Land Surface Heating and the North American Monsoon Anticyclone: Model Evaluation from Diurnal to Seasonal

PATRICK KELLY AND BRIAN MAPES
Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida

(Manuscript received 6 July 2009, in final form 23 March 2010)

ABSTRACT

Data from several regional and global models (including model-based analysis data) are compared with field data from the North American Monsoon Experiment (NAME), from observatioanal sites as well as satellite retrievals. On the regional scale (NAME tier 1.5), sensible heating is shown to exceed latent and is furthermore concentrated in the lower half of the troposphere, so in considering the North American monsoon (NAM) midlevel anticyclone, the authors focus on radiative and turbulent energy fluxes at the surface. Models exhibit large discrepancies in their simulation of the mean diurnal cycle of these fluxes as well as in their sensitivity of evaporative fraction to recent rainfall. Most of the models examined have too much net radiation due to excessive shortwave surface flux (too little cloud) and too much sensible heating. These high biases in sensible heating appear to drive overpredictions of both the daily and seasonal rise of 500-hPa heights in the NAM anticyclone. This diurnal–seasonal resemblance suggests that calibrating surface heating processes using readily field-observed diurnal variations could lead to improvements in seasonal-time-scale NAM simulations.

1. Introduction

The North American monsoon (NAM) is a circulation pattern that brings summer rains to northwest Mexico and the southwest United States (see Adams and Comrie 1997 for a review). Synoptically, summer rainfall of the NAM involves moisture circulated by a “dirty ridge,” that is, an anticyclone in the lower-middle free troposphere marked by deep convection. This anticyclone can be thought of as the top part of a thermal low that develops in the desert Southwest (Tang and Reiter 1984; Rowson and Colucci 1992). Increased precipitation in northwest Mexico and the southwest United States occurs as enhanced easterlies along the southern flank of the anticyclone bring moist air from the Gulf of Mexico into the region (e.g., Bryson and Lowry 1955; Higgins and Wang 1997; Douglas et al. 1993). Moisture injection at low levels is also accomplished by gulf surges into the thermal low over the warm Gulf of California (e.g., Bordoni et al. 2004; Bordoni and Stevens 2006; Gochis et al. 2004; Johnson et al. 2007).

Model prediction of NAM rainfall remains a particularly difficult challenge in this region of complex terrain (Maddox et al. 1995; Dunn and Horel 1994). The North American Monsoon Experiment of 2004 (NAME-2004) was an international field campaign that collected unprecedented meteorological observations during summer 2004 over southwest North America (Higgins et al. 2006). To harvest the fruits of this campaign for model improvement, a “climate process team” was formed, with academic and operational scientists collaborating to connect observations to model-based seasonal to interannual climate prediction efforts. One part of this team’s activities was to organize the second phase of the North American Monsoon Model Assessment Project (NAMAP2; Gutzler et al. 2009). The goal of NAMAP2 is to evaluate model progress since the original NAMAP activity (Gutzler et al. 2005) and to deepen model evaluation with additional diagnosis made possible by NAME-2004 field data. The full suite of NAMAP2 modeling results is publicly available at the NAMAP2 Web atlas (available online at http://www.rsmas.miami.edu/users/pkelly/NAMAP2.html).

If model performance is to be improved, we must connect seasonal-time-scale errors in the simulation of the NAM to deficiencies in physical processes, which are mainly parameterized in the vertical domain and act on...
much shorter time steps (minutes to hours). Rainfall is the main impact variable of the NAM, so rainfall prediction is certainly of interest (as considered in Gutzler et al. 2009). However, rainfall-producing processes are spotty and intermittent, with multiple nonlocal causes that evade the point-based observations and evaluations considered here. Our working hope is that, at the seasonal scale, there may be a predictability to moisture delivery by the continental-scale NAM flow that does not hinge entirely on simulating the details of where and when this moisture falls as rain.

For this anticyclone-dominated monsoon flow over a semiarid domain, our hypothesis (supported below and by Kelly 2008) is that latent heating by the rain itself is not the central driving process but merely a partial feedback on a flow feature driven primarily by sensible heating. Based on this reasoning, we focus here on surface heating and the chain of land–atmosphere interaction processes that govern it as the drivers of the NAM anticyclone.

