Towards Assessing NARCCAP Regional Climate Model Credibility for the North American Monsoon: Current Climate Simulations*

MELISSA S. BUKOVSKY, DAVID J. GOCHIS, AND LINDA O. MEARNS
National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 31 July 2012, in final form 1 May 2013)

ABSTRACT

The authors examine 17 dynamically downscaled simulations produced as part of the North American Regional Climate Change Assessment Program (NARCCAP) for their skill in reproducing the North American monsoon system. The focus is on precipitation and the drivers behind the precipitation biases seen in the simulations of the current climate. Thus, a process-based approach to the question of model fidelity is taken in order to help assess confidence in this suite of simulations.

The results show that the regional climate models (RCMs) forced with a reanalysis product and atmosphere-only global climate model (AGCM) time-slice simulations perform reasonably well over the core Mexican and southwest United States regions. Some of the dynamically downscaled simulations do, however, have strong dry biases in Arizona that are related to their inability to develop credible monsoon flow structure over the Gulf of California. When forced with different atmosphere–ocean coupled global climate models (AOGCMs) for the current period, the skill of the RCMs subdivides largely by the skill of the forcing or “parent” AOGCM. How the inherited biases affect the RCM simulations is investigated. While it is clear that the AOGCMs have a large influence on the RCMs, the authors also demonstrate where the regional models add value to the simulations and discuss the differential credibility of the six RCMs (17 total simulations), two AGCM time slices, and four AOGCMs examined herein. It is found that in-depth analysis of parent GCM and RCM scenarios can identify a meaningful subset of models that can produce credible simulations of the North American monsoon precipitation.

1. Introduction

Many climate models are challenged by the complexity of the North American monsoon (NAM) system (Gutzler et al. 2005). While several existing studies that employ RCMs (e.g., Anderson and Roads 2002; Saleeb and Cotton 2004; Xu et al. 2004; Castro et al. 2007a,b; Cerezo-Mota et al. 2011; Chan and Misra 2011; Gao et al. 2012) have individually demonstrated simulation fidelity, a more systematic analysis of multiple models running under a common comparison framework has been lacking. In this study, we take advantage of the comprehensive set of dynamically downscaled climate simulations produced as a part of the North American Regional Climate Change Assessment Program (NARCCAP). NARCCAP is a major multimodel assessment effort aimed at understanding regional climate change in North America and its uncertainties (Mearns et al. 2009, 2012). In NARCCAP, six regional climate models (RCMs) are used to downscale four different coupled atmosphere–ocean global climate models (AOGCM, hereafter GCM), with 12 statistically planned combinations, and a reanalysis product. Two atmosphere-only global climate model (AGCM) time-slice simulations have also been completed.

Climatologically, NAM system-related precipitation starts in late May–early June in southern Mexico (MX), migrates northward to the southwestern United States by early July, and ends approximately two months later (Higgins et al. 1999). About 40% of the annual total precipitation in the southwest United States and about 70% of it in northwest Mexico falls during the monsoon season (Douglas et al. 1993; Higgins et al. 1997). The monsoon system itself is characterized by a seasonal

* Supplemental information related to this paper is available at the Journals Online website: http://dx.doi.org/10.1175/JCLI-D-12-00538.s1.

Corresponding author address: Melissa S. Bukovsky, NCAR/IMAGE, P.O. Box 3000, Boulder, CO 80307.
E-mail: bukovsky@ucar.edu

DOI: 10.1175/JCLI-D-12-00538.1

© 2013 American Meteorological Society
shift in the North American subtropical high and regional wind patterns (e.g., Adams and Comrie 1997; Johnson et al. 2007). Near-surface onshore flow into northwest Mexico and Arizona (AZ) and the northward low-level jet (LLJ) along the Gulf of California (GOC) during this season are important in transporting the moisture necessary for precipitation (Douglas 1995; Douglas et al. 1998). These features and much of the area’s topography, in general, are not well resolved by most GCMs (Mitchell et al. 2002; Collier and Zhang 2007; Lee et al. 2007). However, higher-resolution models do not always produce a competent simulation in the area either (Gutzler et al. 2009). Given that this is a semiarid region with high vulnerability to both droughts and episodic flooding, proficient simulation of the monsoon system for prediction and projection purposes is important.

This study therefore focuses on characterizing model skill in the current or baseline climate period in the NARCCAP regional model simulations as it relates to precipitation and the causal mechanisms behind identified precipitation biases. In this way, we are working to assess the credibility of the NARCCAP simulations through a process-based analysis approach. Models, data, and the evaluation methodology are described in the next section. The RCM simulations forced by a reanalysis product are analyzed in section 3, and simulations forced by GCMs in a current climate scenario are scrutinized in section 4. A discussion and summary of findings are provided in sections 5 and 6.

2. Models, methods, and datasets

a. Regional and global models

Six RCMs1 and two AGCMs (or time slices) from NARCCAP are used for this analysis (Mearns et al. 2007). They are as follows:

• Canadian RCM (CRCM; Caya and Laprise 1999),
• Scripps Experimental Climate Prediction Center version of the regional spectral model (ECP2; Juang et al. 1997),
• Third-generation Hadley Centre RCM (HRM3; Jones et al. 2003),
• Fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model as run by the Iowa State University modeling team (MM5I; Grell et al. 1993),
• International Centre for Theoretical Physics RCM, version 3 (RCM3; Giorgi et al. 1993a,b; Pal et al. 2007),
• Weather Research and Forecasting model as run by the Pacific Northwest National Laboratory (PNNL) modeling team (WRFG; Skamarock et al. 2005),
• Atmospheric component of the Geophysical Fluid Dynamics Laboratory global climate model, version 2.1 (GFDL; Anderson et al. 2004), and
• Atmospheric component of the National Center for Atmospheric Research (NCAR) Community Climate System Model (CCSM), version 3 (CAM; Govindasamy et al. 2003).

