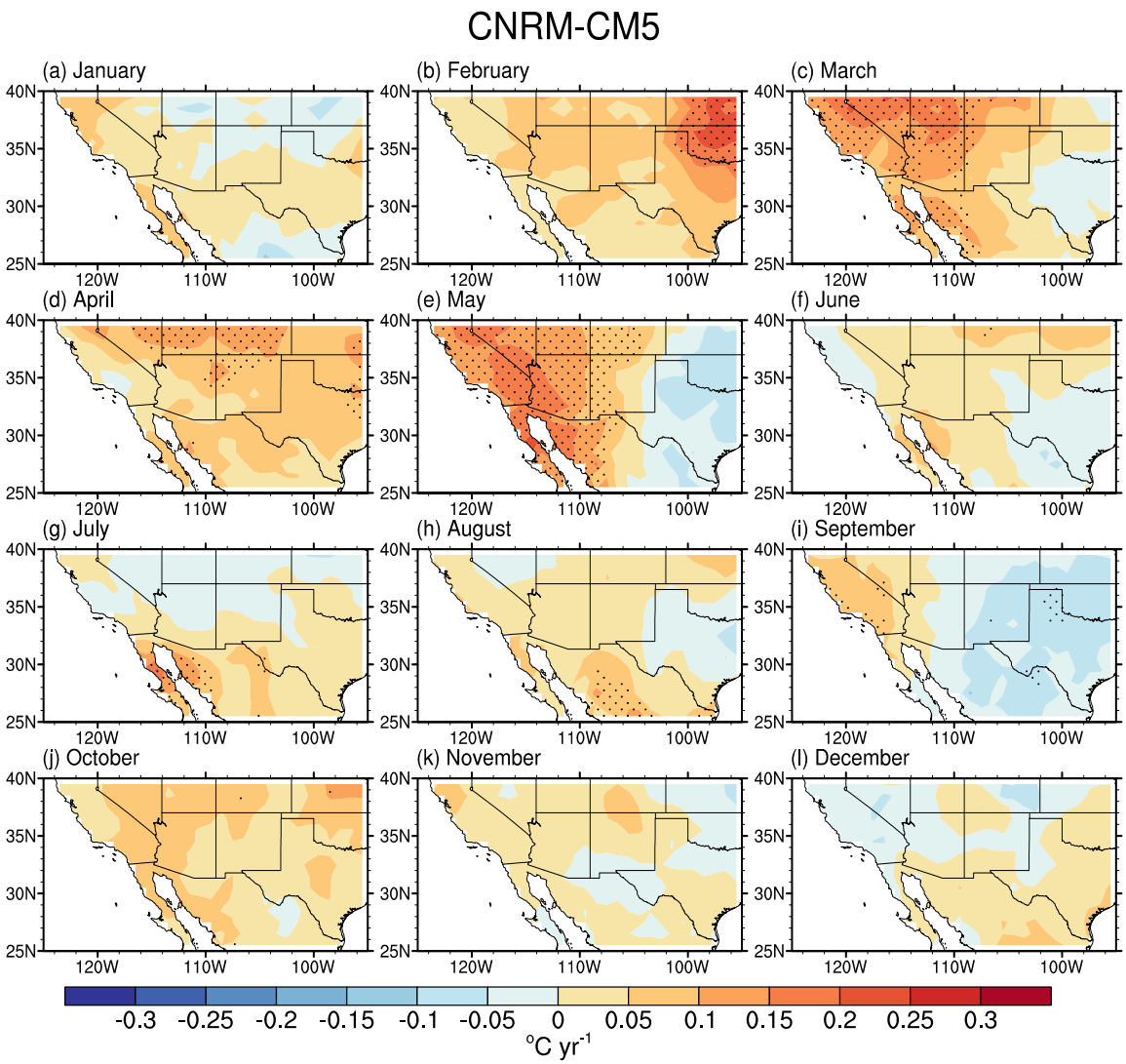


Figure A2-52. Monthly temperature trends for CNRM-CM5. Stippled regions indicate statistically significance (90% level).

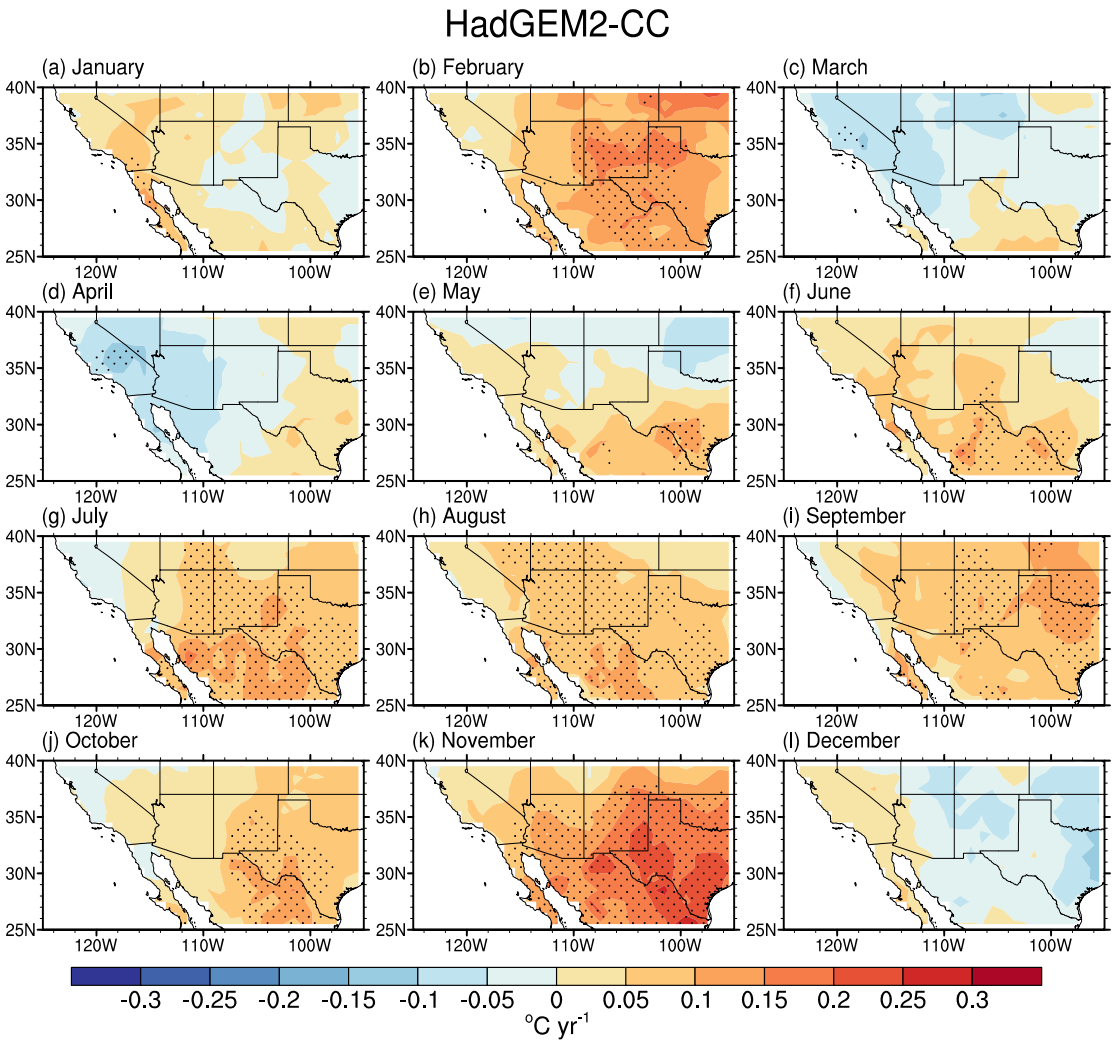


Figure A2-53. Monthly temperature trends for HadGEM2-CC. Stippled regions indicate statistically significance (90% level).