Section 2 describes the data used. Section 3 evaluates NAMAP2 models in terms of the simulation of solar radiation, net radiation, and sensible versus latent partitioning of the resulting net flux into the atmosphere. Section 4 discusses the sensitivity of evaporative fraction to recent rain events. Finally, section 5 demonstrates that model systematic errors in these processes, as deduced on diurnal time scales, parallel the models' seasonal systematic errors in lower-tropospheric thickness and thus the NAM midlevel anticyclone.

2. Data sources

a. Observations

Point observations of downwelling shortwave (SW), net absorbed SW (downwelling minus upwelling component), net radiation (Rnet), and sensible heat flux (SHF) and latent heat fluxes (LHF) come from the Ameriflux database (available online at http://public.ornl.gov/ameriflux/). Because the core of the NAM anticyclone is over southeast Arizona, three nearby stations in that region are used as the baseline for model evaluation (see Fig. 1): Audubon Grasslands, Santa Rita, and Walnut Gulch. Data from the 3 sites are sufficiently similar that a 3-site average is used as a baseline for difference plots below (see Fig. 2). Data are available at 30-min intervals and are composited to form diurnal cycle averages for July–August 2004.

To supplement in situ data, satellite estimates of surface albedo, net SW, and net radiation from the National Aeronautic and Space Administration’s (NASA) Clouds and the Earth’s Radiant Energy System (CERES) are also consulted for July–August 2004. CERES datasets are derived on a $1^\circ \times 1^\circ$ global grid. Surface radiation budget calculations are performed with two sets of algorithms, known as primary and quality-check algorithms. The primary SW algorithm is adapted from Pinker and Laszlo (1992) and the primary longwave (LW) algorithm is an adaptation of Fu et al. (1997). The quality-check SW algorithm is the Langley Parameterized Shortwave Algorithm (Gupta et al. 2001), and the quality-check LW algorithm is described in Gupta et al. (1992).

We use 500-hPa geopotential height fields (Z500) during the 2004 season as a metric for identifying the synoptic footprint of the NAM. The 500-hPa level marks the mean vertical position of the monsoon anticyclone, whose northward march during boreal summer coincides with the onset of the NAM (Higgins and Wang 1997). Observations of geopotential height at Tucson, Arizona, come from the National Oceanic and Atmospheric Administration’s radiosonde database (available online at http://www.esrl.noaa.gov/raobs/). A July–August 2004 average is used, consistent with surface flux observations and because the mature phase of the NAM is in midsummer (e.g., Barlow et al. 1998). To calculate a seasonal index of the Z500 anticyclone over land for section 5, July–August 2004 average of Z500 over the East Pacific box (EPAC) is subtracted from raw observations at Tucson (Fig. 1). Z500 data over EPAC come from the National Centers for Environmental Prediction (NCEP) global reanalysis dataset version 1 (Kalnay et al. 1996). Simply using the reanalysis data for both land and sea regions gives similar results.
b. Models

The various participating models and their key characteristics are listed in Table 1. As part of NAMAP2, modeling groups generated special high-resolution column datasets at the NAME-2004 sounding sites. These are called Model Output Location Time Series (MOLTS) data, from which we extracted surface fluxes for the grid box containing Tucson to compare against flux site observations in section 3. The spatial distance between Tucson and the center of observations is generally less than individual model gridcell spacing (see Fig. 1), although landscape heterogeneity must always be remembered as a challenge for point versus gridcell comparisons. Model bias is defined as the model value minus observed value. Radiative fluxes are defined as positive into the ground, while sensible and latent fluxes are defined as positive out of the ground, in order to make all numbers usually positive during the daytime.

We also consulted MOLTS data from a total of 12 different sites in the NAM region (not shown) in section 5b to determine if inferences of surface moisture–heat flux sensitivities at Tucson are representative of model processes, even though observations are not available everywhere. For this larger intermodel comparison, we used June–September (JJAS) means to maximize data usage and to increase our sample size; July–August data give similar results and conclusions. MOLTS data were produced at a total of 25 different sites, but we found it necessary to omit coastal sites, where model grid cells are contaminated by subgrid-scale water due to inadequate horizontal resolution in complex geography.

