SPECIAL
Land–Atmosphere Exchanges of Energy and Water
COLLECTION

Diagnosing the Nature of Land–Atmosphere Coupling: A Case Study of Dry/Wet Extremes in the U.S. Southern Great Plains

JOSEPH A. SANTANELLO JR. AND CHRISTA D. PETERS-LIDARD

NASA GSFC Hydrological Sciences Branch, Greenbelt, Maryland

AARON KENNEDY

University of North Dakota, Grand Forks, North Dakota

SUJAY V. KUMAR

NASA GSFC Hydrological Sciences Branch, Greenbelt, Maryland, and Science Applications International Corporation, McLean, Virginia

(Manuscript received 14 February 2012, in final form 1 June 2012)

ABSTRACT

Land–atmosphere (L–A) interactions play a critical role in determining the diurnal evolution of land surface and planetary boundary layer (PBL) temperature and moisture states and fluxes. In turn, these interactions regulate the strength of the connection between surface moisture and precipitation in a coupled system. To address model deficiencies, recent studies have focused on development of diagnostics to quantify the strength and accuracy of the land–PBL coupling at the process level. In this paper, a diagnosis of the nature and impacts of local land–atmosphere coupling (LoCo) during dry and wet extreme conditions is presented using a combination of models and observations during the summers of 2006 and 2007 in the U.S. southern Great Plains. A range of diagnostics exploring the links and feedbacks between soil moisture and precipitation is applied to the dry/wet regimes exhibited in this region, and in the process, a thorough evaluation of nine different land–PBL scheme couplings is conducted under the umbrella of a high-resolution regional modeling test bed. Results show that the sign and magnitude of errors in land surface energy balance components are sensitive to the choice of land surface model, regime type, and running mode. In addition, LoCo diagnostics show that the sensitivity of L–A coupling is stronger toward the land during dry conditions, while the PBL scheme coupling becomes more important during the wet regime. Results also demonstrate how LoCo diagnostics can be applied to any modeling system (e.g., reanalysis products) in the context of their integrated impacts on the process chain connecting the land surface to the PBL and in support of hydrological anomalies.

1. Introduction

Quantification of the land surface influence on extremes such as flood and drought is critical for both short-term weather and climate prediction. These dry and wet regimes are modulated by the strength and sensitivity of the land–atmosphere (L–A) coupling and, in particular, how anomalies in soil moisture are translated into and through the planetary boundary layer (PBL), ultimately favoring or suppressing the triggering and support of clouds and precipitation. Improved understanding of L–A coupling is thus essential as a changing climate leads to evolving regions of dry and wet regimes, as well as locations where strong coupling is a dominant mechanism.

Recent studies have looked at the inherent L–A coupling strength (Koster et al. 2004) and predictability (van den Hurk et al. 2010; Koster et al. 2010) in models based on the role of soil moisture anomalies on seasonal precipitation. This was performed in a global climate model context using a large number of ensemble simulations, and therefore parsing out the reasons for differences in coupling strength amongst models (and inherent land surface and PBL physics) remains a difficult task. A companion effort has since been launched that focuses on local L–A coupling (LoCo; Santanello et al.

Corresponding author address: Joseph A. Santanello Jr., NASA GSFC Code 614.3, Bldg. 33, Room G220, Greenbelt, MD 20771.
E-mail: joseph.a.santanello@nasa.gov

DOI: 10.1175/JHM-D-12-023.1

© 2013 American Meteorological Society

2009, 2011) in coupled models by diagnosing land–PBL interactions at the process level using a regional, high-resolution test bed. The methodology and diagnostics developed in LoCo can be applied to any model or observations, and it is particularly well suited to isolate the impacts of land surface perturbations on the PBL (and vice versa) that are crucial for sustaining flood and drought conditions.

With these issues in mind, this paper presents results from case studies of dry/wet extremes in the U.S. southern Great Plains (SGP) to evaluate the performance of and coupling between a range of land surface models (LSMs) and PBL schemes (PBLs) by employing recently developed diagnostics of LoCo. Specifically, the goals of this study are to determine the following: 1) How well are extreme conditions represented in offline LSMs coupled to a high-resolution regional model, and what is the sensitivity in each regime to the choice of LSM–PBL parameterization and their coupling? 2) What are the characteristics of the local L–A coupling during dry/wet extremes, and how do LoCo processes act to support such events? 3) How well do large-scale, coarse-resolution models represent LoCo during dry/wet extremes?

This comprehensive analysis builds upon the work of Santanello et al. (2009, hereafter S09, 2011, hereafter S11) by applying the diagnostics developed in S09 and S11 to extended case studies and performing a thorough evaluation of coupling behavior in a range of models against observations. The case studies chosen for these experiments are composed of extreme dry and wet conditions in terms of soil moisture and precipitation relative to normal, and are therefore ideally suited to capture a wide range of variability in L–A interactions and coupling.

The paper follows with a summary of recent LoCo research and diagnostics in section 2, and by a description of the case studies, models, and observations employed therein in section 3. Results—including surface energy balance and LoCo metrics from diurnal cycle, composite, and reanalysis evaluations—are then presented in section 4. Finally, section 5 summarizes the conclusions and a discussion of the impact of the results on current and future research of coupling and its impact on extremes.

2. Background

A thorough review of LoCo research and the related diagnostic framework can be found in S09 and S11. The goals of the current work are to bring the methodologies of these studies to bear on evaluating the land–PBL coupling during climatological extremes in an array of LSMs and PBL schemes and performing a thorough evaluation of the schemes themselves. This research is a core component of the Global Energy and Water Cycle Study (GEWEX) Land Atmosphere System Study (GLASS; van den Hurk et al. 2011), which coordinates a community working group on studies related to L–A coupling (S11). LoCo is focused on the diurnal cycle and local-/regional-scale processes in models and observations, and in particular on quantifying the impact of the land condition on the atmosphere (through the PBL) and vice versa. There is a great deal of effort being put forth to better understand extremes, including their representation and predictability, primarily in global climate models and over large spatial scales (Koster et al. 2004; van den Hurk et al. 2010; Hirschi et al. 2011). The initial communication and interaction between the land and atmosphere always occurs on local scales, making the process-level understanding and focus of LoCo research essential in order to fully understand the impact of L–A coupling on dry and wet extremes.

As described in S11, from a LoCo perspective the land–PBL coupling can be broken down into a series of links in a ''process chain'' that connects soil moisture to precipitation:

$$\Delta\mathrm{SM} \underset{\text{(i)}}{\rightarrow} \Delta\mathrm{EF}_{\mathrm{sm}} \underset{\text{(ii)}}{\rightarrow} \Delta\mathrm{PBL} \underset{\text{(iii)}}{\rightarrow} \Delta\mathrm{ENT} \underset{\text{(iv)}}{\rightarrow} \Delta\mathrm{EF}_{\mathrm{atm}} \Rightarrow \Delta P/\mathrm{Clouds}, \quad (1)$$

where EF is the evaporative fraction, defined as

$$\mathrm{EF} = \frac{\mathrm{Qle}_{\mathrm{sfc}}}{\mathrm{Qh}_{\mathrm{sfc}} + \mathrm{Qle}_{\mathrm{sfc}}}, \quad (2)$$

and is a function of the sensible ($\mathrm{Qh}_{\mathrm{sfc}}$) and latent ($\mathrm{Qle}_{\mathrm{sfc}}$) heat fluxes at the land surface. From Eq. (1), the impact of soil moisture (ΔSM) on clouds and precipitation (ΔP) is therefore dependent on the sensitivities of (i) the surface fluxes ($\mathrm{EF}_{\mathrm{sm}}$) to soil moisture, (ii) PBL evolution to surface fluxes, (iii) entrainment fluxes at the PBL-top (ENT) to PBL evolution, and (iv) the collective feedback of the atmosphere (through the PBL) on surface fluxes ($\mathrm{EF}_{\mathrm{atm}}$) (Santanello et al. 2007; van Heerwaarden et al. 2009). Equation (1) describes a complex set of dependent relationships, and as Siqueira et al. (2009) and Ek and Holtslag (2004) highlight, the full set of L–A interactions (including those of negative feedbacks) is shown to be critical to understanding the full SM–P relationship.