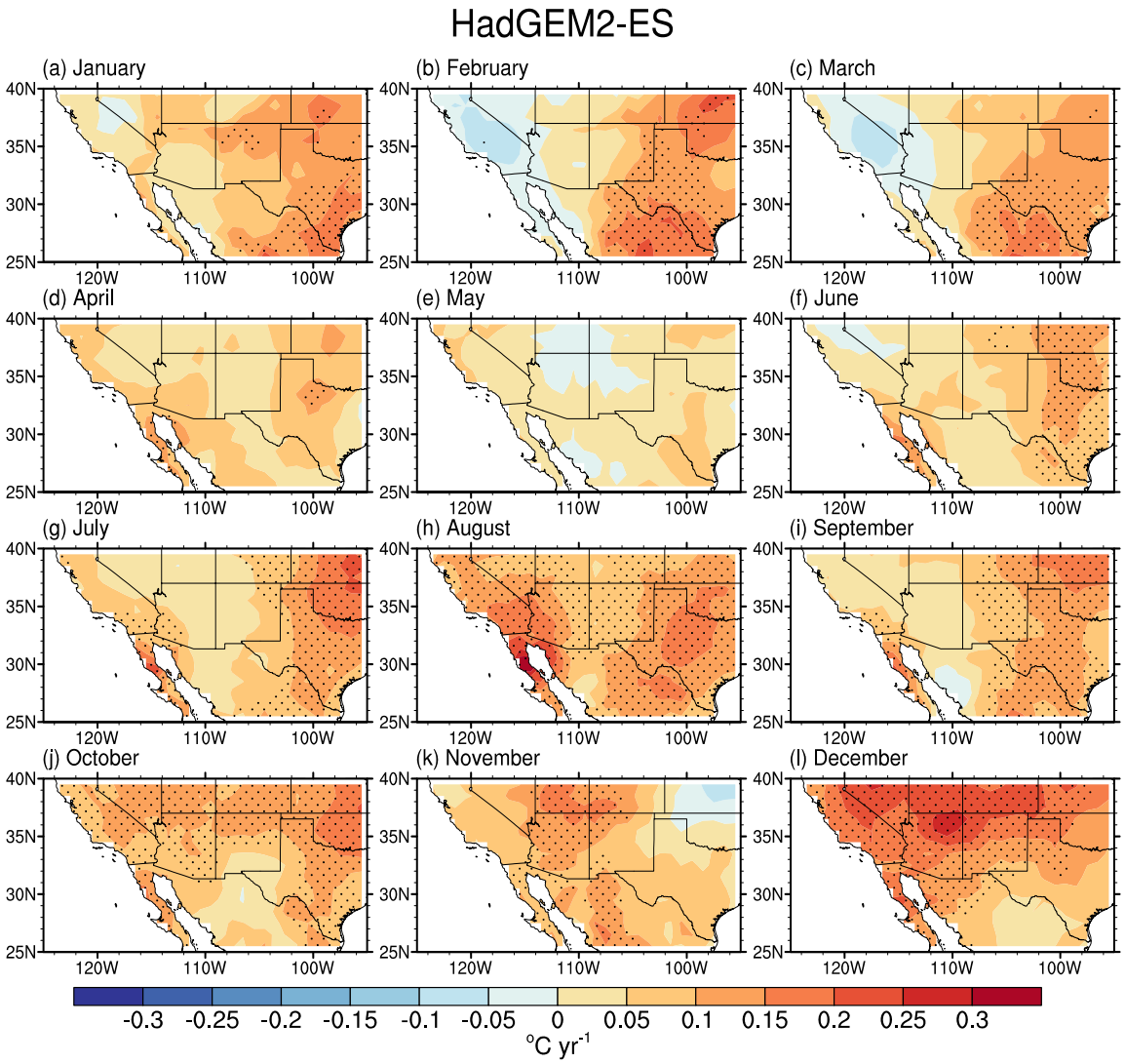


Figure A2-54. Monthly temperature trends for HadGEM2-ES. Stippled regions indicate statistically significance (90% level).

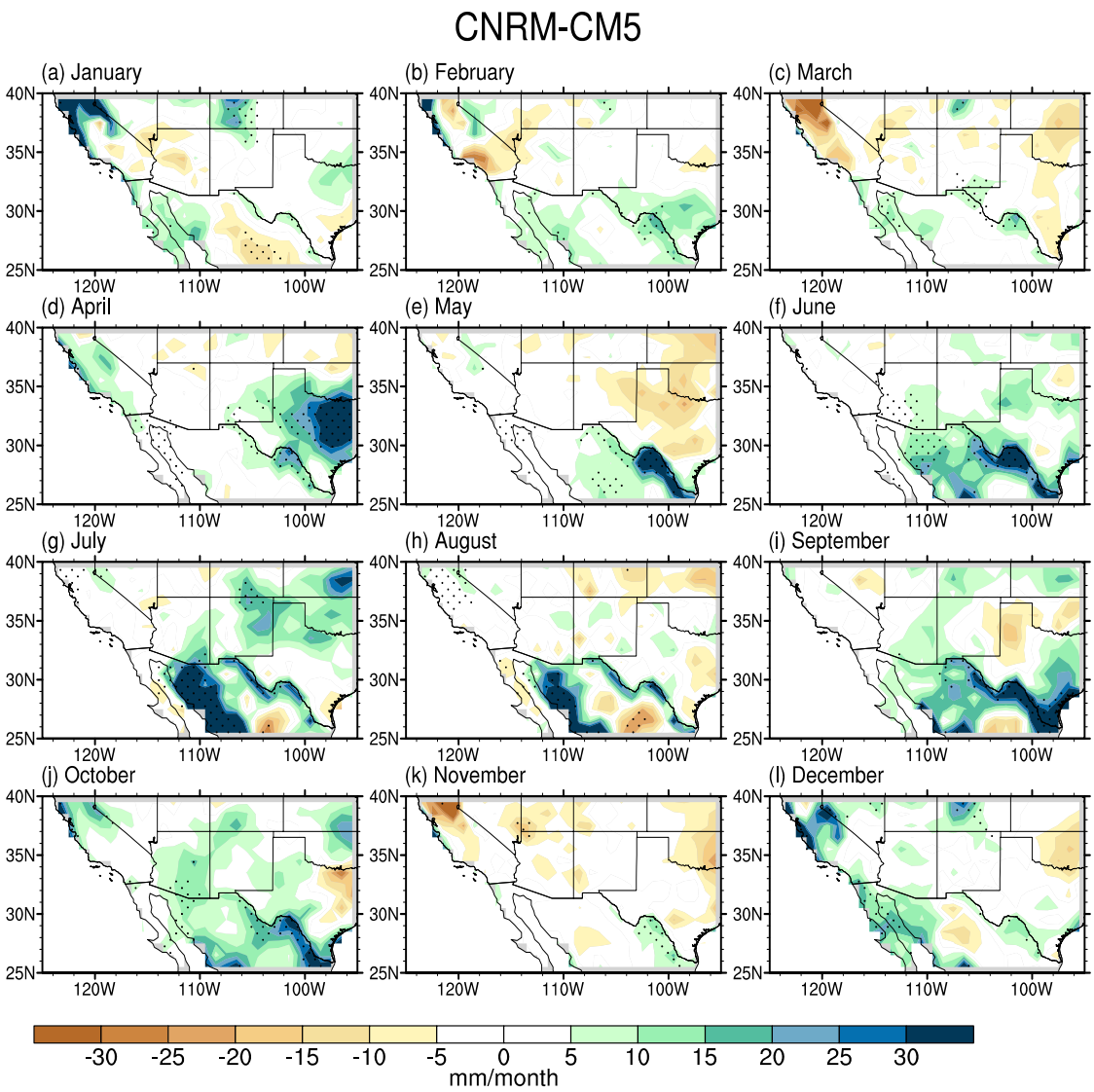


Figure A2-55. Monthly precipitation differences (mm/month) between CNRM-CM5 output and NARR with a statistical significance of 90% (stippled).