3. Surface flux evaluations

a. Local diurnal composites

The mean diurnal cycle of surface fluxes in models is evaluated in Fig. 3. Diurnal composites of the bias (model minus the observations of Fig. 2) are plotted, and the 24-h mean is listed in Table 2. All of the models except the Regional Spectral Model (RSM) overestimate downwelling SW radiation during summer 2004 by

---

1 Downwelling SW fluxes are not available for the GEOS and CAMb,c models.
several tens of W m$^{-2}$ (Fig. 3a), suggesting an under-representation of cloudiness and/or weak cloud–radiation interaction. The finite-volume model (FVM) has the largest error, with a mean bias of nearly 1.90 W m$^{-2}$ over the 24-h period.

Absorbed net SW radiation errors are of the same sign as downwelling SW errors—positive for all models except RSM—although less in magnitude (Fig. 3b). Again, the FVM has far too much SW radiation in the afternoon, with a peak bias of almost 180 W m$^{-2}$. Mean daily biases in the three different Community Atmosphere Model (CAM) simulations are all approximately 50 W m$^{-2}$, while SW errors in the National Oceanic and Atmospheric Administration’s (NOAA) operational Global Forecast System (GFS) and Climate Forecast System (CFS) for instance, are partially compensated by their higher albedo values, which act to reduce the net SW biases. Alternatively, the RSM has too little downwelling SW flux. Its net SW bias is also reduced however (becomes less negative), because it has a more absorptive surface (lower albedo).

Net radiation biases (Table 2 and Fig. 3c) resemble Fig. 3b although the errors are generally smaller as models that overestimate incoming SW radiation also overestimate outgoing LW radiation. Again, the RSM is an outlier: It has a daily mean net radiation bias of +20 W m$^{-2}$ despite too little daytime SW flux. Positive net radiation biases at night may suggest a temperature inversion (cold surface) and perhaps a too emissive (moist or cloudy) boundary layer. The GFS captures the diurnal cycle of net surface radiation the best, with negligible biases throughout the 24-h period. Considering model performance

### Table 1. Key characteristics of various NAMAP2 model and reanalysis data analyzed in this study. Acronym listed is how each dataset is referenced throughout the text.

<table>
<thead>
<tr>
<th>Model</th>
<th>Acronym</th>
<th>Affiliation</th>
<th>Reference</th>
<th>Horizontal resolution</th>
<th>Ensemble size</th>
<th>SST prescription</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate Forecast System</td>
<td>CFS</td>
<td>NOAA/Climate Prediction Center (CPC)</td>
<td>Saha et al. (2006)</td>
<td>T126</td>
<td>5</td>
<td>Multiplatform merged (MPM)</td>
</tr>
<tr>
<td>Global Forecast System</td>
<td>GFS</td>
<td>NOAA/CPC</td>
<td>Campana and Caplan (2005)</td>
<td>T126</td>
<td>4</td>
<td>MPM</td>
</tr>
<tr>
<td>National Center for Atmospheric Research (NCAR) Community Atmosphere Model version 3</td>
<td>CAMa</td>
<td>NCAR</td>
<td>Collins et al. (2006)</td>
<td>1.0° × 1.25°</td>
<td>1</td>
<td>Hadley</td>
</tr>
<tr>
<td>NCAR Community Atmosphere Model version 3</td>
<td>CAMb</td>
<td>Scripps Institution of Oceanography (SIO)</td>
<td>Collins et al. (2006)</td>
<td>T42</td>
<td>1</td>
<td>40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40)</td>
</tr>
<tr>
<td>NCAR Community Atmosphere Model version 3</td>
<td>CAMc</td>
<td>SIO</td>
<td>Collins et al. (2006)</td>
<td>T42</td>
<td>3</td>
<td>MPM</td>
</tr>
<tr>
<td>Regional Spectral Model</td>
<td>RSM</td>
<td>SIO</td>
<td>Juang and Kanamitsu (1994)</td>
<td>30 km</td>
<td>1</td>
<td>MPM</td>
</tr>
<tr>
<td>Finite-volume general circulation model</td>
<td>FVM</td>
<td>NASA/Goddard Space Flight Center (GSFC)</td>
<td>Lin (2004)</td>
<td>0.25° × 0.36°</td>
<td>2</td>
<td>MPM</td>
</tr>
<tr>
<td>Goddard Earth Observing System Model version 5</td>
<td>GEOS</td>
<td>NASA/GSFC</td>
<td>Rienecker et al. (2008)</td>
<td>0.5° × 0.67°</td>
<td>5</td>
<td>MPM</td>
</tr>
<tr>
<td>North American Regional Reanalysis</td>
<td>NARR</td>
<td>NOAA/CPC</td>
<td>Mesinger et al. (2006)</td>
<td>0.33°</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
collectively, seasonal (July–August) radiation is overestimated an average of about 18 W m$^{-2}$ (Table 2).