To this end, a methodology that simultaneously addresses the components of Eq. (1) was tested by S09 and extended by S11, and employs the ''mixing diagram'' approach as introduced by Betts (1992). The power of this diagnostic lies in its ability to exploit the covariance of 2-m potential temperature (θ) and humidity (q) to quantify the components of the LoCo process chain, and

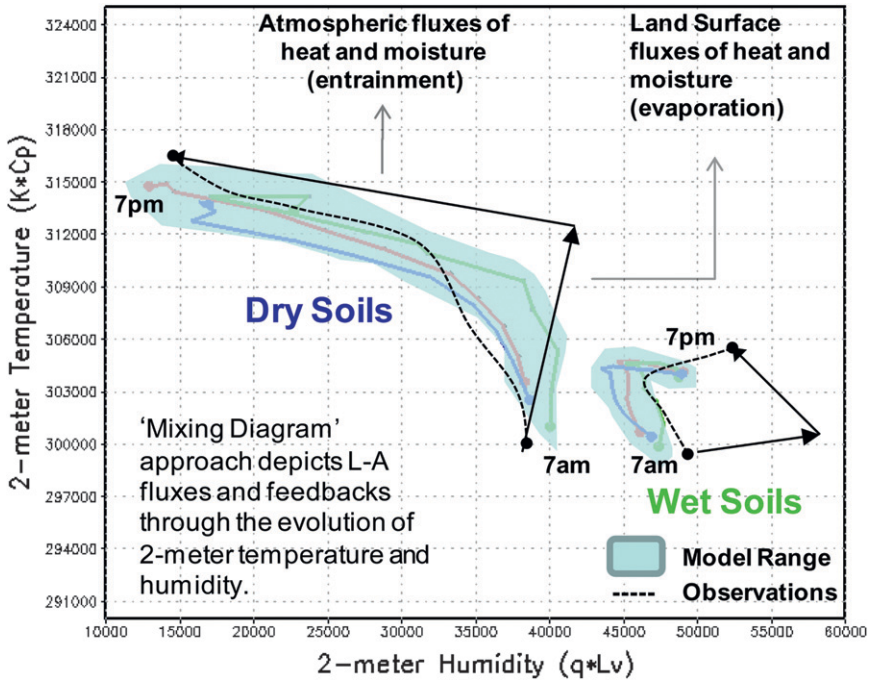


FIG. 1. Mixing diagram showing the diurnal coevolution (0700–1900 UTC) of 2-m specific humidity and 2-m potential temperature on 12 Jun 2002 at a dry and wet soil location as simulated by a coupled mesoscale model (derived from S09, their Figs. 2–5). The shaded regions for each indicate the model range for different LSM–PBL scheme couplings (red, green, and blue) vs what was observed (dashed black). Also shown for the dry site are the vectors that represent the fluxes of heat and moisture from the land surface vs those from the atmosphere due to entrainment, both of which are quantified using this approach.

a key advantage to this approach is that the calculations require only routine variables from observations and models. For a full description of this approach and implementation for LoCo studies, the reader is again referred to S09 and S11.

A summary of mixing diagram results from these studies is shown in Fig. 1, where simulations were run using a fully coupled regional modeling system, each with a different LSM–PBL scheme combination. The results show that soil moisture anomalies (dry versus wet) lead to different patterns of θ and q evolution throughout the day, as well as vector components that represent the contribution of heat and moisture from the land surface versus that from the top of the PBL via entrainment. In addition, derived metrics such as the surface and entrainment Bowen ratios ($\beta_{sfc} = Qh_{sfc}/Qle_{sfc}$, $\beta_{ent} = Qh_{ent}/Qle_{ent}$), and the entrainment ratios of heat and moisture ($A_h = Qh_{ent}/Qh_{sfc}$, $A_{le} = Qle_{ent}/Qle_{sfc}$) are useful diagnostics of the LSM–PBL coupling that can be easily derived from mixing diagrams.

Mixing diagrams diagnose the land and PBL fluxes simultaneously, and therefore provide the components of the full PBL budget of heat and moisture, which serves as the second core LoCo diagnostic. As shown in S09 and S11, how anomalies and/or errors in the surface fluxes computed by a particular LSM–PBL coupling are then translated into the atmospheric water and energy cycle can then be quantified using this approach.

The third LoCo diagnostic that has been developed is the relationship between mean EF and PBL height (PBLH), which serves as a bulk integrative measure of the state of the land surface and the PBL. How EF and PBLH are sensitive to a particular LSM–PBL coupling can therefore be reflective of how extreme conditions manifest themselves in surface drying (wetting) and the corresponding response of PBL growth (decay).

The fourth and final core diagnostic to be applied in this study was presented by S11 and relates to the sensitivity of the LoCo process chain to surface conditions and ultimate response of the PBL in promoting or suppressing clouds and precipitation. This lifting condensation level (LCL) deficit, defined as the difference between actual PBL height reached and the LCL, quantifies how the coupled system responds to a particular land–PBL coupling and condition for both dry and moist processes.

In employing this array of diagnostics, S09 and S11 have shown that the spread and sensitivity in model results due to different LSM–PBL combinations can be

evaluated against observations in the LoCo context and ultimately used to pinpoint the weaknesses in the land and/or atmospheric component of the model and their inherent coupling. Overall, these diagnostics provide a pathway to study both the individual and collective factors determining LoCo [Eq. (1)], and most importantly, can be applied equally to any model and location of interest. However, the work of S09 and S11 was limited to developing and demonstrating the methodology for a few single diurnal cycles, rather than specific or thorough evaluation of the schemes themselves over longer time periods. The experiments and results that follow below are the final piece of this project, designed to be a full implementation of LoCo diagnostics in a regional modeling test bed and focused on scheme evaluation during climatologically and programmatically relevant case studies over multiday periods.

3. Model and site description

a. NU-WRF

The Advanced Research Weather Research and Forecasting Model (ARW-WRF; Michalakes et al. 2001) is a state-of-the-art mesoscale numerical weather prediction system. Derived from the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; Anthes and Warner 1978), ARW-WRF has been designated as the community model for atmospheric research and operational prediction and is ideal for high-resolution (e.g., 1 km) regional simulations on the order of 1–10 days. ARW-WRF has an Eulerian mass dynamical core and includes a wide array of radiation, microphysics, and PBL options as well as two-way nesting and variational data assimilation capabilities.

Recently, work has been performed to develop a National Aeronautics and Space Administration (NASA)-Unified WRF (NU-WRF; https://modelingguru.nasa.gov/community/atmospheric/nuwrf) modeling system at NASA's Goddard Space Flight Center (GSFC). NU-WRF is built upon the ARW-WRF model, and incorporates and unifies NASA's unique experience and capabilities by fully integrating the GSFC Land Information System (LIS; Kumar et al. 2006; Peters-Lidard et al. 2007), the WRF/Chem enabled version of the Goddard Chemistry Aerosols Radiation Transport (GOCART; Chin et al. 2000) model, GSFC radiation and microphysics schemes, and the Goddard Satellite Data Simulation Unit (SDSU; Matsui et al. 2009) into a single modeling framework. Overall, NU-WRF will provide the modeling community with an observation-driven integrated modeling system that represents aerosol, cloud, precipitation, and land processes at satellite-resolved scales (roughly 1–25 km).

The land–PBL interface is a core component of NU-WRF, and has been performed through the coupling of LIS–WRF by Kumar et al. 2008 (LIS–WRF). LIS consists of a suite of LSMs and provides a flexible and high-resolution representation of land surface physics and states, which are directly coupled to the atmosphere. The advantages of coupling LIS–WRF include the ability to spin up land surface conditions on a common grid from which to initialize the regional model, flexible and high-resolution (satellite based) soil and vegetation representation, additional choices of LSMs that continue to expand in range and complexity, and various plug-in options such as land data assimilation, parameter estimation, and uncertainty analysis.

The work of S09 and S11 has demonstrated LIS–WRF as a successful test bed for L–A interaction studies and LoCo because of its land–PBL scheme flexibility and high resolution. Since this time, there have been significant upgrades to both LIS and WRF including new functionality and LSMs in LIS and additional PBLs in WRF. The development of NU-WRF now ensures that the most recent versions of LIS (currently version 6.2) and ARW-WRF (currently version 3.3) are coupled and tested, and are used exclusively for the 2006/07 simulations described in section 3b. For this study, NU-WRF has been specified with a 5-s advection time step, Ferrier microphysics, Rapid Radiative Transfer Model (RRTM; Skamarock et al. 2005) longwave radiation, Goddard shortwave radiation, and the Monin–Obukhov surface layer scheme. The North American Regional Reanalysis (NARR) was used for atmospheric initialization and lateral boundary conditions using 3-hourly nudging. The vertical resolution of NU-WRF was specified as 43 vertical levels, with the lowest model level ~24 m above the surface, which was designed to improve resolution in the lower layers (PBL) relative to default configurations of the model.