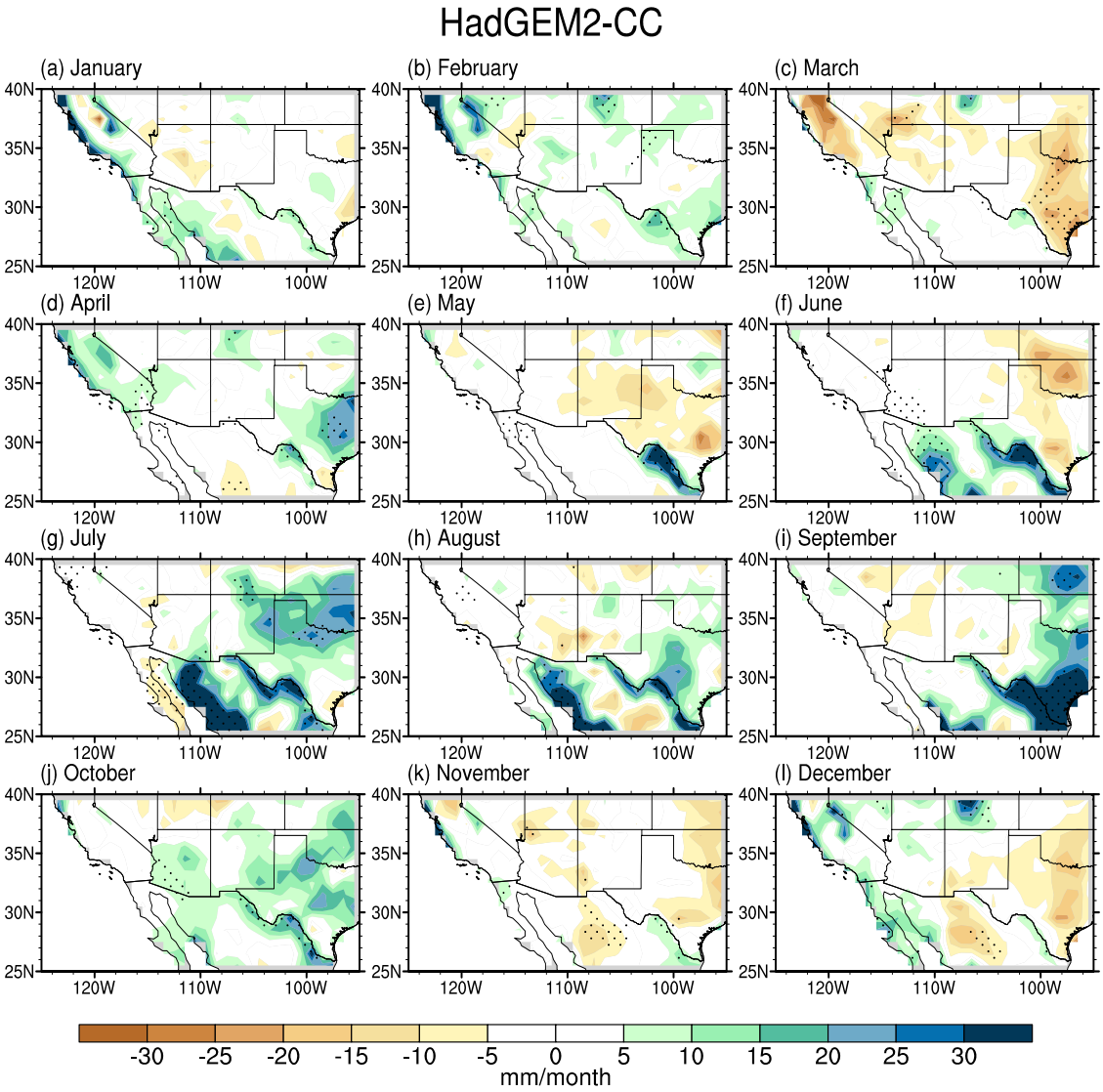


Figure A2-56. Monthly precipitation differences (mm/month) between HadGEM2-CC output and NARR with a statistical significance of 90% (stippled).

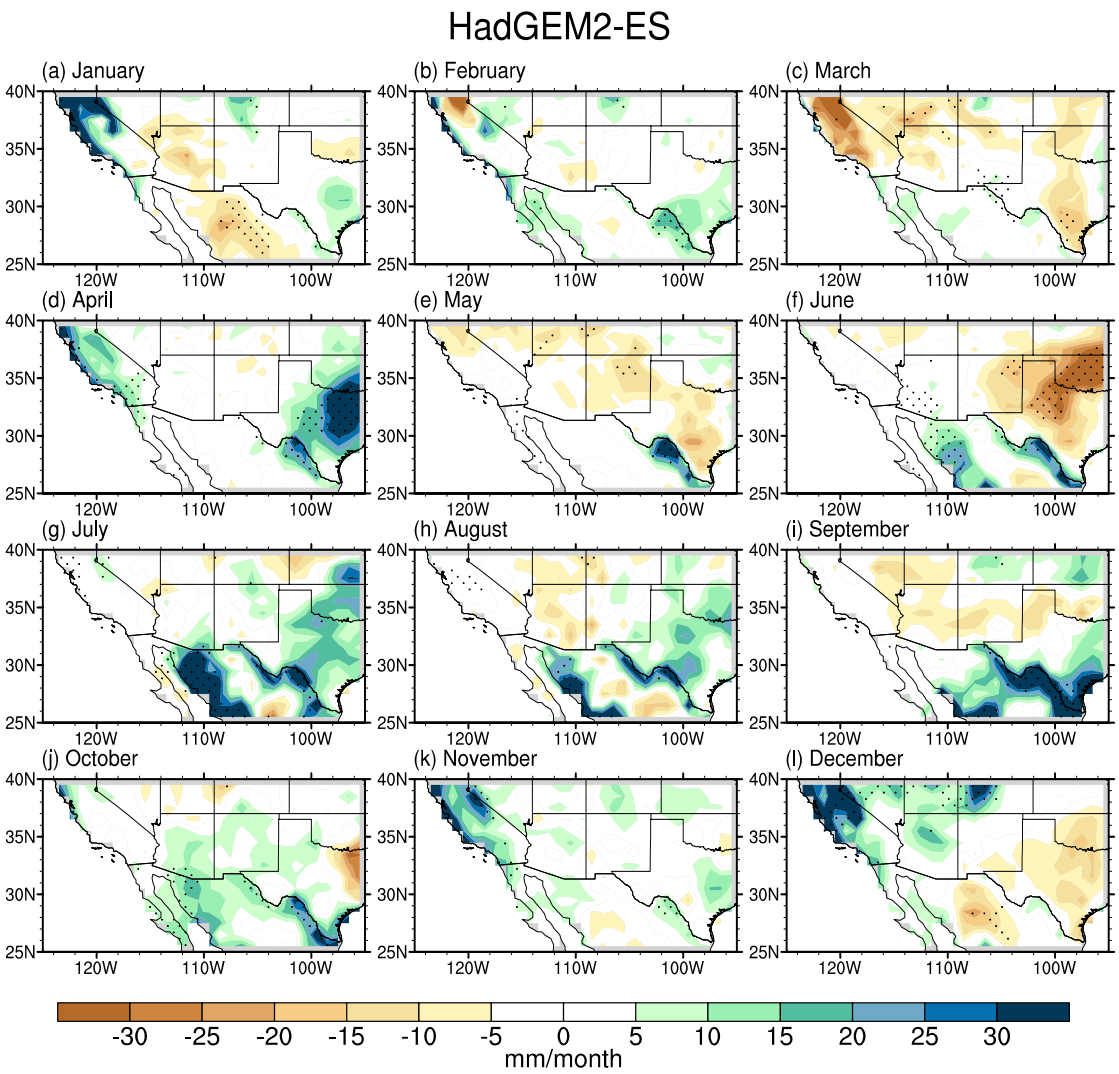


Figure A2-57. Monthly precipitation differences (mm/month) between HadGEM2-ES output and NARR with a statistical significance of 90% (stippled).

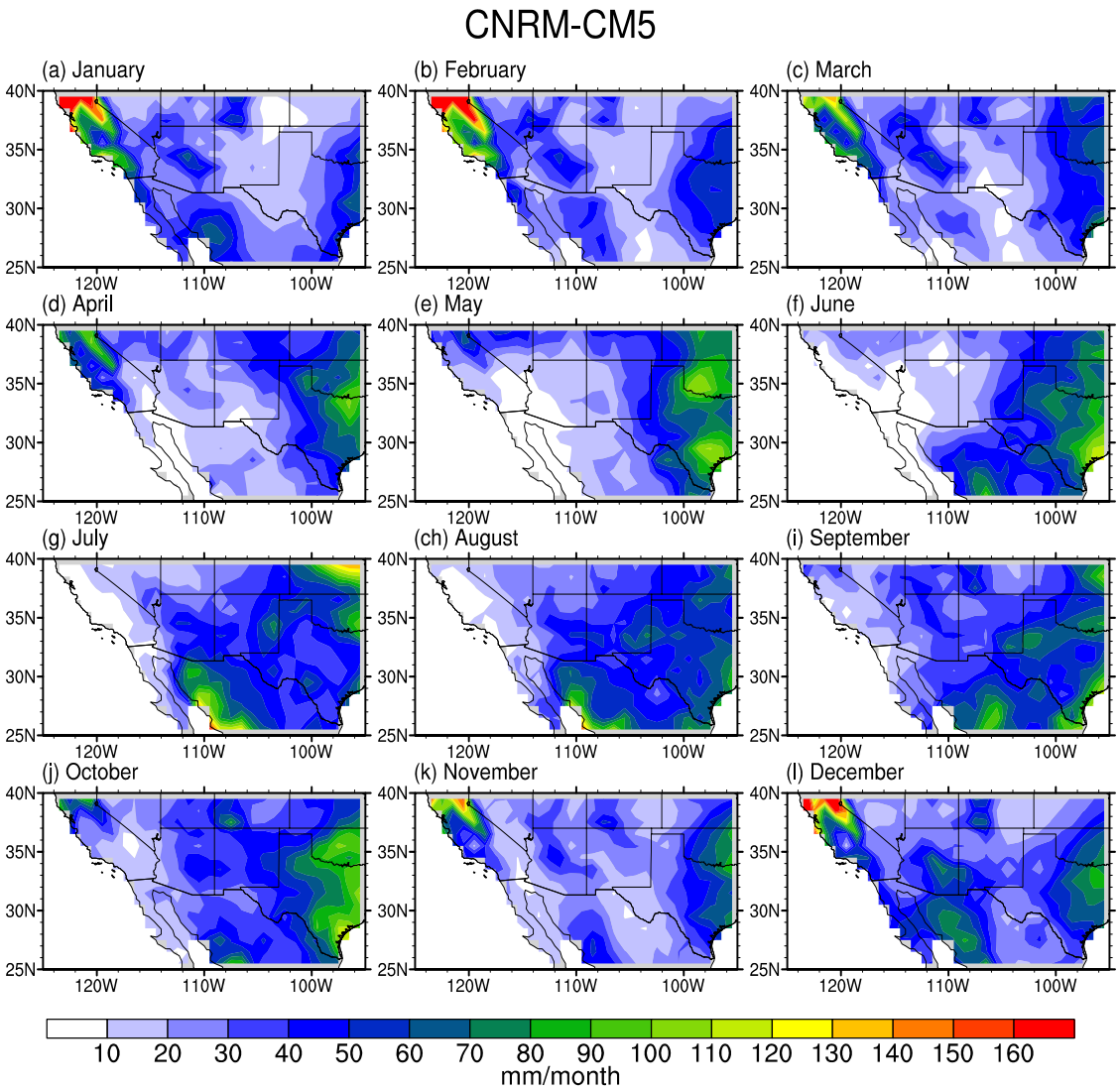


Figure A2-58. Monthly precipitation root mean square difference (mm/month) between the CNRM-CM5 downscaled output and NARR.