All RCMs and AGCMs have a horizontal resolution of 50 km. The CRCM and ECP2 are the only two models that include some form of interior nudging, a “relaxation” toward the large-scale driving conditions within the interior of the domain, which constrains these regional models to more closely follow their driving reanalysis or GCM. The impact of the nudging is illustrated in Mearns et al. (2012), where these two models more closely reproduce observed temperature and precipitation variability and average magnitude in the regions selected for analysis when forced with a reanalysis product. Further details on NARCCAP and the configuration of the models can be found in Mearns et al. (2012), online (www.narccap.ucar.edu), and/or in the listed references. However, as they are of greatest interest given our analyses, the convective parameterization and number of vertical levels used in each RCM is listed in Table 1.

We examine two sets of simulations using these RCMs. The first set acquires initial and boundary conditions from the National Centers for Environmental Prediction (NCEP)/Department of Energy (DOE) Reanalysis II (hereafter NCEP; Kanamitsu et al. 2002). These simulations are for the period of 1980–2004. The second set is composed of simulations forced by four GCMs for the period of 1971–2000 (or 1999, depending on the GCM). Three out of the four are simulations that are

| Table 1. Convective parameterization scheme and number of vertical levels used in each NARCCAP RCM. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| RCM             | Levels          | Convective parameterization scheme                             |
| CRCM            | 29              | Bechtold–Kain–Fritsch (Bechtold et al. 2001; Kain and Frisch 1990) |
| ECP2            | 28              | Simplified Arakawa–Schubert (Pan and Wu 1995)                  |
| HRM3            | 19              | Mass flux with downdraft and momentum transport (Gregory and Rowntree 1990; Gregory and Allen 1991; Gregory et al. 1997) |
| MM5I            | 23              | Kain–Fritsch 2 (Kain 2004)                                     |
| RCM3            | 18              | Grell with Fritsch–Chappell closure (Grell 1993; Fritsch and Chappell 1980) |
| WRFG            | 35              | Grell–Devenyi (Grell and Devenyi 2002)                         |

1 All RCM acronyms are as used in the NARCCAP model archive.

NOVEMBER 2013 BUKOVSKY ET AL. 8803
included in the World Climate Research Programme (WCRP)’s Coupled Model Intercomparison Project phase 3 (CMIP3) archive. The GCMs, their horizontal resolution, and their ensemble member number, where applicable, are

- NCAR CCSM, version 3.0 (Collins et al. 2006), at T85 (1.4° × 1.4°) with 5 members;
- Canadian Global Climate Model, version 3 (CGCM; Flato et al. 2000), at T47 (1.9° × 1.9°) with 4 members;
- GFDL Climate Model, version 2.1 (GFDL; Anderson et al. 2004), at 2.0° × 2.5° with 2 members; and
- Hadley Centre Climate Model, version 3Q0 (HadCM3Q0, hereafter HADCM; Gordon et al. 2000; Pope et al. 2000), at 2.5° × 3.75° (not included in the CMIP3 archive). This version differs from the version used to perform the CMIP3 simulations in that it includes a flux adjustment and an aerosol cycle.

The two different AGCMs are based on the atmospheric components of the CCSM and GFDL GCMs and are run using lower boundary conditions derived from observed SSTs and sea ice.

Simulations were completed with all six RCMs using NCEP for forcing. Of the possible 24 RCM-GCM pairings, 12 are being completed as part of NARCCAP. Each RCM uses boundary conditions from two different GCMs and each GCM drives three RCMS, with combinations chosen using a balanced fractional factorial design. The GCM-forced simulations used in this analysis are summarized in Table 2. When in combination, an RCM and its driver will be referred to as, for example, RCM3-gfdl, with the forcing simulation or reanalysis in lower case. When not in combination, all acronyms will be in standard upper case.

<table>
<thead>
<tr>
<th></th>
<th>CCSM</th>
<th>CGCM</th>
<th>GFDL</th>
<th>HADCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRCM</td>
<td>O</td>
<td>O</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ECP2</td>
<td>X</td>
<td>—</td>
<td>—</td>
<td>O</td>
</tr>
<tr>
<td>HRM3</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>O</td>
</tr>
<tr>
<td>MM5I</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>RCM3</td>
<td>—</td>
<td>O</td>
<td>O</td>
<td>—</td>
</tr>
<tr>
<td>WRFG</td>
<td>O</td>
<td>O</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

terrain from one of the RCMS (terrain does vary slightly between the RCMS, but one is fairly representative of the others), important topographic features in the region, and locations for the cross sections shown later. The increase in resolution to 50 km from the resolution of any of the parent GCMs or NCEP has a substantial influence on the resolved topography in this region. For comparison, the surface elevation fields from two of the parent GCMs (one finer and one coarser) are shown in Figs. 1b and 1c.

Model simulation data used herein include what is available at the time of writing. Omissions of specific simulations from various figures often result because the necessary variables are not yet available for analysis. Also, because of missing data/times near the end of simulations, all GCM-driven RCMS and AGCM time slices are analyzed from 1971 to 1999 for consistency across analyses while trying to use as much available data as possible. For June–September (JJAS) seasonal analysis using the NCEP-driven RCMS and NARR, the analysis period is always 1980–2004, and for analysis that spans the entire year, the period is 1980–2003 (as the end of 2004 is not available in all simulations).