1) LAND SURFACE MODELS

The LSMs employed in LIS for this study are the Noah LSM version 2.7.1 (Noah; Ek et al. 2003), the Community Land Model version 2.0 (CLM; Dai et al. 2003), and the Hydrology Tiled European Centre for Medium-Range Weather Forecasts (ECMWF) Scheme for Surface Exchanges over Land (HTESSEL; Balsamo et al. 2009). Each model dynamically predicts water and energy fluxes and states at the land surface, but they vary in specific parameterizations and representation of soil and vegetation properties and physics. For example, Noah and HTESSEL solve moisture and heat transport through 4 discrete soil layers while CLM solves for 10 layers. In addition, treatment of vegetation types and properties (such as height, coverage, and

density) and canopy fluxes differ between the three LSMs.

The Noah model employed in this study is version 2.7.1 and is identical to the version of Noah packaged in the original version of ARW-WRF version 2.2. Noah is used operationally by the National Centers for Environmental Prediction as the LSM for the North American Mesoscale (NAM) model and the Global Forecasting System (GFS). CLM and HTESSEL are unique to NU-WRF (i.e., not in official ARW-WRF releases), and it should be noted that CLM is an earlier version of the LSM for the National Center for Atmospheric Research (NCAR)'s coupled Community Climate System Model (CCSM; Gent et al. 2011), while HTESSEL is the LSM employed in the operational ECMWF Integrated Forecast Scheme (IFS; ECMWF 2011) for prediction and data assimilation, where the version employed here is identical to that used in the Global Land–Atmosphere Coupling Experiments (GLACE). As such, these LSMs are well supported and developed, and capture a wide range in complexity (e.g., layering and vegetation physics) and coupled application (e.g., mesoscale to global climate model).

2) PBL SCHEMES

In ARW-WRF, there are three options for PBLs in version 2.x and nine available in version 3.x. For this study, we focus on the three robust and well-tested PBLs that are typically employed over full diurnal cycles (including convective and stable conditions) rather than some of the newer schemes that are targeted for more narrow applications. The simplest of the three is the Medium-Range Forecast (MRF; Hong and Pan 1996) scheme, which is based on nonlocal K theory (Troen and Mahrt 1986) mixing in the convective PBL and where the diffusion and depth of the PBL are a function of the Richardson number (Ri_{cr}). The Yonsei University (YSU; Hong et al. 2006) scheme is based on the MRF and the nonlocal K theory implementation, but includes explicit treatment of entrainment and counter gradient fluxes. Finally, the Mellor–Yamada–Janjic (MYJ; Janjic 2001) scheme is the most complex of the three, and employs nonsingular level 2.5 turbulent kinetic energy (TKE) closure (from Mellor and Yamada 1982) with local K vertical mixing. In the MYJ scheme, the length scale is a function of TKE, buoyancy, and shear, and the PBL height is diagnosed based on TKE production. Overall, these three PBL schemes span the range in complexity (first order to TKE) and application (single column to full 3D) of those participating in the GEWEX Atmospheric Boundary Layer Study (GABLS).

To address LoCo under the NU-WRF framework, simulations were performed across the array of LSMs and PBLs described above, with each enabling a different LSM–PBL combination for a total of 9 (3×3) representations of L–A coupling. The remainder of the NU-WRF setup is identical for each. The results of each simulation are then evaluated using the LoCo diagnostic approaches of S09 and S11 (described in section 2), where the processes and feedbacks generated by each LSM–PBL can be evaluated and contrasted.

b. *Experimental design: 2006/07 extremes*

As shown by Koster et al. (2004) and others, the SGP region has been identified as a hotspot for L–A coupling in terms of the strength of interactions and feedbacks and its role as a transitional zone of soil moisture and vegetation conditions. Because of this, and the large record of observational data in this region, S09 and S11 focused on experiments conducted for the two golden days during the International H_2O Project in June 2002 (IHOP_2002; Weckworth et al. 2004), and evaluated simulations using data from the Atmospheric and Radiation Measurement test bed located in the region (ARM-SGP). The ARM-SGP region has also recently been the focus of studies on extreme conditions observed during the 2006/07 period. Significant low anomalies of clouds and precipitation in the 2006 water year (October–September) were immediately followed by conditions of high cloudiness and rainfall in 2007. The dry–wet contrast from 2006 to 2007 is unprecedented in the last century of data, with 2006 being the second driest and 2007 the seventh wettest year on record. Further details can be found in thorough observational analysis of the period performed by Dong et al. (2011). These dry–wet extremes have also been chosen as a focal point for integration projects designed by the NASA Energy and Water Cycle Study (NEWS; NEWS Science Integration Team 2007). This unique period combined with the strong nature of L–A interactions in this region make it an ideal case study to employ NU-WRF for studies of LoCo.

Based on the ARM-SGP data, the summers [June–August (JJA)] of 2006 and 2007 were analyzed to find an ideal case study for each. The 14–20 July 2006 experiment consists of a lengthy dry-down period with little synoptic disturbance in which the land was free to interact and evolve with the atmosphere on primarily local scales. The case study of 14–20 June 2007 focuses on a period with scattered precipitation every 1–2 days in portions of the ARM-SGP domain, interspersed with brief dry downs in which conditions were clear and/or cloudy.

As was performed for the IHOP_2002 experiments in S09 and S11, each of the three LSMs in LIS was run offline (uncoupled) for an approximately 4-yr period prior to the start time of the 2006 and 2007 case studies

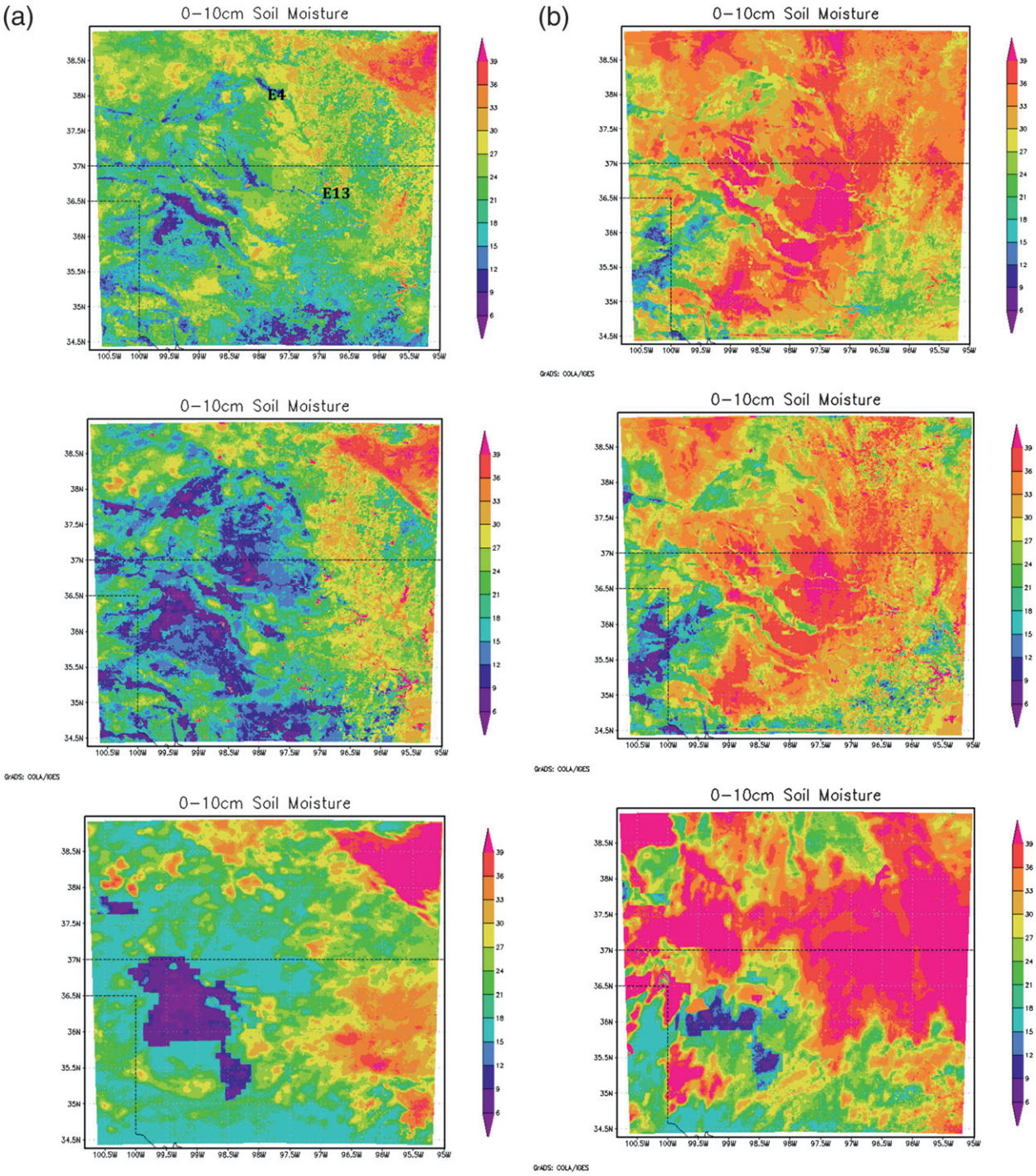


FIG. 2. (a) Soil moisture ($m^3\ m^{-3} \times 100$) in the upper 0–10-cm layer valid at 0000 UTC on 14 Jul 2006 as simulated from a 3.5-yr spinup of the (top) Noah, (middle) CLM, and (bottom) HTESSEL models over the 1-km LIS–WRF domain in the SGP. The ARM-SGP central facility (CF) at Lamont, OK, is also shown. (b) As in (a), but valid at 0000 UTC on 14 Jun 2007.

to create equilibrated, or ''spun-up,'' land surface states for initialization of LIS–WRF. Forcing data from the North American Land Data Assimilation System (NLDAS-2) project were used to drive each of the offline LSM runs. Using these resultant spun-up surface fields as initial surface conditions for the 2006/07 case studies, NU-WRF simulations were then performed over a single, high-resolution domain (500 × 500; see Fig. 2), centered

over Oklahoma and Kansas with a horizontal resolution of 1 km and time step of 5 s. The remainder of the model configuration was then ensured to be consistent with that of the experiments performed by S09 and S11.