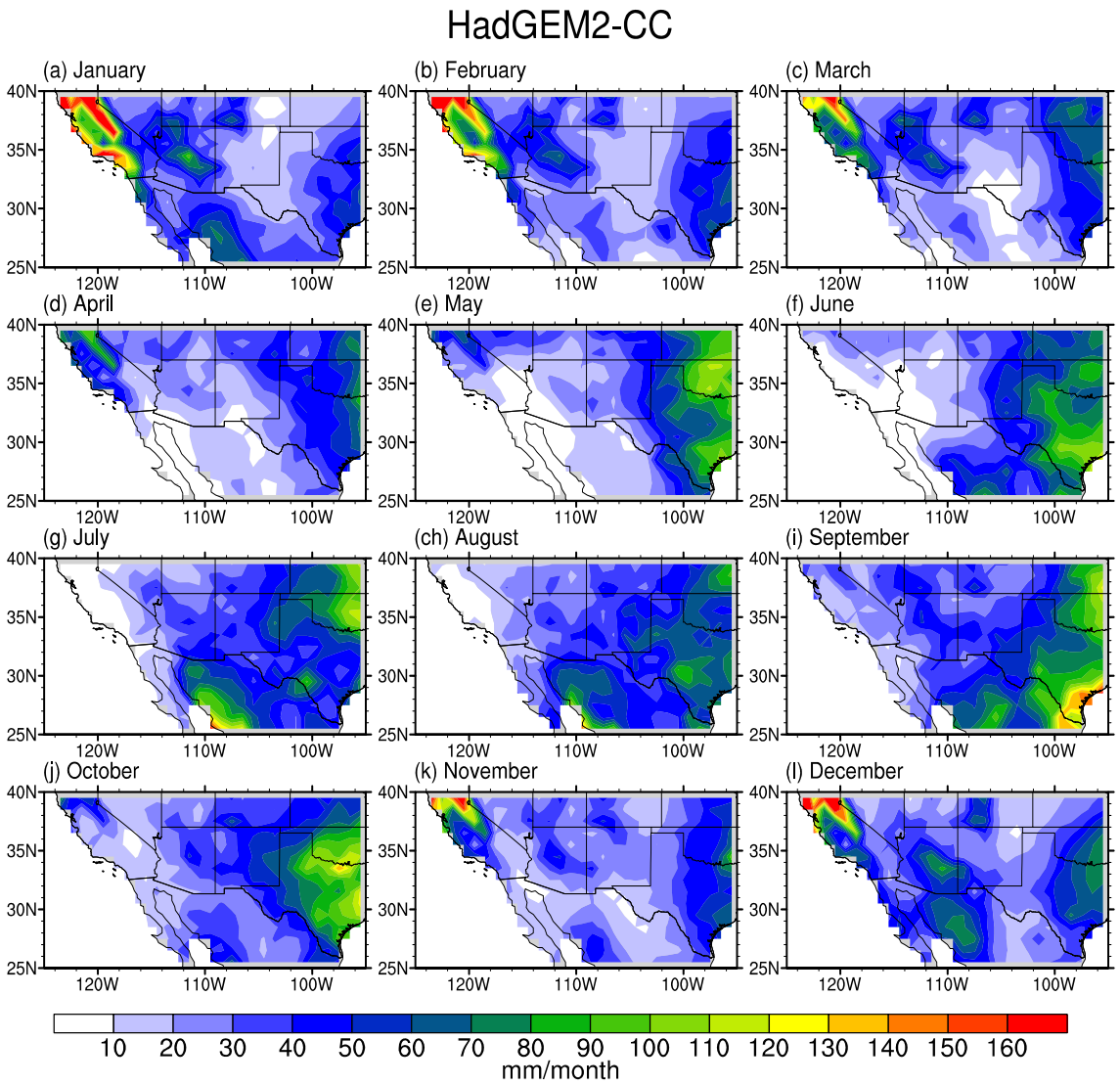


Figure A2-59. Monthly precipitation root mean square difference (mm/month) between the HadGEM2-CC downscaled output and NARR.

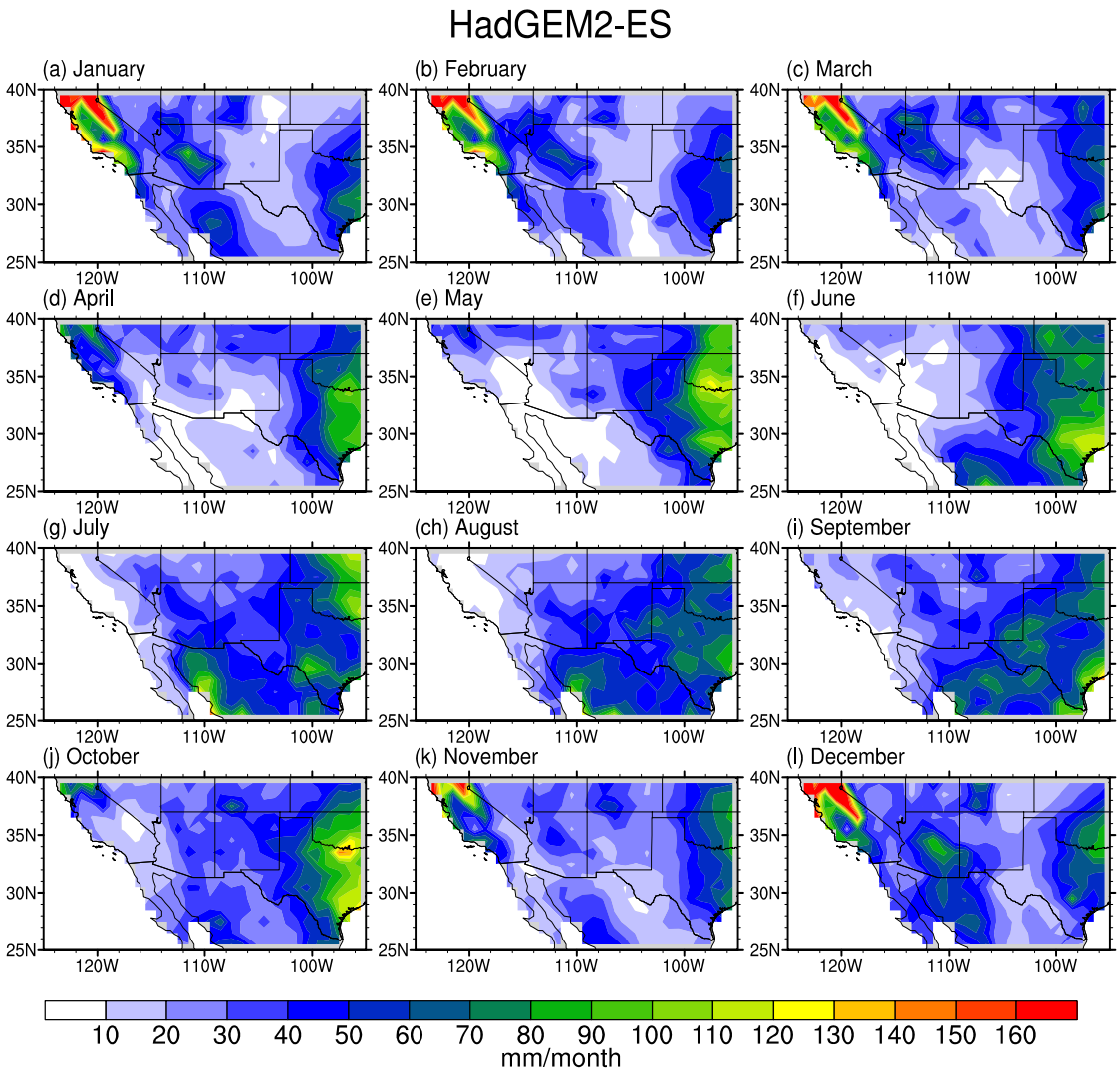


Figure A2-60. Monthly precipitation root mean square difference (mm/month) between the HadGEM2-ES downscaled output and NARR.