### c. Verification datasets

Several observationally based datasets will be used for model evaluation because no single dataset is comprehensive in terms of our needs and because this provides a simple view of the variability/uncertainty between datasets. In this analysis, we use datasets as follows:

- North American Regional Reanalysis (NARR; Mesinger et al. 2006) at 32-km horizontal resolution and 3-h temporal resolution with data from 1980 onward. [Unlike most reanalyses, precipitation is essentially assimilated where observations are available, making these data particularly useful over the contiguous United States; however, problems do exist elsewhere as noted in Bukovsky and Karoly (2007) and Ruane (2010a,b).]
- NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) at ½° latitude/longitude resolution and 6-h temporal resolution, globally available from 1979 onward. [Relatively few analyses of the CFSR have been conducted so far, but it is included here for comparison with the NARR, which has known problems with winds in the NAM region (documented below).]
- University of Delaware analysis, version 2.01 (UDEL; Willmott and Matsuura 1995), at ½° latitude/longitude resolution with gridded observationally based land-only monthly average precipitation and temperature from 1900 to 2008.

### Table 2. NARCCAP GCM-forced simulations. Planned combinations are marked with either O or X. Those used for this work are indicated with an ‘‘O’’. Those not yet available are indicated with an ‘‘X’’. Note that the simulations with the ECP2 may be incomplete. The death in June 2008 of the project co-investigator responsible for all of the ECP2 simulations greatly affected all of the work with this model.

<table>
<thead>
<tr>
<th></th>
<th>CCSM</th>
<th>CGCM</th>
<th>GFDL</th>
<th>HADCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRCM</td>
<td>O</td>
<td>O</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ECP2</td>
<td>X</td>
<td>—</td>
<td>—</td>
<td>O</td>
</tr>
<tr>
<td>HRM3</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>O</td>
</tr>
<tr>
<td>MM5I</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>RCM3</td>
<td>—</td>
<td>O</td>
<td>O</td>
<td>—</td>
</tr>
<tr>
<td>WRFG</td>
<td>O</td>
<td>O</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
Colorado State University North American Monsoon Experiment upper-air and surface gridded analyses (NAME; Ciesielski and Johnson 2008) at 1° latitude/longitude horizontal resolution and hourly analysis of observations collected 1 July–15 August 2004 during the North American Monsoon Experiment enhanced observing period.

- Tropical Rainfall Measuring Mission satellite-derived precipitation (TRMM; Huffman et al. 1995, 1997) at 0.25° latitude/longitude horizontal resolution available every 3 h, starting in December 1997.
- National Oceanic and Atmospheric Administration (NOAA)’s Climate Prediction Center (CPC) daily United States and Mexico precipitation analysis (US-MEX; Higgins et al. 1996) at 1° latitude/longitude horizontal resolution over land only from January 1948 to April 1999 that includes about 300 gauge stations in Mexico prior to 1990 and about 600 after 1990.
- NOAA high-resolution blended analysis of daily sea surface temperatures (SSTs) and ice cover (Reynolds et al. 2007) at 0.25° latitude/longitude horizontal resolution generated from satellite data and in situ observations with data available after September 1981.

3. NCEP-driven simulations

Monsoon season (JJAS) average precipitation and precipitation bias from the NCEP-driven RCMs are shown in Figs. 2 and 3, respectively. There is a widespread average dry bias over most of this region in most of the RCMs. The CRCM is the outlier in that it is approximately 72% wetter than UDEL in the common land area shown in Fig. 2. The spatial distribution of precipitation along the Sierra Madre Occidental, the Mogollon Rim, and in New Mexico is well represented by the models despite the bias. This behavior is important, as the correlation of precipitation to prominent terrain features is an important aspect of the NAM precipitation climatology. The spatial distribution in the CRCM, given the noise in the field and the high bias, does not have as clean of a spatial distribution as the other RCMs. Four precipitation estimates are given in Fig. 2 for verification, as each has drawbacks. UDEL is only available over land and as monthly averages, NARR has an incorrect minimum in precipitation along country boundaries (Bukovsky and Karoly 2007; Ruane 2010a). TRMM is only available during the last seven years of.
the simulations, and US-MEX is only available through 1998 and is coarser in resolution than the RCMs.

Seasonal average fields of precipitation, however, do not present enough information to determine whether or not the RCMs are simulating precipitation as part of a credible monsoon system. That is, a low or high bias is a necessary but not sufficient condition to determine model failure and/or a lack of useful information. Likewise, knowing more about the potential causes of the bias may give us useful information that will help us determine our confidence in the RCMs. As such, additional analyses proceed along those lines.

Given the generally reasonable quality of the NCEP boundary conditions, the RCMs should be able to produce a clear signal for the monsoon season precipitation within the annual precipitation cycle. For the most part they do, as illustrated in Fig. 4. To justify the use of NARR precipitation in Fig. 4 (the only dataset that spans the entire period, covers the entire region, and has submonthly information), we offer Fig. 5: the verification of the NARR against other precipitation datasets that span shorter periods. Together, Figs. 4 and 5 show that spread in model performance in any region is larger than that in the various precipitation datasets, as discussed next.