SHF and LHF (Figs. 3d,e) show how this net surface radiation is partitioned back to the atmosphere. The sum of latent and sensible fluxes is approximately equal to net radiation, since ground heat storage is negligible beyond a few hours. All models have excessive sensible heating in the afternoon (Fig. 3d; Table 2). The three NCEP models (GFS, CFS, and RSM) are all at the high end. Daytime peak biases range from $\sim$90 W m$^{-2}$ in the CAMa, to $\sim$230 W m$^{-2}$ in the CFS. Models that have large positive sensible heating errors tend to have too little LHF biases, but the Goddard Earth Observing System (GEOS) and CAMa models manage to have too much latent as well as sensible flux (Fig. 3e).

b. Regionally averaged comparisons

Here we build upon our inferences of local model biases made above and extend them to regional model evaluations of surface heating processes. Table 4 lists model biases of surface net SW and net radiation relative to CERES averaged across the tier 1.5 domain (see Fig. 1). As before, data were averaged over July–August 2004. Evaluation of surface radiative heating at Tucson is characteristic of model performance across the broader NAM region: All models have excessive net SW flux, except RSM. Moreover, the collective model bias of seasonal net radiation is $\sim$18 W m$^{-2}$ at both Tucson and for tier 1.5 (Tables 2 and 4). Satellite-measured net surface radiation must be interpreted cautiously, since the LW component is not directly measured. However, this general consistency between ground-referenced model biases in southwest Arizona and satellite-referenced model biases across the entire NAM region suggests conclusions.

Table 2. Seasonal mean (July–August 2004) model biases of surface fluxes at Tucson, AZ. Bias is defined as model value minus 3-site composite averaged observations. All values are in units of W m$^{-2}$.

<table>
<thead>
<tr>
<th>Model</th>
<th>SWdown</th>
<th>SWnet</th>
<th>Rnet</th>
<th>SHF</th>
<th>LHF</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAMa</td>
<td>59.5</td>
<td>47.8</td>
<td>30.9</td>
<td>21.6</td>
<td>24.1</td>
</tr>
<tr>
<td>CAMb</td>
<td>—</td>
<td>46.9</td>
<td>8.5</td>
<td>28.4</td>
<td>–13.4</td>
</tr>
<tr>
<td>CAMc</td>
<td>—</td>
<td>52.1</td>
<td>3.9</td>
<td>34.9</td>
<td>–28.8</td>
</tr>
<tr>
<td>CFS</td>
<td>48.3</td>
<td>21.5</td>
<td>18.0</td>
<td>41.5</td>
<td>–6.8</td>
</tr>
<tr>
<td>GFS</td>
<td>36.3</td>
<td>22.2</td>
<td>7.0</td>
<td>44.2</td>
<td>–23.0</td>
</tr>
<tr>
<td>RSM</td>
<td>–16.8</td>
<td>–2.1</td>
<td>21.3</td>
<td>56.7</td>
<td>–17.7</td>
</tr>
<tr>
<td>FVM</td>
<td>89.5</td>
<td>54.8</td>
<td>27.8</td>
<td>56.1</td>
<td>–13.2</td>
</tr>
<tr>
<td>GEOS</td>
<td>—</td>
<td>24.8</td>
<td>16.2</td>
<td>26.3</td>
<td>14.5</td>
</tr>
<tr>
<td>NARR</td>
<td>59.7</td>
<td>43.3</td>
<td>24.5</td>
<td>56.3</td>
<td>7.9</td>
</tr>
<tr>
<td>Mean</td>
<td>46.1</td>
<td>34.6</td>
<td>17.6</td>
<td>40.7</td>
<td>–6.3</td>
</tr>
</tbody>
</table>
drawn from a single site are representative of model performance. Implications of this excessive surface net radiation in models—driven by positive net SW flux biases—for the heating of the lower-middle free troposphere and the simulation of the NAM anticyclone are discussed below in section 5.