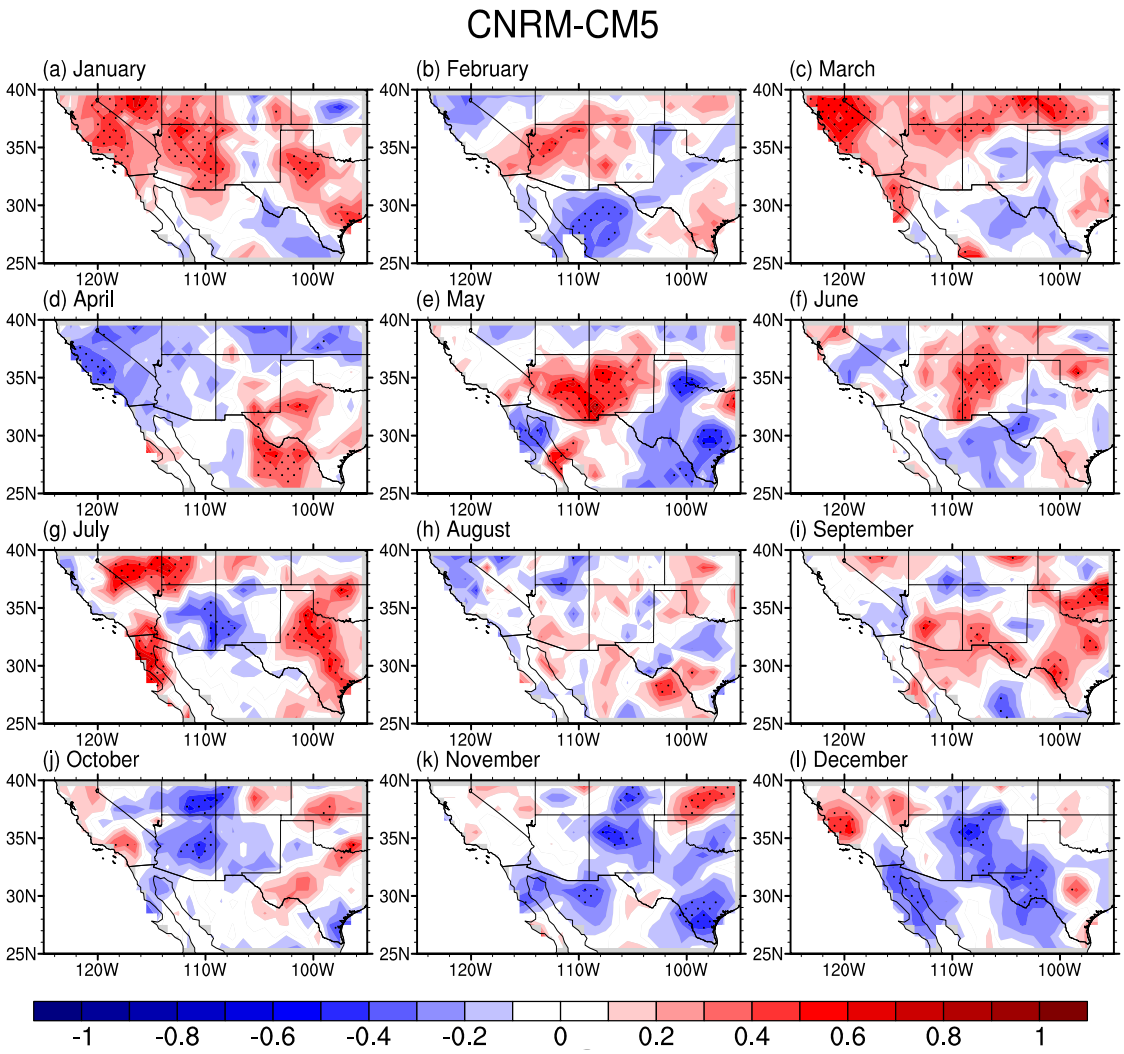


Figure A2-61. Monthly precipitation correlations between CNRM-CM5 output and NARR with a statistical significance of 90% (stippled).

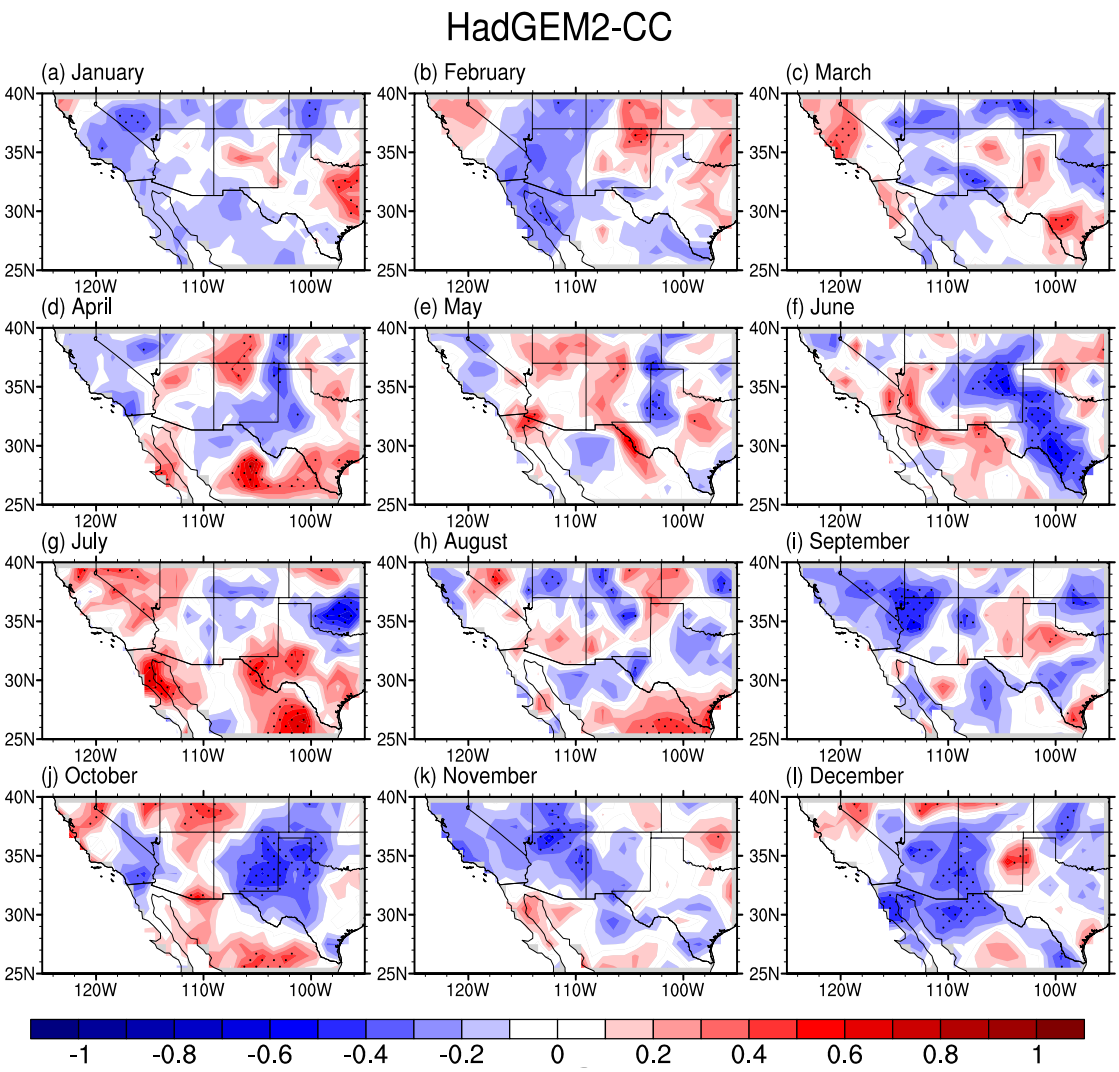


Figure A2-62. Monthly precipitation correlations between HadGEM2-CC output and NARR with a statistical significance of 90% (stippled).