Over the entire core region (Fig. 4a), the RCMs all produce a clear seasonal signal for the monsoon precipitation albeit with varying degrees of model bias, as supported by Fig. 3. Within the subregions, however, performance is mixed as the RCMs capture the Mexican subregion of the monsoon much better than they do the Arizona subregion. Over the Mexican subregion (Fig. 4c), the models all show a clear monsoon signal, and with one exception the magnitude and time of onset is well captured. The average time of monsoon onset for this region is within about 5 days of what is indicated by the NARR. The exception is the RCM3, which rains too early and too heavily over the Sierra Madre Occidental.
Over AZ (Fig. 4b), the CRCM high precipitation bias again stands out during monsoon season and also during the preceding dry season; however, the time of monsoon onset is reasonably captured. The dry bias in all other RCMs is strong in AZ. In the MM5I, it is so intense that there is no monsoon signal in the precipitation climatology. All other RCMs contain some relative increase in precipitation near monsoon onset, relative to the well-captured spring dry season, that lasts through monsoon season.

The 2-m specific humidity (not shown) in all of the RCMs ranges from 0.7 to 2.4 g kg\(^{-1}\) or from 8% to 31% drier than the NARR (from the RCM3 to the MM5I), except the CRCM, which is about 6.4 g kg\(^{-1}\) or 90% more moist (not shown). This problem extends above the surface for all models except the CRCM, as illustrated in a seasonal average profile for the point nearest Phoenix, Arizona (Fig. 6). The seasonal average dry bias extends up to at least 700 hPa. This not only affects the amount of moisture available for convection but also the amount of moist instability, or moist static energy, on average. One may also notice the HRM3-ncep warm bias (Fig. 6). It exists over almost all of North America (Mearns et al. 2012), but it is weaker in the NAM region than it is elsewhere. The CRCM has an as yet unexplained high bias in specific humidity at 2-m in the Southwest only (i.e., this is not a problem in the rest of North America), and the bias does not appear to exist aloft. All CRCM simulations and the CGCM simulation have the same problem.\(^2\) This excessive near-surface moisture in the CRCM is a likely contributor to its excessive rainfall. Not only is there more moisture available from the lowest level (if not aloft), but

\(^2\)Note that the CRCM and the CGCM are similar models. They use different parameterizations for convection and some modifications are made to account for the different resolutions. Additionally, there may be differences in the order the parameterizations are called, and of course one is a limited area model and the other is a global model.
FIG. 4. The 1980–2003 5-day average precipitation climatology for (a) the core NAM region, (b) AZ, and (c) MX.
also a surface-based air parcel in this environment is much more unstable in the CRCM, on average, than in any other RCM (not shown but can be inferred from Fig. 6).

To some extent the low-level dry bias in several RCMs, particularly in AZ, reflects the inability of the RCMs to develop low-level onshore flow and the GOC LLJ. Low-level flow is exhibited through the near-surface moisture flux field in Fig. 7. A vertical cross section of specific humidity and temperature along the GOC showing flow tangent to the cross section is given in Fig. 8. Comparisons here are given against the NARR and the CFSR; however, the NARR does have northward flow over the GOC during this season that is too strong, as noted in Mo et al. (2005). In an attempt to quantify the bias in NARR and CFSR (as the latter has not been as well documented), Figs. S1 and S2 in the supplementary material compare NARR and CFSR to NAME for July 2004. NARR certainly has a stronger northward flow component along the GOC at low levels compared to NAME observations during this month, and this understandably affects the calculated moisture flux field (Figs. 7g,j) but surprisingly not the Phoenix moisture profile in Fig. S2. Thus, when comparing the RCMs to NARR, we do so qualitatively, looking for correct direction in the RCMs more so than an absolute magnitude that matches NARR’s. Winds in the CFSR may also be similarly biased at low levels, given the comparison between Figs. 7i and 7k but clearly not to the same extent as NARR and not through as great of a depth (Fig. S1). Nevertheless, it is clear that some of the RCMs, particularly ECP2 and MM5I, do not contain an average low-level northward flow component along the GOC that is strong enough during JJAS (although a weak sea breeze does develop in the afternoon in the RCMs; not shown). Obtaining an onshore flow component into Mexico is much less

![Fig. 5. Comparison of 5-day average precipitation climatologies between (a) NARR and US-MEX (USMX) for 1980–98 for AZ and MX and (b) NARR and TRMM for 1998–2004 for the core NAM region, AZ, and MX.](image-url)
problematic for the RCMs (Fig. 7), and it is here that the RCMs perform better.

Midlevel moisture for the monsoon core region is also sourced from the Caribbean basin and the Gulf of Mexico (GOM) via the Caribbean and Great Plains low-level jets (CLLJ and GPLLJ, respectively; e.g., Mo et al. 2005; Schiffer and Nesbitt 2012). Cerezo-Mota et al. (2011) state that NCEP is drier than NARR and the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) over the GOM (for JJAS 1980–2004 at 850 hPa, by around 1 g kg$^{-1}$; light dashed green), and isobars (hPa, horizontal light purple lines). For the RCMs, the temperature profile is drawn as the solid line (colors indicated in the key), and the dewpoint temperature is the dashed line. NARR (CFSR) temperature and dewpoint profiles are both solid blue (orange) lines with dewpoint to the left.

4. AOGCM-driven simulations and time slices

a. Precipitation

Figure 9 shows precipitation from the GCM-driven simulations and time slices. Simulations of average
precipitation from any given RCM now are quite different from the corresponding NCEP-forced simulation. For example, while the WRFG-ncep performed fairly well in Fig. 2, here it is much drier whether driven by the CCSM or CGCM. Similarly, the CRCM-ccsm and CRCM-cgcm simulations are also much drier. The opposite is true of all of the RCMs forced by the GFDL. The AGCM time slices are also included in Fig. 9. Both have problems developing precipitation in AZ, similar to the RCMs when forced with NCEP. Also, the CAM time slice does not capture the spatial pattern of the precipitation very well, and its spatial distribution is somewhat different than what is observed. For reference, precipitation from the forcing GCMs is also included in Fig. 10 over a much expanded area. As expected, given that they do not represent the region’s topography as well, their average precipitation field is much smoother overall, and they do not simulate the precipitation in the southwest United States well. They all have the southeast–northwest-oriented rain belt along the western coast of Mexico, though with large variations in magnitude.