Figure 2 shows the upper layer (0–10 cm) soil moisture values over the ARM-SGP domain as generated by the spinups for all three LSMs valid at 0000 UTC on 14 July 2006 and 14 June 2007. The advantages of using LIS for this purpose are evident in the high spatial resolution seen in Fig. 2 as a reflection of the inputs of vegetation and soil properties. In addition, differences in particular LSM parameters can be seen such as the coarser native soil texture information in HTESSEL relative to that used in Noah and CLM.

It should also be emphasized that the soil moisture used to initialize coupled NU-WRF simulations is taken from these identically forced spinup runs rather than treated as a specified variable or boundary condition, and therefore is different for each LSM (as shown in Fig. 2). As a result, variability in the coupled results across LSMs is not strictly due to differences created during the coupled simulation, and includes the influence of varying initial soil moisture (and temperature) fields from the spinup runs. We have chosen this approach, rather than controlling for uniform initial soil moisture, for example, in order to consider the performance of each LSM end to end, in the sense that each spins up its own unique soil moisture and temperature states, and then continues to run with consistent configuration (and equilibrated states through spinup) with the coupled simulation.

Overall, Fig. 2 shows soil moisture in the ARM-SGP region varies significantly from dry and heterogeneous (generally <25% volumetric) in 2006 to extremely wet (near saturation) and more homogeneous in 2007. It should be noted that in terms of spanning the range of extremes the spinup results indicate a hydrological condition that is more extreme in the wet year, while 2006 is a below normal but not entirely desiccated regime. Implications of these relative extremes will be discussed as they arise in sections 4 and 5.

c. *Data and evaluation*

The ARM-SGP program provides a wealth of surface flux, meteorological, and hydrological observations along with atmospheric profiles from radiosonde and lidar for a network of sites in and near the winter wheat belts of Oklahoma and Kansas. This includes collocated soil moisture, net radiation, and sensible, latent, and soil heat, along with collocated surface meteorology data that provide the full set of variables needed to calculate the LoCo diagnostics discussed in section 2 and evaluate against model results. In addition, during the summer of 2007 the Cloud and Land Surface Interaction Campaign (CLASIC; http://acrf-campaign.arm.gov/clasic/) field campaign took place within the ARM-SGP domain, and has provided additional sites and data for this period for evaluation purposes.

The core evaluation of these simulations in terms of the surface energy balance components are carried out for the first time employing the Land Surface Verification Toolkit (LVT; Kumar et al. 2012). LVT provides a standardized platform for intercomparing model output (from LIS or other sources) with observations and offers a range of statistical and benchmarking approaches. For these experiments, ARM-SGP data were collected from 24 sites in the domain that measure surface fluxes using eddy correlation (ECOR) and Bowen ratio (EBBR) towers, along with the collocated surface meteorology, gravimetric soil moisture probes, and, where available, radiosonde profile data.

4. Results

To determine the accuracy and impact of land–PBL coupling during the 2006 and 2007 case studies, the analysis is broken down into three components: 1) evaluation of the surface fluxes, 2) application of LoCo diagnostics, and 3) large-scale model intercomparison.

a. *Land surface energy balance*

Surface turbulent fluxes of sensible (Qh) and latent heat (Qle) serve as the principal communication and transport of heat and moisture between the land and atmosphere. In coupled models, they also provide the lower boundary condition, and from an atmospheric perspective represent the only variables of physical interest and impact from the LSM. As a result, the accuracy and sensitivity of surface fluxes simulated by each LSM–PBL coupling is of first-order importance in ultimately assessing LoCo (section 4b).

1) DOMAIN-AVERAGE FLUXES

Domain-average root-mean-square error (RMSE) and bias statistics of Qh, Qle, and soil heat flux (Qg) for each coupled simulation were calculated using LVT. Specifically, each of the 24 ARM-SGP sites was evaluated against the nearest NU-WRF 1-km grid cell at each observation time step (30 min) over the full 7-day period of each case study.

(i) Dry regime

Overall, the 2006 results in Fig. 3a show that large RMSEs in Qle (>60 W m^{-2}), Qh (>50 W m^{-2}), and Qg (>40 W m^{-2}) exist in all LSM–PBL combinations. The confidence intervals of the observations are 28.8,

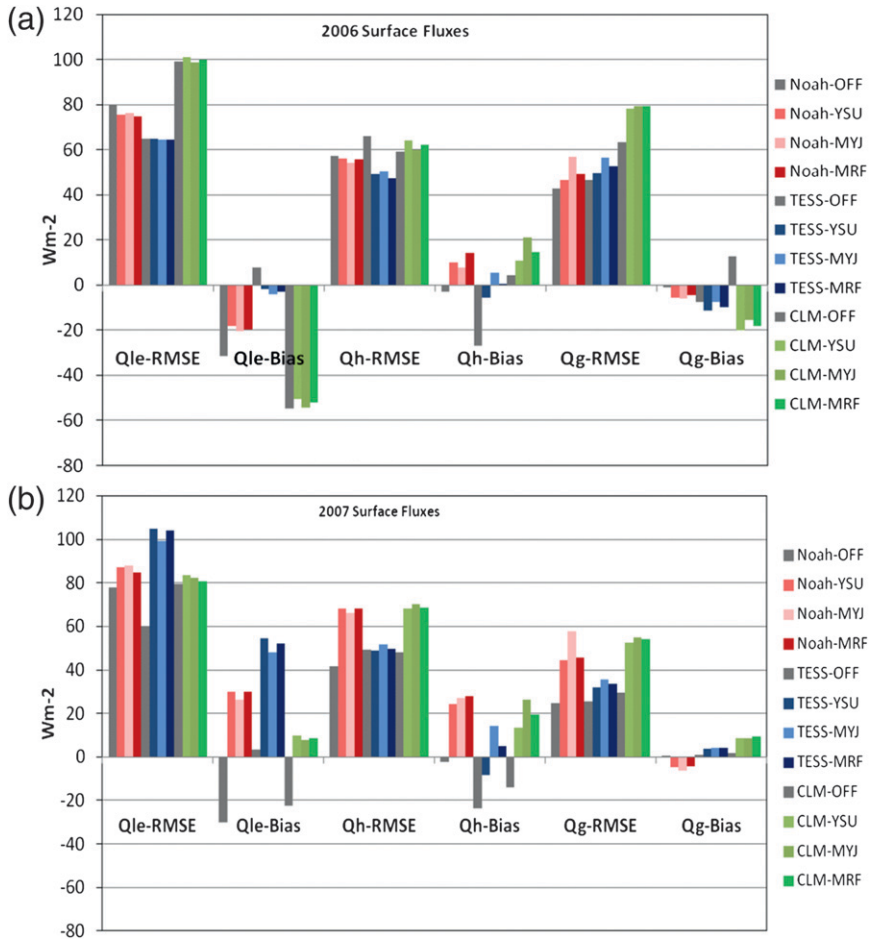


FIG. 3. Domain-average RMSE and bias statistics for the land surface fluxes of latent (Qle), sensible (Qh), and ground (Qg) heat as simulated by the LSM–PBL couplings in LIS–WRF over the (a) 14–20 Jul 2006 and (b) 14–20 Jun 2007 periods. Also shown are the offline simulations of each LSM spinup (OFF) continued through the periods driven by NLDAS-2 atmospheric forcing.