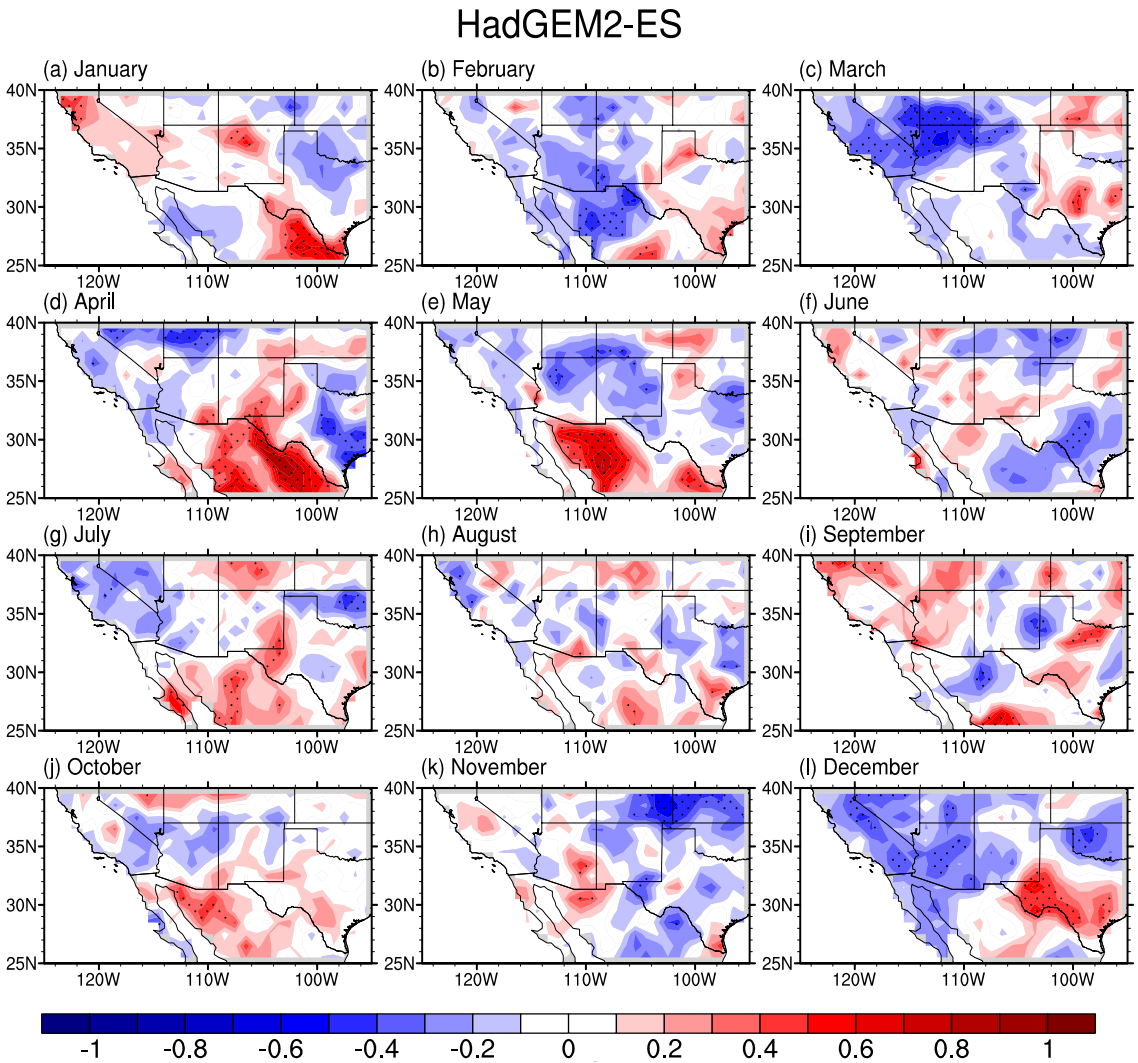


Figure A2-63. Monthly precipitation correlations between HadGEM2-ES output and NARR with a statistical significance of 90% (stippled).

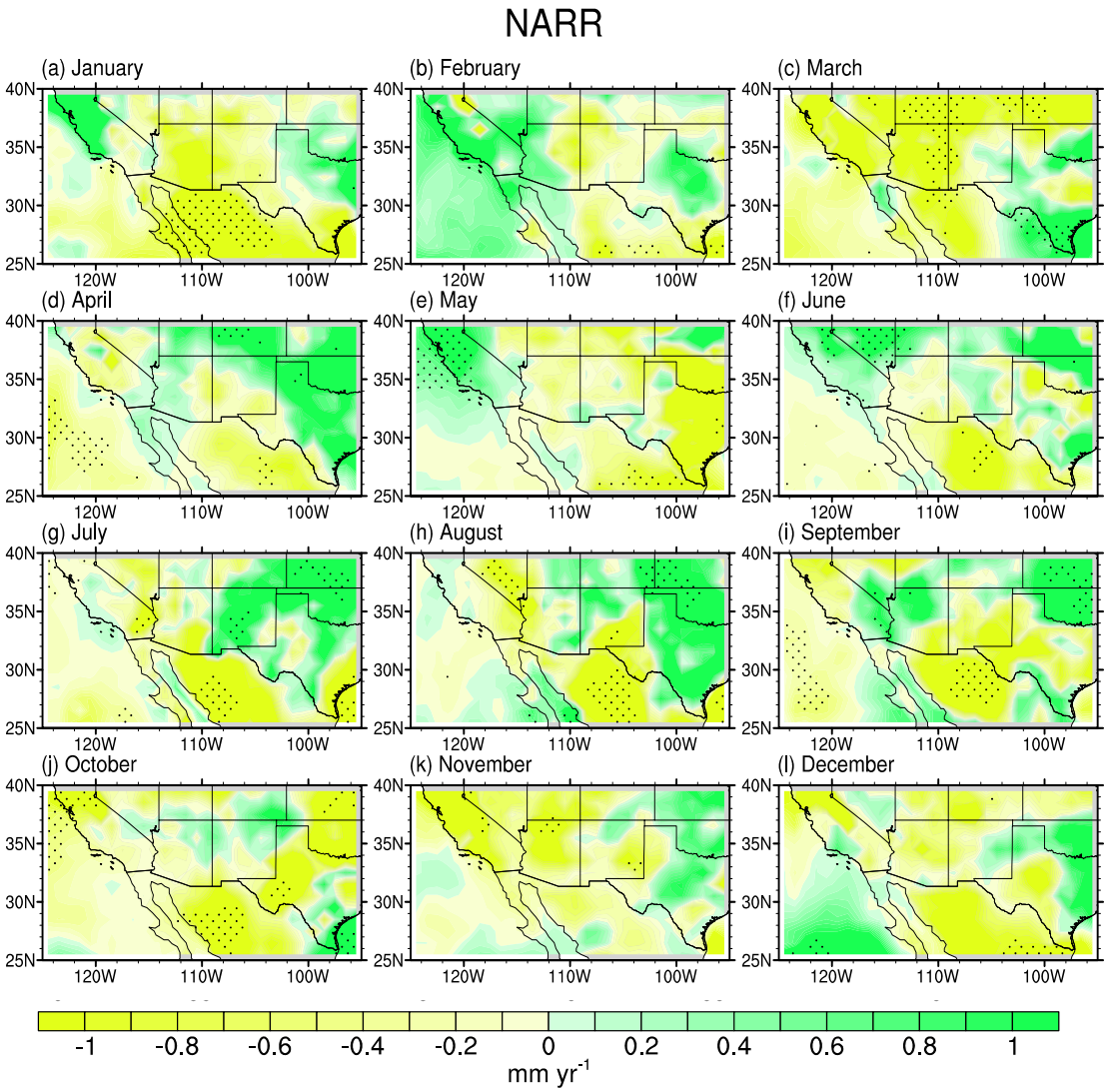


Figure A2-64. Monthly precipitation trends for CNRM-CM5. Stippled regions indicate statistically significance (90% level).