Precipitation annual cycle climatologies from these simulations are shown in Fig. 11 (full core region), Fig. 12 (AZ), and Fig. 13 (MX). The panels in each are subdivided by forcing GCM, as the skill of the RCMs in this measure is determined by the associated GCM. Proficiency in capturing the monsoon signal in the precipitation is highly variable now, relative to the skill when forced by NCEP. When forced by the CGCM, the CRCM captures the monsoon precipitation signal best. In the RCM3-cgcm, the signal is weak to absent, depending on the subregion, and the simulation of monsoon precipitation in the WRFG-cgcm is poor across the
board, a clear switch from its skill when driven by NCEP. The CCSM-forced RCMs fail to produce a monsoon precipitation signature. The CRCM-ccsm has the least bias, but it is still inadequate. On the other hand, the CAM time slice is reasonably skillful, particularly in Mexico, where the timing and magnitude of precipitation are considerably better than in most of the other simulations in JJAS. As with the NCEP-driven RCMs, the CAM time slice also has a dry bias in AZ. (The CAM time-slice high bias in the core-region average later in the JJAS season is largely due to excessive precipitation over the GOC; not shown). Both the CCSM- and CGCM-forced simulations produce an annual cycle over the full core region that is in contrast to
the climatology produced by the GCMs themselves, as shown in Fig. 14. The CGCM produces a very reasonable precipitation annual cycle during this period, not exhibiting problems with timing or magnitude, and the CCSM, while wet in the monsoon season and in the cool season, at least has a monsoon precipitation signal, unlike the RCMs it forces. The GFDL-driven RCMs may contain a monsoon signal in Fig. 11, but it is masked by the highly excessive precipitation in the latter part of the year. This is a problem in the core-region area average more so than in the subregions because the excessive precipitation is produced mainly over water. The same is also true in the GFDL GCM, as is clear in Fig. 14. In AZ, it appears that the HRM3-gfdl may be capturing the monsoon precipitation signature, and the same might be said of all of the RCMs in Mexico as well, but there is clearly something wrong that starts near the end of the monsoon season. The GFDL time slice does not have

Fig. 9. The 1971–99 JJAS average precipitation from the GCM-driven RCMs, the AGCM time slices, and UDEL.
this problem, unlike the RCMs and the GFDL GCM, but it does have a strong low bias in AZ and a high bias in Mexico. As with the CAM time slice, the GFDL time slice also does not contain the cool season wet bias seen in the corresponding RCM simulations. The two time-slice simulations show more skill in simulating the annual precipitation climatology than almost all of the GCM-driven RCMs, the GFDL time slice in AZ during the monsoon season being the exception. Similarly, the HADCM-forced simulations have more skill in simulating precipitation during the monsoon season than most of the other RCM–GCM combinations. The season is shifted (i.e., it starts early), similar to what is seen in the HADCM, but the magnitude of the precipitation is improved relative to most of the others. Problems in the timing and magnitude of monsoon precipitation in the GCMs utilized in NARCCAP are fairly typical of the problems in many other CMIP3 GCMs. As shown in Lin et al. (2008), while most simulate monsoon rainfall, there are problems with early onset, a high bias at the end of the monsoon season through the end of the year and/or late retreat, and a high precipitation bias during the NAM season. Three of the four NARCCAP-chosen GCMs coincidentally sample all of these problems.

b. Diagnostics discussion

Since the skill of the RCMs subdivides by the skill of the parent GCM, it is clear that a substantial amount of the precipitation bias in these simulations results from problems in the GCM boundary conditions and is not completely caused by the RCMs.

In the CGCM- and CCSM-driven simulations, moisture supply appears to be the likely main issue leading to the overall lack of monsoon system precipitation. Atmospheric low-level moisture contains a strong dry bias in the CGCM and CCSM, as illustrated in Fig. 15 in the 850-hPa specific humidity field, off the west coast of California and the Baja Peninsula, south of the Baja Peninsula, and over the GOM, as compared with NCEP. As the boundaries of the RCMs cut through these regions, the RCMs inherit these biases and then exacerbate them, given their inherent bias.

Using the WRFG and the CRCM as examples, as they are both forced by these two GCMs, various problems with the CGCM and the CCSM become clear. Both the WRFG and CRCM develop onshore low-level monsoon flow into Mexico and AZ, unlike some of the other RCMs when forced with NCEP that do not produce northerly flow in the northern GOC, as illustrated in Fig. 7. However, their near-surface moisture flux (Fig. 16) and representation of the GOC LLJ (Fig. 17) is weaker in the CGCM- and CGCM-driven simulations and more so in the latter. Their flow remains representative of
their performance when forced with NCEP (as it does in all of the RCMs, so they still have a better representation of the low-level flow over the GOC than several of the other RCMS; not shown), but it is affected by the GCM bias. This implies not only less moisture transport (and less low-level moisture, as is illustrated in Fig. S6 of the supplementary material), but also less forcing for precipitation as orographic lift decreases.