19.7, and 2.3 W m^{-2} for Qle, Qh, and Qg, respectively. Large biases also are present and indicate that the Bowen ratio (evaporative fraction) is overestimated (underestimated) by all the runs. Overall, the differences between LSMs are significant (at the 95% confidence interval), where CLM performs worst in terms of RMSE and bias, while HTESSEL simulates the surface energy balance best and is notably unbiased in all three flux components.

In addition, statistics were computed for fluxes simulated by each LSM run in offline mode during the 7-day case studies (i.e., continuation of the spinup runs), using best-available atmospheric forcing from NLDAS-2. When compared against the coupled runs, these results show that running NU-WRF with Noah and HTESSEL (regardless of PBL scheme) tends to improve the fluxes of Qh and Qle versus running them with prescribed forcing offline. Qg, on the other hand, shows slight degradation in all coupled runs. It should be noted that the typical magnitude of Qg is much less than that of Qh and Qle during the daytime, so the errors in Qg seen here are quite large in terms of the proportion of their daily average—the implications of which will be discussed in the next section.

The relative sensitivity of surface flux errors to the choice of LSM versus PBL scheme can also be discerned from this analysis. Clearly, during the dry conditions of 2006 it is the choice of LSM that is critical as the spread across PBL scheme choices given the same LSM is negligible. This is not an unexpected result given that the LSMs control the calculation of surface fluxes, but the degree to which the PBL scheme modulates the atmospheric feedback is important to quantify and in the case of the dry regime appears to be quite minimal.

(ii) Wet regime

In contrast, results from 2007 (Fig. 3b) show a different hierarchy of LSM performance, stratification, and coupling impact. The confidence intervals of the observations are 15.0, 11.0, and 1.7 W m^{-2} for Qle, Qh, and Qg, respectively. HTESSEL performs worst for Qle, significantly overestimating evaporation and nearly doubling the RMSE seen during the dry regime. This corresponds to a very wet soil moisture condition (not shown) that is near saturation and thereby able to sustain a freely evaporating surface that responds directly to the net radiation at the surface. CLM produces the most accurate and least biased Qle, but is worst in terms of Qg. Once again, Noah is in the middle in terms of the energy balance accuracy relative to the other two LSMs. Soil heat flux Qg is relatively unbiased in all runs in both 2006 and 2007 because of the diurnal cycle of Qg (positive during the day, negative at night), which balances out overall. The overall sensitivity of surface fluxes to the choice of PBL scheme is somewhat higher in this wet regime, but from a solely land surface energy balance perspective it remains the choice of LSM that is of first-order importance once again. How these sensitivities then stratify in terms of the PBL response will be examined in a later section evaluating LoCo diagnostics.

The biases in 2007 also indicate that there is an excess of net radiation, as the cumulative bias of the three flux components is largely positive, and because this is a wet period much of that extra energy goes toward Qle. This will be investigated further in the next section. Also contrary to 2006, coupled simulations using NU-WRF versus offline LSM runs using NLDAS-2 result in significantly worse RMSE and bias statistics (e.g., Qle for HTESSEL), with the sign of the bias often reversing as well (e.g., Qle for Noah and CLM). This suggests that the potential impact of running a fully coupled model versus prescribing best-available atmospheric forcing to an offline LSM is much larger in the wet regime. In particular, it is the differences in simulated versus observed cloud cover and the impact on incoming radiation at the surface that are the major factors.

Along these lines, additional NU-WRF simulations of the 2007 case (not shown) using the Noah LSM showed considerable insensitivity of the surface energy balance components to the soil type (texture) map used for the ARM-SGP domain or to the Noah ''Czil'' parameter used in flux computations. That the impact of using spatially constant or unrealistic values of soil texture (and in turn associated hydraulic properties) is diminished during the wet extreme period supports that this is an atmosphere-limited regime of evaporation, and that PBL dynamics should play a larger role than details in the soil (already near saturation) scheme. In contrast, during the dry case and soil-limited regime, proper soil type specification results in improvement in fluxes on the order of 10–30 W m^{-2} overall.

2) MEAN DIURNAL CYCLES

While the cumulative averages over the domain and 7-day periods in Fig. 3 provide an assessment of the bulk behavior of the different LSMs and their sensitivities to different atmospheric components, it is also instructive to examine the diurnal cycles of the flux components both overall and at individual sites.

(i) Dry regime

Figure 4 shows the mean diurnal cycles (7-day averages) of Qle, Qh, Qg, and net radiation (Rn) for different LSM–PBL couplings during the 2006 period. The domain-average Noah + YSU results (Fig. 4a) show that Rn is simulated well, but that the Bowen ratio is overestimated (Qh large, Qle small) along with Qg. This is likely due to soil moisture being too low and unable to produce the evaporation observed. The offline Noah run (Fig. 4b) shows that Rn is ~100 W m^{-2} less than observed during the daytime, which translates into even lower Qle flux than when coupled (as confirmed by Fig. 3a). The reduced energy, incidentally, reduces Qh as well to match closely with observations. As discussed earlier, Qg is too large (positive) during the day and vice versa at night, leading to a small net bias over the full cycle.

The diurnal cycles for HTESSEL (Fig. 4c) confirm that Qh and Qle are simulated quite closely to observations over the domain average. Net radiation is slightly overestimated, and is primarily reflected in too-large Qg during the daytime. When looking at individual sites for the Noah + YSU and HTESSEL + MYJ runs (Figs. 4d,e), net radiation is simulated very well compared to observations. Once again, Noah produces a Bowen ratio that is too high, while HTESSEL simulates evaporation quite well and more of the flux error is seen in Qg. It is also evident that each of the LSMs produces a phase error in daytime Qg, with an earlier peak than observed. This is likely due to the differences in observed (0–5 cm based on heat flow plates) versus modeled (0–10-cm temperature gradient and parameterized approach) Qg estimation, but should not be disregarded as having negligible impact on coupling, as will be discussed later.

(ii) Wet regime

Similarly, focusing on the diurnal cycles during the wet regime (Fig. 5) yields insight on how well each LSM partitions the incoming energy into fluxes that ultimately drive the coupling behavior. The coupled runs confirm a very large overestimation (~150–250 W m^{-2} daytime

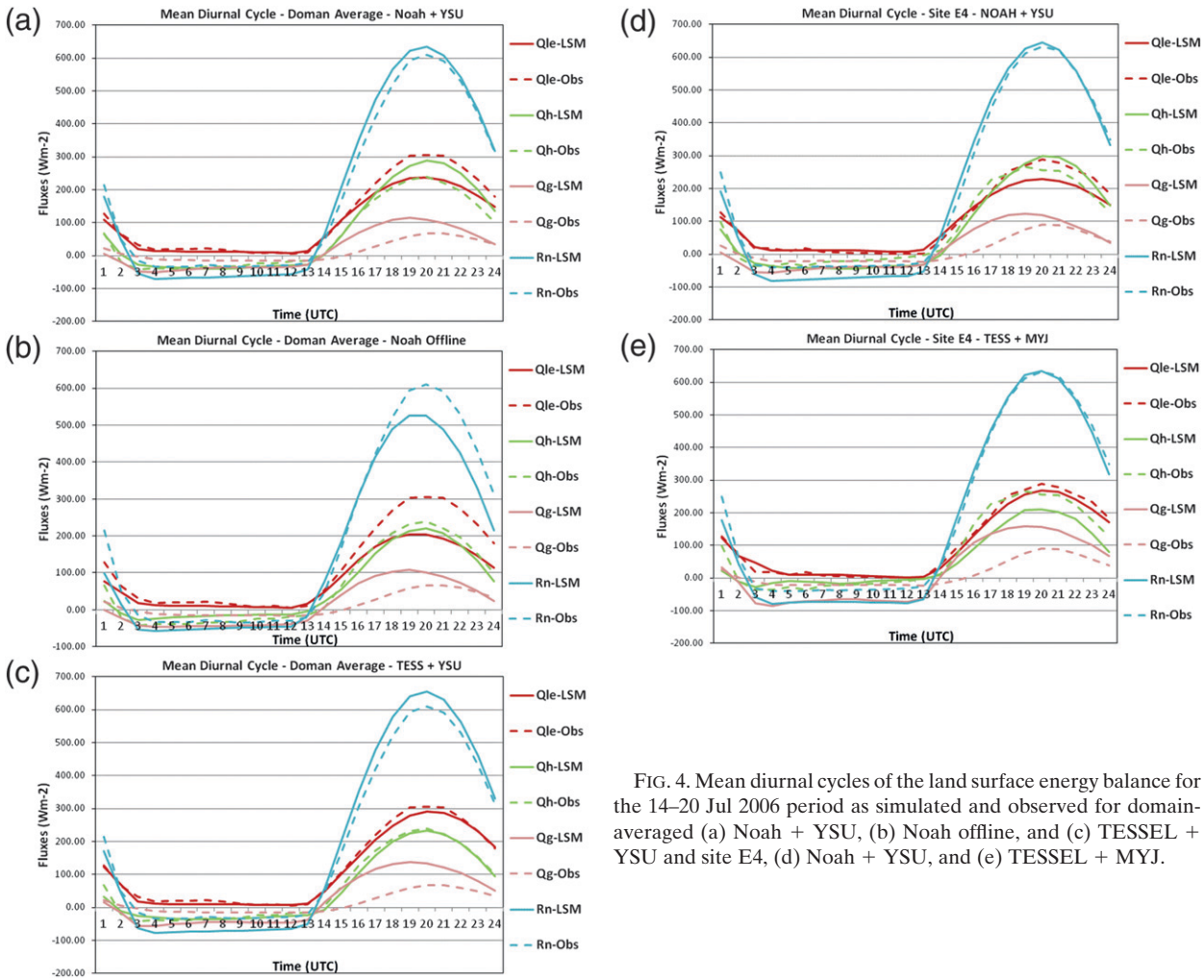


FIG. 4. Mean diurnal cycles of the land surface energy balance for the 14–20 Jul 2006 period as simulated and observed for domain-averaged (a) Noah + YSU, (b) Noah offline, and (c) TESSEL + YSU and site E4, (d) Noah + YSU, and (e) TESSEL + MYJ.