4. Evaporative fraction–rain sensitivity

a. One site and month in detail

Latent and sensible fluxes are highly dependent on available soil moisture (and hence, prior rainfall), as well as on surface flux parameterizations schemes. These dependences are obscured in seasonal mean data, so Fig. 4 shows high-frequency time variation of latent and sensible fluxes over a month. Precipitation is also shown, from which soil moisture variations may be broadly inferred. We use time series at Santa Rita, Arizona, in observations and model data interpolated to Tucson,2 for August 2004. This month was chosen as an illustrative example because both models and observations have a good sample of rain and no-rain days.

Despite the sporadic nature of rain events, day-to-day changes in peak latent fluxes are quite small and gradual in observations (top row of Fig. 4): Sensible flux is comparable to latent flux (∼200 W m⁻²) even on intense rain days in the beginning and middle of the month, while sensible flux is ∼3 times larger than latent flux during a dry week at the end of the month. Fluxes in the NARR (second row of Fig. 4), which assimilates observed precipitation, are similarly unresponsive to rain events: LHF remains small even on the rainiest days. In contrast, CAM fluxes are extremely sensitive to rain events and exhibit a very bimodal distribution: LHF is enhanced and SHF is reduced by a factor of about 4 on rainy days. This extreme surface wetness sensitivity apparently involves canopy-captured rain reevaporation, aided by a lack of cloud shading during rain events, as finite convective cloud fraction without condensed water content made no impact on the radiation scheme. This weakness has been noted by others (e.g., DeMott et al. 2007; Guo et al. 2006) and improved upon in subsequent CAM model revisions (Lawrence et al. 2007). Results from the CFS, FVM, and GEOS models lie somewhere in the middle, with surface flux sensitivities larger than the NARR, but less extreme than the CAM’s.

Evaporative fraction (EF) is defined as EF = LHF/(SHF + LHF) and conveniently summarizes the relationships seen in the individual fluxes. In Fig. 5, the solid and dashed lines represent composite averages of EF for “wet” and “dry” tercile days, respectively, for the 3-station observation and again at Tucson in models. Terciles are defined as the 10 days in August 2004 when the EF value at 0900 LST is the highest and lowest, respectively. The EF at 0900 LST time is chosen to avoid the vagaries of using highly variable precipitation, yet still separate days based on the initial morning conditions for boundary layer development and lower troposphere heating (thickening) later in the afternoon.

EF differences between wet and dry composites are consistent with inferences made above: EF increases by about 0.2 from dry to wet days in observations and the NARR, but much more in the models. The CFS, FVM, and GEOS models all exhibit a relatively similar structure, with EF differences of about 0.4, while wet–dry differences are clearly excessive in the CAM panels.

b. Seasonal consistency across sites

A larger sample of data comparisons confirms that the results above are representative of model performance. Figure 6 shows scatterplots of daily mean latent flux normalized by net radiation versus daily precipitation at 12 model grid sites scattered across Mexico and the southwest United States from June to September. Since

---

2 Results for the GFS, RSM, and CAMc are omitted for brevity. The GFS and RSM use the same NOAH LSM as NARR, while CAMc uses the same LSM as CAMb. Results are qualitatively similar to those shown.
daily rainfall data are very noisy, we use the June–September period to increase our sampling size of non-zero rainy days. Each square represents an individual day at a particular grid location. A simple box-car moving average (using a rainfall bin size of 0.5 mm) was applied to each model to estimate a fit to the data (heavy curve in each panel).