In the CGCM-driven simulations, problems with flow in and around the GOC and with the low-level humidity are worse than they are in the CCSM-driven simulations (Figs. 17 and S6). Not only does the CGCM have a stronger humidity bias than the CCSM (Fig. 15), which the RCMs inherit, but that problem is also exacerbated by a problem with GOC SSTs. As illustrated in Fig. 18, the CGCM-forced simulations have a cold SST bias in the GOC, which would inhibit the flux of moisture from the water surface and increase thermodynamic stability. While this could also compel a stronger land–sea contrast/land–sea breeze, it does not because there is also a cold bias over land in both the CRCM-cgcm and the WRFG-cgcm that is worse than what is seen in the NCEP-driven simulations, particularly over the western California and Sonoran deserts (Fig. S7 of the supplementary material and illustrated for Phoenix in Fig. 19). This 2-m temperature cold bias is also reflected in lower heights associated with the subtropical ridge in these RCMs (not shown). Interestingly, neither bias is seen in the CGCM, where the strength and placement of the subtropical high at 500 hPa is very close to that in the NARR, and the near-surface temperature over land in the Southwest is not colder than that seen in NCEP (not shown). Thus, the exact cause of this cold bias in the RCMs parented by the CGCM is as yet unknown. The same problem with the GOC SSTs does not exist in the CCSM-driven simulations, but the WRFG-cscm also has a tendency to be colder than the CCSM and the NCEP-driven simulation in the Southwest over land (Fig. S7d). Near Phoenix, these problems combined lead to an even greater deficit in low-level atmospheric moisture than

**Fig. 11.** The 1971–99 5-day average precipitation climatology for the core NAM region from the GCM-driven RCMs and AGCM time slices. Each panel corresponds to a GCM with NARR (black lines) as in Fig. 4. Note that the GFDL-driven RCMs panel is on different vertical scale in this figure only.
seen in the NCEP-driven simulations (cf. Fig. 6 and Fig. 19; GCM JJAS average soundings are also shown in Fig. 19 for reference). The same is true for the Mexican region, as can be seen in Fig. S6, though here the lack of rainfall is more noteworthy because the Mexican NAM rainfall component is clearly a subregion where the RCMs are likely to have greater skill, as indicated in the NCEP-driven simulations. These biases help explain why the CGCM-forced simulations are even less humid than the CCSM-forced simulations and why both have problems with precipitation in this region while their GCM counterparts do not to the same extent.

In the CRCM-cgcm simulation, monsoon rainfall is still captured in the annual cycle climatology, particularly in Mexico, though not in the other CGCM-driven simulations. This may be due to the similarities between the two models. The CRCM may be better able to overcome some of the biases it is receiving from the CGCM because they are so similar and therefore still be able to generate precipitation during the monsoon season. That is at least one hypothesis for its more skilful performance over the other CGCM-driven simulations. It is also possible that the high near-surface moisture bias in the CRCM is leading to the somewhat smaller negative precipitation bias in this simulation. The high near-surface moisture does lead to substantially more moist instability in all CRCM simulations as well (not shown). As atmospheric moisture content appears to be responsible for the low precipitation bias, and the cause is not a more fatal flaw in the large-scale setup, this could be an example of a beneficial or compensating bias. This high near-surface moisture bias also exists in the CCSM-forced simulation as well, but the CRCM-ccsm does not show as much of an improvement over the other RCMs here as it does when forced by the CGCM, so it could also be a combination of the two factors.

The most obvious feature in the GFDL-forced simulations is the excessive precipitation that starts toward the end of the monsoon season. Most of this precipitation falls over water, which is why its signal is strongest in the full core region average. This excessive precipitation looks to be forced by an anomalous long-term average low-pressure center that forms in the GFDL GCM just southwest of the Baja Peninsula (Fig. 20),
likely as a result of excessive precipitation in the GFDL GCM. Extra cyclonic activity here forces flow that is more southerly than should exist well past monsoon season. Not only does this mean greater moisture transport into the monsoon region, but it also forces extra rising motion in our analysis region (particularly the southern end of it) through December, although (as suggested by NCEP) monthly average vertical velocities around 600–500 hPa should switch sign or decrease in intensity much earlier in the year (not shown). As the source of the problem lies on/near the southern boundary of the RCMs, it is clearly translating into the RCMs.

Excessive rainfall in the GFDL-driven simulations, particularly that over the included water bodies (both in

---

**FIG. 13.** As in Fig. 11, but for the Mexico subregion.

**FIG. 14.** The 1971–99 precipitation climatology for the core NAM region from the four parent GCMs. The 5-day average climatology from the CCSM, CGCM, and GFDL as in Fig. 11. Monthly average climatology from the HADCM also shown (finer temporal scale precipitation output not available from this simulation) with NARR (black line) as in Fig. 4.
the GOC and near the coastal Pacific Ocean), is likely exacerbated by the high GFDL GCM SSTs during and after the monsoon season (Fig. 18). The warm SST bias on the west coast of the Americas (and Africa) in GCMs is a well-known problem (Gordon et al. 2000; Washington et al. 2000; Collins et al. 2006; Jungclaus et al. 2006; Christensen et al. 2007) and does translate into the RCMs, as they use those SSTs as a 6-hourly updating lower boundary condition. Where there are no water points to interpolate from, the NARCCAP RCMs interpolate SSTs from the nearest water bodies (i.e., the GOC SSTs in the RCMs are interpolated from the Pacific Ocean GCM SSTs). The SSTs in the GOC in some of the RCMs are more realistic because this GCM SST bias exists; however, it is still a warm bias in the RCMs over the adjacent Pacific Ocean (except in the CGCM-driven simulations, where the Pacific SSTs are more reasonable, causing a low SST bias in the GOC). Even with these problems, the GFDL-driven simulations do capture some aspects of the monsoon rainfall, as is clear in the Mexico precipitation climatology (Fig. 13, third panel from top). In AZ, all GFDL-driven simulations except the HRM3-gfdl fail to show a good signal for the monsoon in their precipitation. Why the HRM3-gfdl does and the other GFDL-forced RCMs do not remains to be investigated; however, it may just be because the HRM3 consistently develops near-surface northward flow over the GOC and a GOC LLJ that is stronger than that produced by the RCM3 and the ECP2, regardless of the forcing, which would help produce more precipitation in AZ and give a better signal in the climatology (cf. 850-hPa GOC flow in Figs. S4 and S5; near-surface flux not shown).