peak) of Rn versus that observed, both in domain average (Figs. 5a,c) and at individual stations (Figs. 5d,e) and regardless of LSM or PBL choice (MYJ and MRF as well). In the Noah runs (Figs. 5a,d), this extra energy goes largely to Qle and Qh, which are in turn overestimated relative to observations. CLM, on the other hand, buries much of the extra energy in the soil heat flux and as a result, simulates Qle quite close to observations. This has major implications for the accuracy and nature of the coupling, in that the atmosphere ultimately cares only about the land boundary condition of Qle and Qh.

As for the dry regime, the offline Noah run underestimates Rn in significant contrast to the coupled run overestimates. The reduced energy does not allow for evaporation to match that observed, despite the decent simulation of Qh and Qg (though these terms are quite small relative to Qle). This comparison of offline (good forcing) versus coupled (NU-WRF) net radiation indicates big differences in the simulated cloud field. When traced back to the source, it is the downward shortwave radiation that causes this disparity, where the offline case (NLDAS-2) reflects more substantial cloud fields and limited radiation compared to the coupled runs where NU-WRF is allowed to freely evolve over the 7-day period. These Rn errors are rather instructive from a LoCo perspective, as models often contain biases in clouds, precipitation, and radiation that ultimately impact and feedback upon the surface condition, fluxes, and PBL evolution.

Focusing on individual sites (Figs. 5d,e) again yields insight as to how each LSM handles this particular wet regime. Noah + YSU at site E13 shows the extra Rn spread out amongst all three surface fluxes but weighted more toward Qh and Qg, thereby producing an evaporative flux that is only slightly overestimated. HTESSEL + YSU exhibits the opposite effect of too much radiation. Because of its nearly saturated soil (as discussed above), HTESSEL produces evaporative fluxes that are extremely high, virtually matching the atmospheric

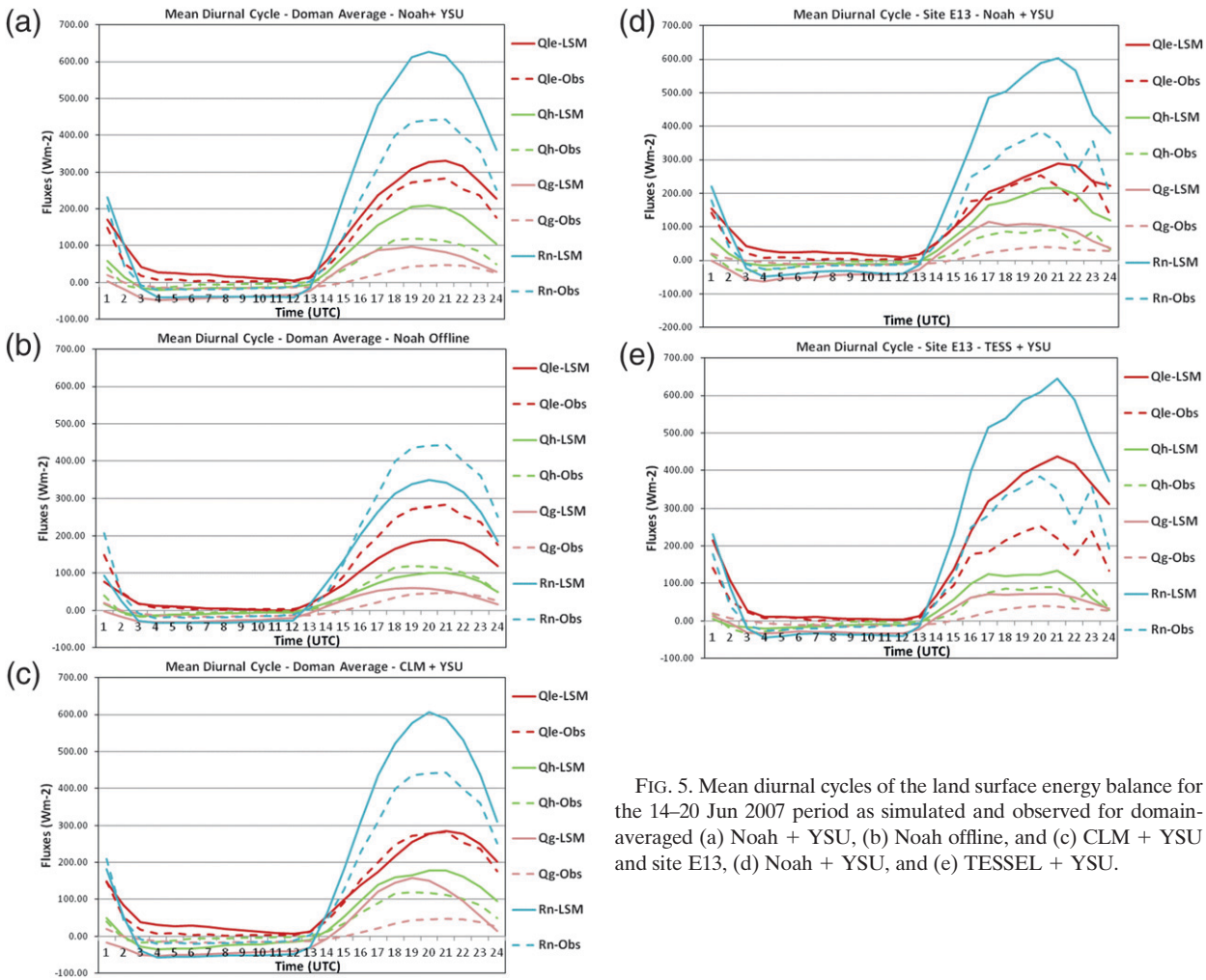


FIG. 5. Mean diurnal cycles of the land surface energy balance for the 14–20 Jun 2007 period as simulated and observed for domain-averaged (a) Noah + YSU, (b) Noah offline, and (c) CLM + YSU and site E13, (d) Noah + YSU, and (e) TESSEL + YSU.

demand that is very high in this case. CLM at individual sites (not shown) is consistent with its domain average, and buries much of the extra Rn in Qg, thereby allowing for the best simulation of the diurnal cycle of EF.

b. Application of LoCo diagnostics

The analysis presented above provides an accounting for how and why the surface fluxes that constitute the lower boundary condition to the atmosphere (i.e., PBL) behave versus observations during dry and wet extremes. How these fluxes translate through the coupled system in terms of T, q, PBL development, thermodynamics, and clouds (e.g., LCL deficit) can then be understood in the context of the LoCo diagnostics.

1) MIXING DIAGRAMS

(i) Dry regime

The behavior of coupled heat and moisture states and fluxes can be captured simultaneously using the mixing diagram approach as presented in S09 and S11. Figure 6 presents composite diagrams of the seven daytime periods of the 2006 period for each LSM coupled to the three PBL schemes and evaluated against observations at the E4 site. From the coevolution of T and q, it is evident that Noah is too warm and dry overall, CLM is significantly too warm, and HTESSEL is closest to observations. This follows with the surface flux analysis (as do the surface Bowen ratio vectors) of the previous section, but these results also show more spread in T2m, Q2m, and fluxes due to the choice of PBL scheme than were evident from the surface analysis alone. The MYJ scheme performs the best relative to the YSU and MRF for all three LSMs, and in particular when coupled with HTESSEL, and the dry air and moist entrainment ratios (A_h, A_{le}) generally follow suit.