Intermodel LHF comparisons in Fig. 6 are generally consistent with Fig. 5: The CAMa and CAMb curves have a steep slope at low rain rates, indicating high LHF sensitivity to even small precipitation amounts, and rapidly asymptotes to near 95% of the upper bound of net radiation. In contrast, LHF is only around 65% of net radiation on even the rainiest days in the GFS, RSM, and NARR, all of which use the same (Noah) land surface model. Again the CFS, FVM, and GEOS models are in the middle of the range here, with maximum LHF around 85% of net radiation during rainy days. In short, there is clear consistency between samples at a single grid cell considered above (Figs. 4, 5, and others not shown) and results averaged over the whole summer (JJAS) and across the entire NAM domain in models (Fig. 6).

Given the complex landscape of the NAM region, it is prudent to examine observations in other types of terrain than southeast Arizona. Identical analysis to Fig. 5 was performed at Tesopasco and Estación Obispo, Mexico (indicated symbols on Fig. 1). The former is an upland site surrounded by tropical deciduous forest (Watts et al. 2007), while the latter is along the coastal plain of the Gulf of California. Both sites had surface radiation and turbulent heat flux measurements during summer 2004. Results in those locations were similar to Arizona observations in showing weak soil moisture–LHF feedbacks, with an EF increase of about 0.2 between wet and dry days (not shown). If this consistency across the region seen in models also holds in observations, as our sample of 5 sites may suggest, then field data can indeed inform large-scale, seasonal model evaluation meaningfully.
5. **Z500: Diurnal–seasonal connection**

Since the overarching goal of this work is to address seasonal model errors, we are led to hope (or more formally, to hypothesize) that diurnal biases in models may correspond fairly directly to biases in the seasonal development of the NAM midlevel anticyclone. This reasoning is premised on the importance of bottom-boundary sensible heating in driving the NAM anticyclone.

As a first step, we assess the continental heat source in NAMAP2 models by comparing averages of SHF against precipitation-induced latent heating over tier 1.5 (Table 5). Tier 1.5 is used for averaging as it encompasses the lateral boundary of the NAM anticyclone (Fig. 1), including the rainy core region of northwest Mexico as well as the southwest United States where precipitation is more marginal. The tier 1.5 spatial domain also extends to the east of these regions (encompassing western Texas and eastern Mexico) since a common modeling flaw is simulation of precipitation too far to the east relative to observations (Gutzler et al. 2009).

Table 5 shows mean July–August 2004 values of SHF and latent heating derived from precipitation for the tier 1.5 domain. All models have a greater amount of seasonal mean sensible heating than latent heating from condensation, except for the relatively rainy CAMb and GEOS5 models. The suite of NAMAP2 models collectively has about 15% more sensible heating than latent heating. Furthermore, sensible heating is more efficient at driving the midlevel anticyclone: it is concentrated in the lower troposphere, which generates anticyclonic potential vorticity at the top of the heating (i.e., in the midlevel anticyclone). In other words, even these comparable values of column-integrated sensible and latent heating point to a greater importance of sensible heating in terms of lower-tropospheric thickening and anticyclone development.

Since sensible heating is a primary process governing NAM anticyclone intensity, systematic model biases in SHF should drive biases in lower-troposphere thickness on both the diurnal and seasonal time scales. We test this hypothesis using geopotential height data in the southwest...
United States, building on the conclusions drawn in section 3. Models that overestimate diurnal thickening of the lower troposphere, and thus rising of the 500-hPa surface’s height (Z500), might be expected to overestimate the NAM anticyclone’s seasonal development as well. On the other hand, this relationship is by no means guaranteed, since Z500 is not simply locally and thermally governed.

To estimate the NAM anticyclone’s seasonal range in models, a continent minus ocean Z500 difference is calculated between the Arizona–New Mexico box (AZNM) and the EPAC box (Fig. 1). All data are averaged in time over July–August 2004 to stay consistent with previous surface energy budget analysis and because the NAM anticyclone is maturely developed in midsummer. We focus on the AZNM region, because the observed midlevel NAM high is centered here (see Fig. 1), and models generally produce a Z500 maximum near this region (see Web atlas). Figure 7 shows a scatterplot of the diurnal range of Z500 (calculated as 0000 minus 1200 UTC; roughly 1700 minus 0500 LST) at AZNM versus its estimated seasonal range, in NAMAP2 models discussed previously, as well as the NCEP global reanalysis dataset version 1 (Kalnay et al. 1996). An observational estimate is also computed using raw observations at Tucson (located within the AZNM domain) minus the EPAC average from NCEP global reanalysis data.