The HADCM-forced RCMs do not suffer from the same problems as described above (analysis not always shown for all models in previous figures to save space). This GCM clearly provides boundary conditions with biases that are less traumatic to the RCMs it forces. Problems are similar to those in the NCEP-driven simulations, and the precipitation climatology in AZ is even improved in both simulations (potentially because of the increased moisture over the GOM and just south of the Baja Peninsula relative to NCEP, as shown in Fig. 15). Early onset is the main problem in these simulations (we do not consider the biases in magnitude a problem relative to what we are seeing in the other simulations). This behavior is also present in the HADCM and is likely due to the HADCM migrating the monsoon band northward too quickly. Interestingly, this is not seen in the CMIP3 HADCM simulation analyzed in Lin et al. (2008).

In this analysis, we have not included the AGCM time slices as they are not impacted by inherited GCM bias as the RCMs are. As the time slices are forced with observed SSTs during this baseline period, their problems are much more similar in nature to those seen in the NCEP-driven RCMs. Variations in performance between the two time slices are due to differences in model

![Figure 15. The 1971–99 JJAS average 850-hPa specific humidity from the (a) CCSM, (b) CGCM, (c) HadCM3, (d) GFDL, and (e) NCEP. Note that in NCEP, below ground values are filled via interpolation.](image-url)
configuration, as they are forced with the same “perfect” lower boundary conditions.

We have also not included an examination of intraseasonal and interannual monsoon variability in this article. As the RCMs correlate fairly well in time with their drivers (not shown), this is largely an inherited trait. Fortunately, as it pertains to the NARCCAP parent GCMs, intraseasonal variability has already been explored in the existing literature on the CMIP3 archive. Likewise, as shown in Lin et al. (2008), most of the CMIP3 GCMs produce the monsoon-related precipitation belt, suggesting that they develop the basic processes responsible for the NAM, and their analysis includes the GCMs used in NARCCAP. The same study examines tropical easterly waves and the Madden–Julian oscillation (MJO) in these GCMs and thus their representation of intraseasonal variability (it was found that most GCMs have a good representation of easterly waves but a poor representation of variance and the MJO). As these issues will become more important when

![Figure 16](image-url)

**FIG. 16.** The 1971–99 JJAS average near-surface moisture flux from the (a),(c) CRCM and (b),(d) WRFG with CCSM- and CGCM-driven simulations.
examining the climate change projections from these simulations, we will save discussion and analysis on the impacts of interannual and intraseasonal variability bias for future work.

5. Discussion

The experimental design of NARCCAP allows us to suggest additional simulations that may be useful to complete in the future (e.g., CRCM and/or WRFG simulations forced by the HADCM might be valuable, given the information presented here). Having each regional model driven by two GCMs and each GCM force three different RCMs has also proven to be a very useful setup when it comes to diagnosing the problems and potential in the various models.

This study presents another example of how difficult it is to determine how to evaluate climate models
FIG. 18. (a)–(j) The 1971–99 JJAS average SSTs from the GCM-forced RCMs. (k) 1982–2004 JJAS average SSTs from NOAA.
(regional or global) to determine their credibility (e.g., Gleckler et al. 2008; Knutti et al. 2010). For example, how important is a large warm bias across central North America when analyzing the NAM? The warm bias in question belongs to the HRM3-ncep, as shown in Mearns et al. (2012) and illustrated in Fig. 6. It has been suggested by some that this simulation be ignored in NARCCAP ensemble analysis as a result. However, the HRM3 is one of the better performers in this region, even with that bias. Either warm/cold or wet/dry, this type of bias, at face value, does not mean that there is no useful information to be gleaned from the simulation. It takes further process-level evaluation of the models to figure that out. The same could be said of a simulation with little of this type of bias as well.

We also ask, given this analysis, whether the compensating bias is always undesirable. We argue that it depends on the biases involved. Of course, perfect simulations would be ideal, but in the face of imperfection we may have some simulations in this analysis that may be better, and potentially more useful, than they would be without them. The high near-surface humidity bias in the CRCM, for example, may be responsible for too much precipitation in the NCEP-driven simulations in the Southwest, but it may also be aiding the CRCM when it is forced by two GCMs that do not provide enough moisture. Furthermore, the HRM3-ncep may not have as strong of a warm bias in the core monsoon region as it does in the plains, but it is still about 3°C too warm in this region, over land only, which could be enhancing the onshore monsoon flow it is producing (though this is not as great of a problem in the GCM-driven simulations and the flow is still acceptable). Also, while we think it is an overall problematic bias, the warm SST bias near the west coast of the continents in many GCMs is the only reason some of the NARCCAP RCM-GCM combinations have reasonable SSTs in the GOC. Without it, the SSTs would be too cold in the GOC in the RCMs, as they are in the CGCM-forced simulations and the flow is still acceptable. This discussion is in no way an argument in favor of this SST bias but perhaps a note to the RCM community simulating this region, or similar areas with water bodies that are unresolved in the chosen driver, that a better way of dealing with water surface temperatures, aside from interpolating them from the nearest ocean point in the coarser parent dataset, is warranted.