The implications for different LSM–PBL coupling are also evident in the thermodynamics (overlain contours). The moist static energy (i.e., equivalent potential temperature; θ_e) is simulated quite well in all runs, with little

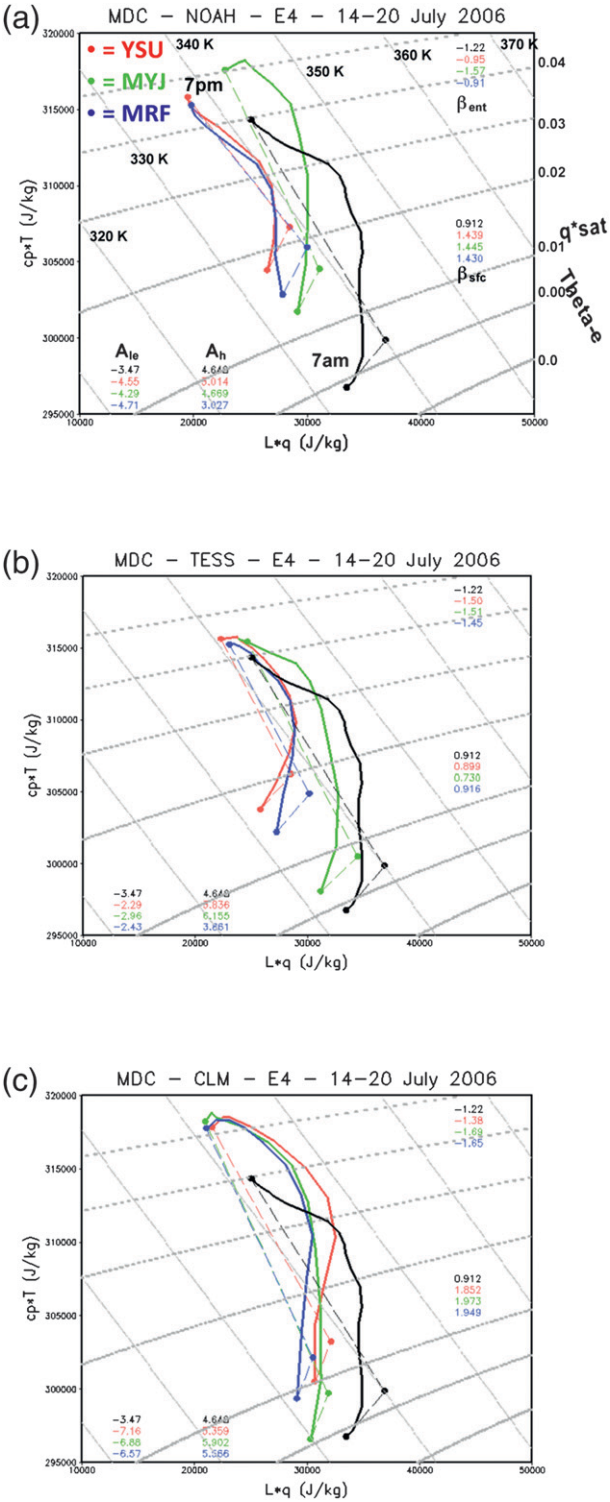


FIG. 6. Mixing diagrams from the LIS–WRF simulations of (a) Noah, (b) TESSEL, and (c) CLM composited over the 7-day case studies of 2006 (14–20 Jul) and 2007 (14–20 Jun), and overlain with lines of constant equivalent potential temperature (θ_e) and saturation humidity (q^*_{sat}) deficit.

diurnal increase during the dry regime. The PBL saturation deficit (q^*_{sat}), on the other hand, differs substantially between runs, and in particular is overestimated by Noah and CLM, indicating a daily PBL that is extremely dry. This type of deficit tends to support dry regimes, as the drier PBL raises atmospheric demand and ensures that the maximum evaporation (given the soil moisture condition) is reached. The diurnal cycles of q^*_{sat} as evaluated in this manner could serve as a useful metric in terms of evaluating whether a particular model or scheme is supporting a dry (or wet) regime, how it relates to observations, and ultimately what is driving the differences (in this case depleted soil moisture, low evaporation, and large PBL growth and entrainment). There is also the potential to identify a threshold of q^*_{sat} that once reached, makes it difficult to transition out of this dry regime and feedback loop.

To synthesize the information content of these diagrams, statistics can be generated based on the differences in the modeled versus observed heat and moisture states. Table 1 shows the RMSE and bias metrics for each LSM–PBL pair as calculated from the cumulate differences in T2m and Q2m over the daytime diurnal cycles in Fig. 6. The values confirm that the largest cumulative errors are seen in the Noah and CLM simulations, regardless of PBL choice. However, there is a distinct advantage to using the MYJ PBL that produces the lowest RMSE and biases for each of the LSMs. These metrics also indicate where the largest errors of the coupled system are manifested in terms of the heat (T2m) versus moisture (Q2m), where Noah tends to overestimate the drying while CLM overestimates the heating in the PBL.

An advantage of transforming the mixing analysis into energy space units (in addition to allowing comparable flux computations) is that combined metrics of the mean absolute error (MAE) and RMSE error can be established as follows:

$$\text{Total RMSE} = \text{RMSE(T2m)} + \text{RMSE(Q2m)} \quad \text{and} \tag{3a}$$

$$\text{Total MAE} = \text{MAE(T2m)} + \text{MAE(Q2m)}, \tag{3b}$$

which summarize the cumulative heat and moisture error in the coupled land–PBL system over the course of the daytime cycle. In particular, the values in Table 1 confirm that 1) the MYJ produces the best coupling for all LSMs and 2) HTESSEL generally outperforms the other LSM–PBL couplings. As a result, it can be concluded that the best-simulated surface fluxes (Qh and Qle) from HTESSEL do, in fact, translate into better LoCo components relative to Noah and CLM. From

TABLE 1. Statistics (J kg^{-1}) based on evolution of T2m and Q2m vs observations, derived from the mixing diagrams in Fig. 6. Boldface values indicate the smallest magnitude error produced by each LSM and associated PBL scheme.

	Noah + YSU	Noah + MYJ	Noah + MRF	TESS + YSU	TESS + MYJ	TESS + MRF	CLM + YSU	CLM + MYJ	CLM + MRF
RMSE T2	4286.08	4955.41	**3690.39**	4033.14	**2141.18**	3467.70	4821.05	**4238.01**	4705.49
RMSE Q2	7676.35	**4010.32**	7541.49	5374.09	**2260.11**	5095.88	**3328.06**	4118.45	4494.36
BIAS T2	3679.64	4909.45	**3108.82**	3611.82	**2076.45**	3082.27	4777.64	**3898.18**	4628.91
BIAS Q2	−7573.25	**−3809.71**	−7386.57	−4993.44	**−2137.12**	−4763.65	**−3239.87**	−4075.62	−4432.84
Total RMSE	11 962.43	**8965.73**	11 231.87	9407.22	**4401.29**	8563.57	**8149.11**	8356.46	9199.85
Total MAE	11 252.89	**8719.16**	10 495.39	8605.25	**4213.57**	7845.91	8017.51	**7973.80**	9061.74

a physical standpoint, significant excess (e.g., CLM + YSU) or deficient (e.g., Noah + MRF) energy in the system has implications for the evolution of thermodynamics (θ_e, RH), PBL evolution (e.g., q^*_{sat}, LCL), and ultimately clouds and precipitation that will become more evident when examining the wet regime.

In terms of the MYJ performance, this can be traced directly to its superior performance in the stable (nighttime) regime and PBL mixing, during which more accurate T2m and Q2m cycles are simulated. This agrees with previous results regarding TKE versus nonlocal schemes during stable conditions (Shin and Hong 2011; Steeneveld et al. 2008), as the YSU/MRF formulations for nighttime mixing result in a dampening in the amplitude of the diurnal cycle. This results in improved morning heat and moisture states in MYJ, which then allows for (but does not ensure) a better daytime diurnal cycle of both fluxes and states. HTESSEL + MYJ is an example of this, whereas CLM + MYJ is an example of an improved initial state not resulting in better LoCo during the day as a function of the specific interaction of the land and PBL schemes.

(ii) Wet regime

The mixing diagrams for the 2007 simulations (Fig. 7) show a distinctly different signature of states and fluxes for the wet regime. Although there appears only small sensitivity to choice of PBL scheme for each of the LSMs, the daytime diurnal cycle is considerably dampened (T2m, Q2m, fluxes, and PBL height) such that the relative spread is still significant. Noah and HTESSEL exhibit somewhat similar patterns that are close to observations in terms of T2m but underestimating Q2m, with a significant impact of the MYJ scheme on both. In the case of Noah, the MYJ schemes produces too large an increase in heat and moisture in the system, while the MYJ scheme coupled to HTESSEL nudges the moisture condition closer to observed. Following from the surface energy balance results, HTESSEL has a much too large evaporative flux at the surface, but the response of the entrainment fluxes and ratios is not significantly degraded as a result. CLM shows more extreme heating and moistening of the system despite having initial states that are closer to observed than for Noah or HTESSEL.