A very good intermodel correlation ($r = 0.88$) is seen. Most models tend to overdo both the diurnal and seasonal thickening of the lower troposphere, consistent with their high biases in net radiation and sensible heat flux (Figs. 3c,d; Table 2). The CFS, GFS, RSM, and FVM models with their excessive net radiation and large positive bias in diurnal SHF, have large diurnal Z500 rises, and also show a positive bias in their seasonal high. The GEOS, NARR, and CAM models have a comparably lower diurnal sensible heating range and similarly lower diurnal and seasonal Z500 ranges.

In summary, diurnal changes of Z500 in the NAM high, driven by sensible heating over the continent, appear to prevail in a rectified way on the seasonal scale, despite large-scale horizontal dynamics that surely blends, and in principle could even counteract, these one-dimensional thermal influences. While this small sample of models makes the result less than conclusive, it does appear that improving the accuracy of diurnal heating processes in models, which can be done based on observations in a single field season, might be expected to improve seasonal NAM simulations and perhaps forecast accuracy.

**FIG. 6.** Scatterplots of daily mean normalized latent heat flux (LHF/Rnet) vs daily precipitation (open squares) at 12 MOLTS sites across Mexico and the southwest United States from JJAS data. Heavy black line indicates box-car moving average.
6. Discussion

A major justification for field campaigns is to help inform model development and improvement. Unfortunately, field observations are often too limited to draw robust, broad conclusions about model performance. Although the NAMAP2 model assessment spans only the 2004 summer season, it appears sufficient to constrain some key processes and indicate model shortcomings.

The initial NAMAP study already highlighted model difficulties and differences in the simulation of turbulent heat fluxes in the surface energy budget (Gutzler et al. 2004). Lack of sufficient observations prior to NAME-2004 made adjudication of these intermodel differences difficult (Gutzler et al. 2005). This follow-up NAMAP2 exercise (see also Gutzler et al. 2009) reiterates this intermodel variability, but now we can conclude that the reason that all models appear to have locally excessive sensible heating is as a result of too much surface shortwave radiation (Fig. 3c). Furthermore, these local heating biases appear to add up to biases on seasonal–regional scales (Fig. 7). Since sensible heating over the tier 1.5 area (Fig. 1) exceeds latent heating (Table 5) and is concentrated over a thinner layer (the lower half of the troposphere), it is a primary process in governing NAM midlevel anticyclone intensity, with implications for this elevated flow’s ability to carry moisture inland beyond the low deserts of the Gulf of California region. This reasoning and its conclusions are consistent with previous modeling work that showed a decrease in NAM precipitation associated with a weaker midlevel circulation as a result of increasing surface albedo over southwest North America (Shaffrey et al. 2002).

In summary, our working hypothesis—that the NAM anticyclone is largely governed by the degree of near-surface sensible heating—appears to be confirmed. Paradoxically, improving dry heating processes in models may play a key role in improving predictions of warm-season precipitation in the western United States at the NAM’s northern edge. The next steps beyond this multimodel assessment will be model experimentation, to further explore the connection between surface heating processes and the seasonal NAM circulation, seeking robust improvements to regional simulations from observation-constrained treatments of radiative and thermodynamic processes.

Acknowledgments. This work was supported by a grant from the NOAA Climate Program Office’s Climate Prediction Program for the Americas (program GC05-039) and by the National Science Foundation under Grant 0731520. We thank the modeling centers, the many people who worked to make NAME-2004 a reality, and the ongoing AmeriFlux project for supplying the data analyzed in this study. The technical support of Lindsey Long at NOAA/CPC is especially appreciated.

REFERENCES

http://www.emc.ncep.noaa.gov/gc_wmb/Documentation/TPB0ct05/T382.TPB.FINAL.htm.