As the forcing conditions for an RCM do have a strong effect on RCM credibility, a conclusion solidified here, an argument could be made for selecting GCMs for nested regional modeling that provide “sensible” boundary conditions for the given study area. For this region, out of the NARCCAP suite, that GCM is clearly the HADCM. Correspondingly, for their analysis, Dominguez et al. (2010) chose two CMIP3 GCMs that most accurately represented selected relevant processes important to precipitation in the southwest United States in winter, and one of their chosen or “best” GCMs was a realization of the HADCM. However, Cerezo-Mota et al. (2011), in examining a CMIP3 realization of
the HADCM (not the one used here), found that it has a dry bias over the GOM and warn that it would not provide realistic boundary conditions and therefore realistic projections for the NAM. This suggests that finding the best GCM for a dynamical downscaling activity is not easy or straightforward. Compounding this issue here would be the size of the NARCCAP domain. Without substantial improvements in GCM simulations,
it may be impossible to find high-quality GCM boundary conditions that would be sufficient for all North American subregions.

While it is outside the scope of this paper to examine future climate changes from these models, the impact of the known biases on the future simulations, and our confidence in the simulations from a climate change standpoint, these are the subject of ongoing work. Clearly, we might have more confidence in the projections from some of these simulations over others, but they may all contain some useful information.

6. Summary

In-depth analysis of both parent GCM and RCM scenarios has identified a reasonable sample of models that can produce credible simulations of North American monsoon precipitation.

In this analysis, we have shown that the six RCMs when forced with the NCEP reanalysis and the two AGCM time slices, forced with observed SSTs, produce monsoon system-related precipitation in the core region in a manner that implies that they are reasonably simulating the basic large-scale mechanisms behind the monsoon precipitation in Mexico, but most of the RCMs and the time slices have a strong dry bias in AZ. The CRCM is the exception, with a wet bias, and the MM5I is the driest, with no discernible monsoon signal in its precipitation climatology. Realistic precipitation in AZ requires that the models produce a LLJ over the GOC and near-surface onshore flow, which not all of them do. Overall, WRFG appears the least biased in this region, arguably followed by the HRM3 and CRCM, although assigning differential credibility here is not straightforward, as none is perfect.

The skill of the RCMs when forced by different GCMs is largely controlled by the driving GCMs. Dry biases and other compounding problems in the CGCM and CCSM overshadow realistic large-scale forcing mechanisms and inhibit precipitation in the RCMs they force. The GFDL GCM, on the other hand, has problems later in the season and forces unrealistic precipitation, such that a monsoon appears to continue through December. The HADCM clearly provides the least problematic boundary conditions to the RCMs and only appears responsible for an early monsoon onset.

We have also shown that the RCMs do add value to the GCM simulations. The spatial distribution of precipitation is superior to that in the GCMs, even when the magnitude is highly biased because of GCM error, but that is not surprising given the increased detail in the region’s terrain. Nonetheless, because of the increase in resolution, the better-defined orography, coastal features of the GOC, and Baja Peninsula, we have shown that some of the RCMs can produce the monsoon system LLJ over the GOC even at 50 km. The RCMs that can simulate this in the NCEP-driven simulations retain this skill in the GCM-driven simulations. This appears to allow the HRM3, for example, to produce a better climatology of precipitation in AZ in the GFDL-driven simulations compared to the other RCMs. The other two models that are more skillful in this aspect, the WRFG and the CRCM, are both forced by the same two GCMs that are plagued with a dry bias, among other issues, and therefore contain a worse simulation of precipitation than the GCMs despite their ability to simulate a GOC LLJ. Whether or not they would improve if forced by other GCMs (the HADCM for instance) is, unfortunately, an unanswerable question at this time.

Acknowledgments. We wish to thank NARCCAP for providing the data used in this paper. NARCCAP is funded by the National Science Foundation (NSF), the U.S. DOE, NOAA, and the U.S. Environmental Protection Agency (EPA) Office of Research and Development. We would also like to thank the NARCCAP modeling team for useful discussions regarding this work. Likewise, we would like to acknowledge the other following data providers: NOAA/OAR/ESRL PSD for the NOAA SSTs and UDEL (www.esrl.noaa.gov/psd/), Colorado State University for TRMM and NAME data (http://rain.atmos.colostate.edu/CRDC/datasets/TRMM_3B42.html; http://tornado.atmos.colostate.edu/name/products/gridded/), NCEP for NARR and the NCEP reanalysis (http://www.emc.ncep.noaa.gov/mmb/reanalysis/; http://www.cpc.ncep.noaa.gov/products/wesley/reanalysis2/), the NOAA–CPC Climate Diagnostics Center for the CPC U.S.–Mexico precipitation dataset (http://www.cdc.noaa.gov/), and the Research Data Archive, maintained by the Computational and Information Systems Laboratory at NCAR, for the CFSR (http://rda.ucar.edu/datasets/ds093.1/).

We also acknowledge the modeling groups, the PCMDI, and the WCRP’s Working Group on Coupled Modeling for their roles in making available the WCRP CMIP3 multimodel dataset. Support for the CMIP3 dataset is provided by the Office of Science, U.S. DOE.

The authors also acknowledge the support of NOAA Climate Program Office Modeling, Analysis, Predictions and Projections (MAPP) Program. Work was supported under Grant NA11AOR4310111.

REFERENCES


Huffman, G. J., R. F. Adler, B. Rudolf, U. Schneider, and P. R. Keerh, 1995: Global precipitation estimates based on a