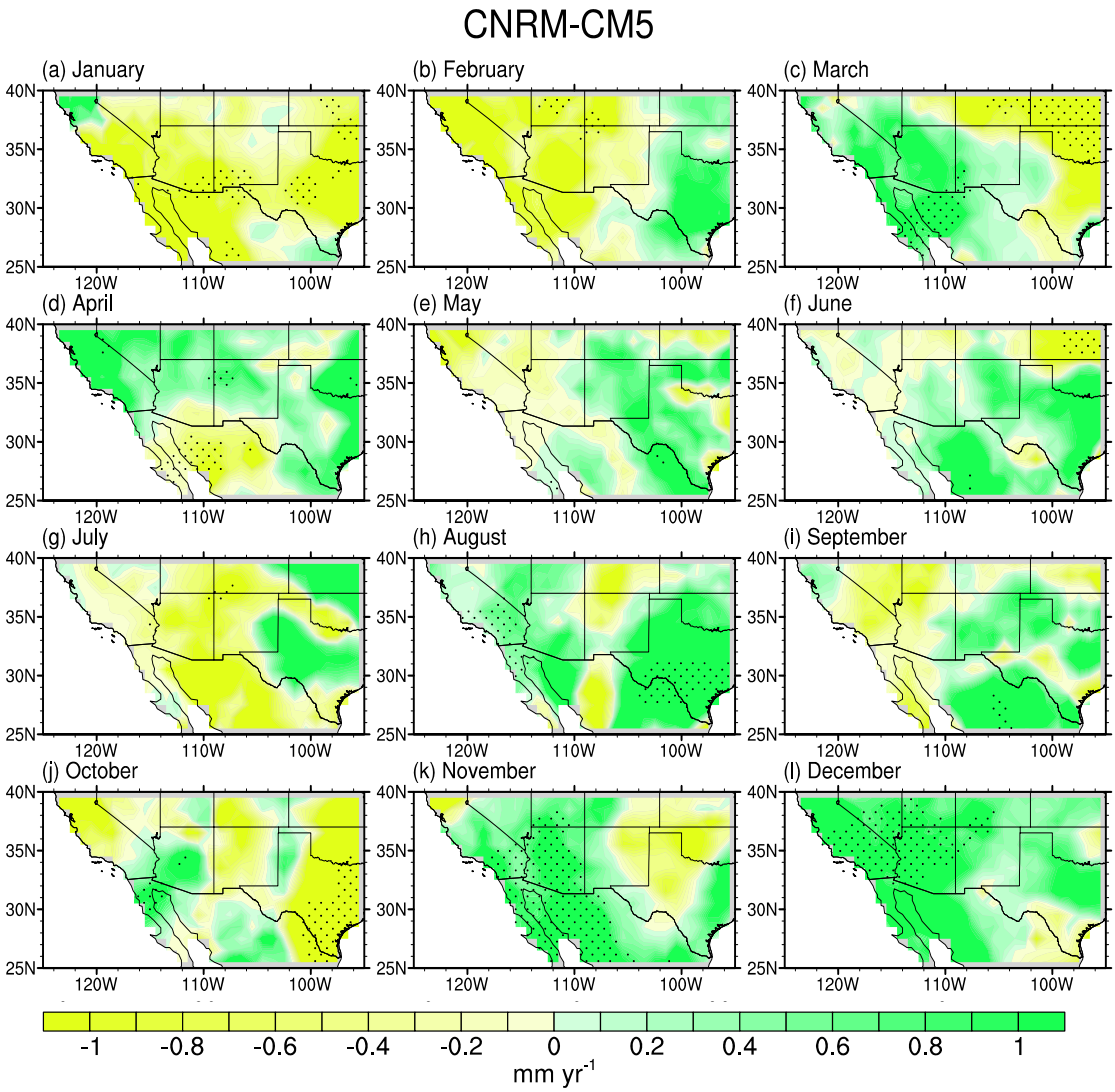


Figure A2-65. Monthly precipitation trends for CNRM-CM5. Stippled regions indicate statistically significance (90% level).

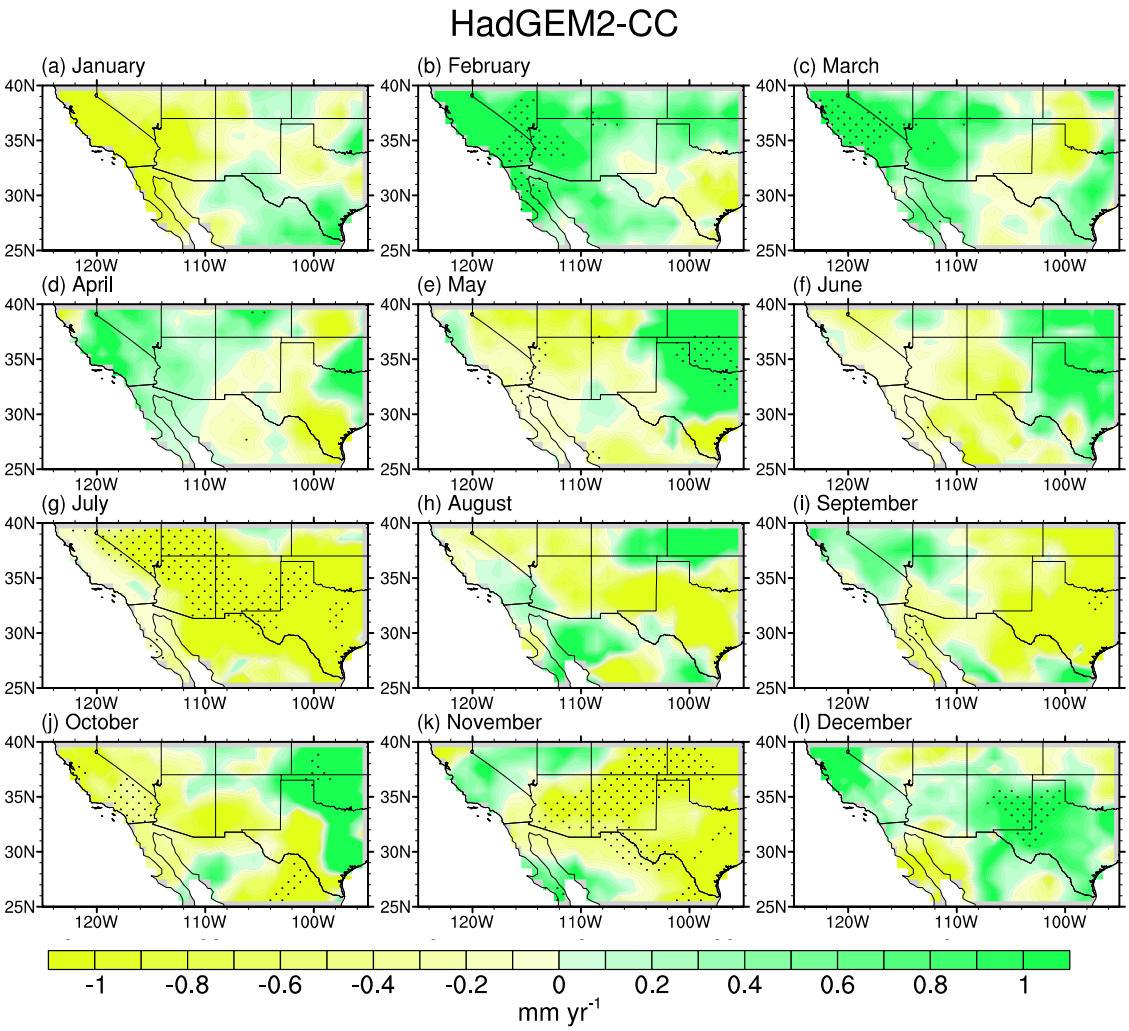


Figure A2-66. Monthly precipitation trends for HadGEM2-CC. Stippled regions indicate statistically significance (90% level).

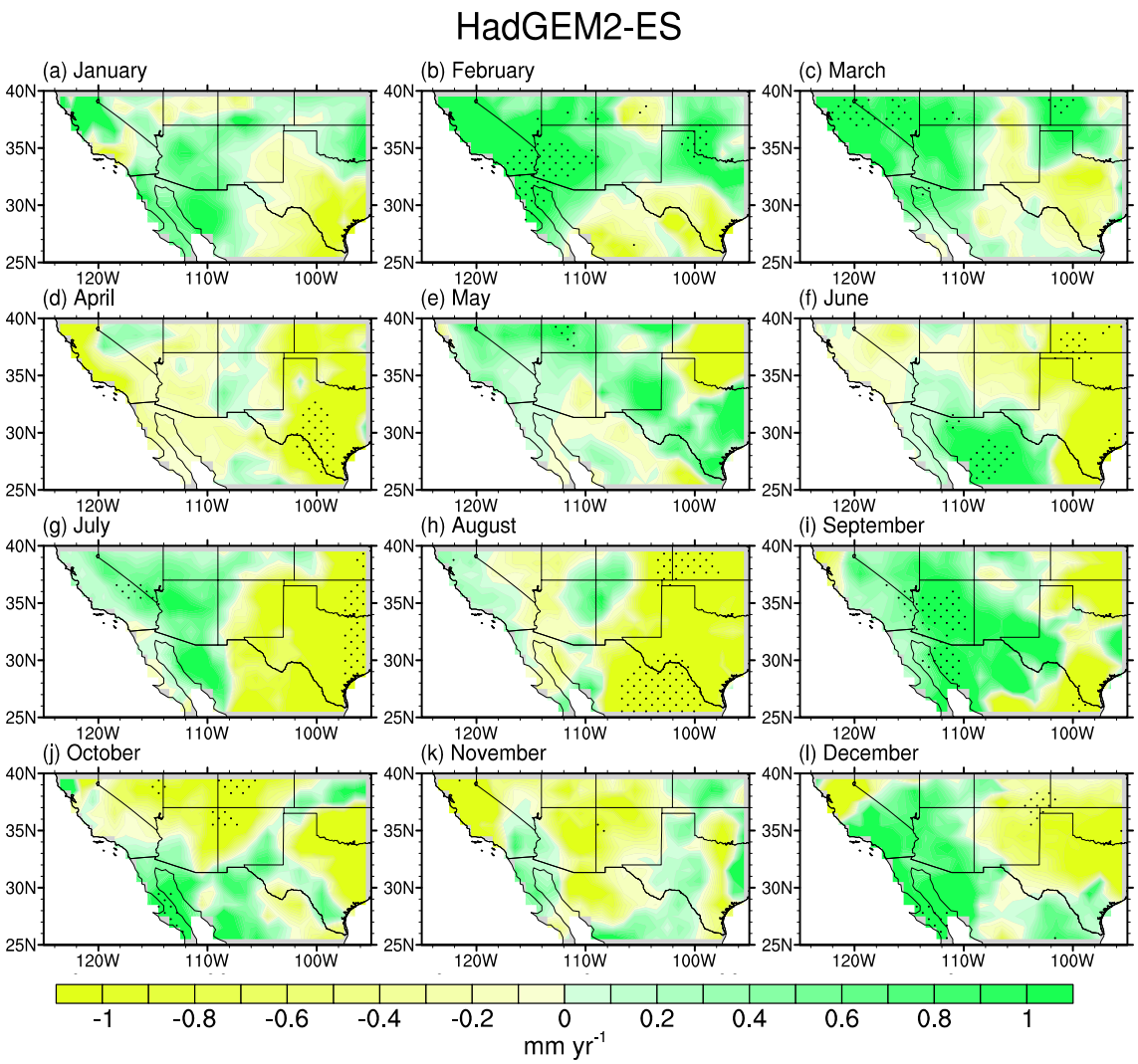


Figure A2-67. Monthly precipitation trends for HadGEM2-ES. Stippled regions indicate statistically significance (90% level).

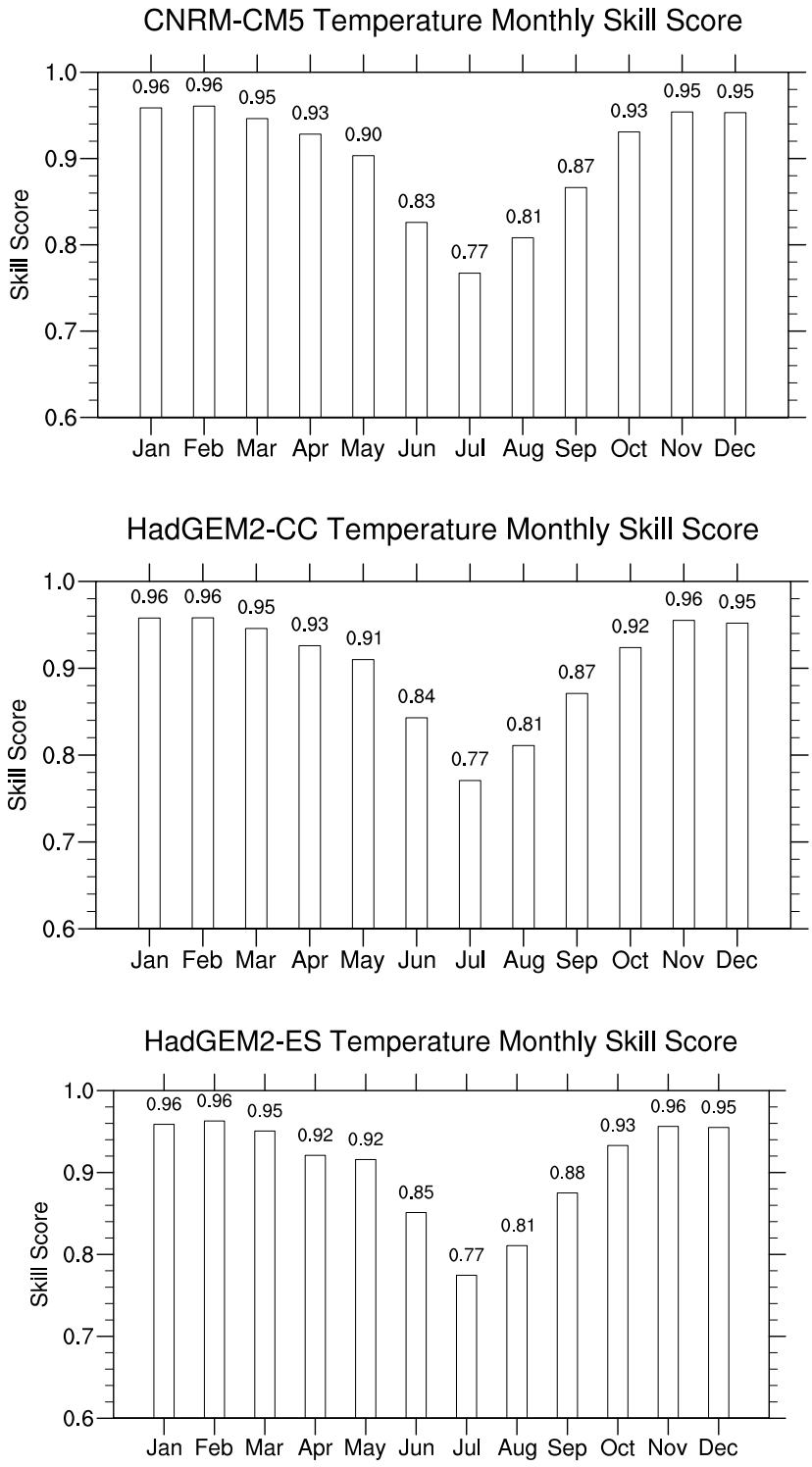


Figure A2-68. Monthly skill scores for temperature for the downscaled CNRM-CM5, HadGEM2-CC, and HadGEM2-ES.

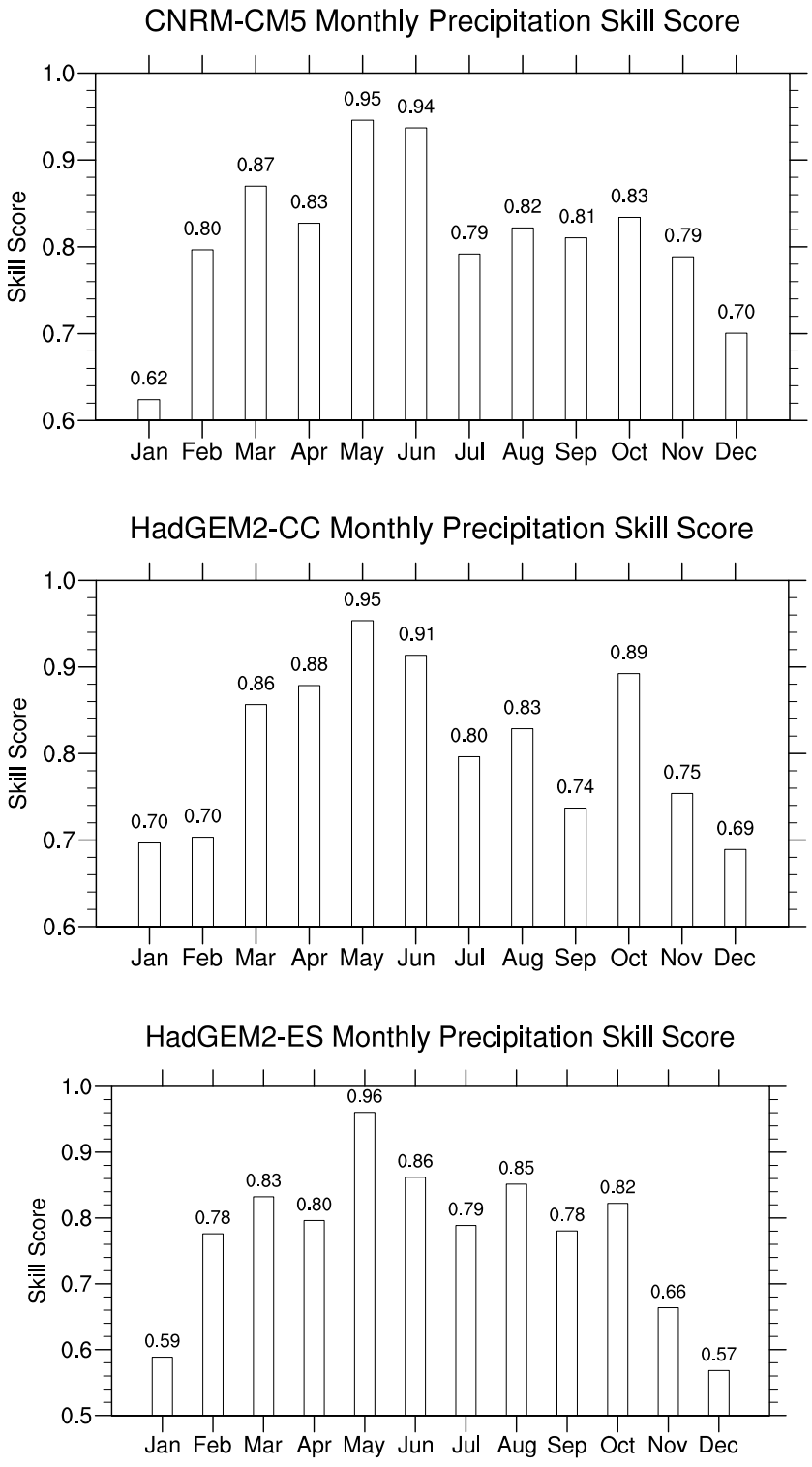


Figure A2-69. Monthly skill scores for precipitation for the downscaled CNRM-CM5, HadGEM2-CC, and HadGEM2-ES.