The metrics in Table 2 also support the relative impacts of the LSM and PBL schemes for this wet regime. There are quite large RMSE and bias values for Q2m regardless of scheme choice, but a sign in the bias that is dependent on both the LSM and PBL choice. Overall, there is more spread due to PBL choice for the wet regime than for the dry regime. In terms of MAE, MYJ again outperforms the other PBL schemes regardless of LSM choice, and HTESSEL produces the lowest errors regardless of PBL scheme. For all LSM–PBL pairings, the magnitudes of RMSE and MAE are much larger during the dry regime (as expected given the larger diurnal amplitude in T2m and Q2m). However, there is substantially greater variance in the (hourly) errors during the wet regime as evidenced by the differences in RMSE and MAE for each.

It should be noted that the large overestimation in Rn in 2007 was not evenly reflected in the coupled diagnostics. Interestingly, CLM was identified as having the best Qle and HTESSEL the worst (overestimated), but here it is evident that HTESSEL actually underestimates the moisture bias and CLM vice versa. This supports the idea that the PBL plays a much larger role during the wet regime, and that the magnitude (and even sign) of a bias in the LSM can be outweighed by the atmospheric component.

2) PBL BUDGETS

Following the approach of S09 and S11, the full PBL budgets of heat and moisture can be derived directly from the mixing diagram analysis. Figure 8 shows how each LSM–PBL coupling performs relative to each other and observations for the surface, entrainment, and total fluxes of Qle and Qh for the sites and composites shown in Figs. 6–7. It is immediately evident how the LSM choice is dominant for surface flux partitioning and magnitude, and also how the impact of that choice is felt through the PBL and total fluxes of the coupled system. In 2006, the LSMs all underestimated the available energy (Rn − Qg), primarily as a result of overestimation

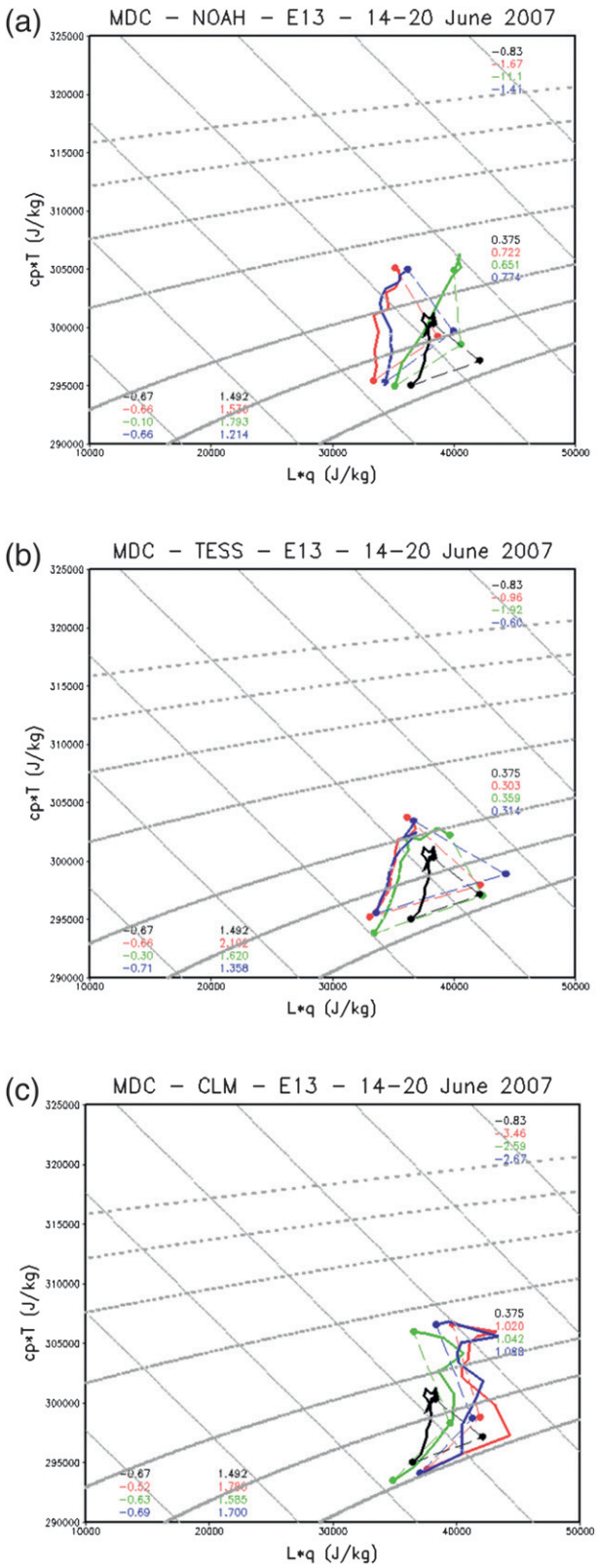


FIG. 7. As in Fig. 6, but for the 14–20 Jun 2007 case.

of Qg (as seen in the diurnal cycle analysis), with CLM performing worst. The stratification by LSM remains strong in the entrainment and total fluxes as well, while the MYJ scheme for each LSM is noticeably closer to observations than the other PBL schemes. This result is largely consistent with statistics presented in Tables 1 and 2, which is important to consider for LoCo as the PBL budget results are based on the flux components while the MAE values are based on the T2m and Q2m evolution (i.e., states).

In the wet regime, there is a noticeable shift in all components of the PBL budget toward higher Qle and Qh, as expected. The Rn overestimation is seen in the surface fluxes, with HTESSEL significantly exceeding the observed Qle. The extra energy then is propagated to the coupled system such that the entrainment and total fluxes are overestimated, particularly in terms of Qh. There is also much more spread both within and across LSMs than in the dry case, indicating (and supported by earlier analyses) that the choice of PBL is more important and at least on par with the choice of LSM during wet regimes. Overall, HTESSEL still produces the best entrainment and total heat and moisture budgets, despite the large bias in surface evaporation.

3) EVAPORATIVE FRACTION VERSUS PBL HEIGHT

A third diagnostic of the LoCo behavior is the relationship of EF to PBLH, which serves as integrated measures of the land and PBL condition, respectively. As in S09 and S11, PBLH was postprocessed from NU-WRF output using a bulk Richardson number approach so as to ensure consistency across PBL schemes (as opposed to default internal calculations). In Fig. 9, the overall shift from dry to wet regime is as expected in terms of lower PBLH and higher EF. It is the spread due to LSM and PBL choice, again, that is of interest, and shows larger spread in PBLH due to PBL choice in 2006, and likewise for EF due to the LSM choice in 2007. PBLH is insensitive to the choice of LSM or PBL scheme in the wet regime because of the limited PBL growth, 7-day averaging, and very low day–day standard deviation.

Observations show that HTESSEL performs best in both the dry and wet years, producing nearly exact PBLH and EF values for each. This is consistent with the results of the mixing diagrams and PBL budgets above. From a land surface perspective, however, it is clear that this may be the right answer for the wrong reasons and highlights the importance of not simply focusing on individual metrics of the land or PBL in order to assess scheme coupling and deficiencies. For example, that HTESSEL is close to the mean observed EF (and Bowen ratio) in both years masks out the fact that a very large Rn leads to overestimation of Qle (and Qh). In addition,

TABLE 2. As in Table 1, but for the 14–20 Jun 2007 case.

	Noah + YSU	Noah + MYJ	Noah + MRF	TESS + YSU	TESS + MYJ	TESS + MRF	CLM + YSU	CLM + MYJ	CLM + MRF
RMSE T2	2693.39	4042.83	**2570.96**	1735.03	**1466.35**	1597.78	3613.76	**3297.83**	3804.91
RMSE Q2	3391.67	**1888.69**	2993.62	2521.81	**1952.25**	2445.54	3823.87	**1543.29**	2975.94
BIAS T2	2160.09	3518.09	**2093.55**	1205.27	**747.36**	1158.88	2666.09	**1910.55**	2889.18
BIAS Q2	−3327.33	**1149.51**	−2943.65	−2412.52	**−1295.72**	−2299.12	3499.8	**494.55**	2592.73
Total RMSE	6084.94	5931.53	**5564.60**	4256.88	**3418.57**	4043.28	7437.55	**4841.05**	6780.95
Total MAE	5487.46	**4667.60**	5037.19	3617.85	**2043.41**	3457.93	6165.89	**2405.10**	5481.90

PBLH as a metric may actually integrate LoCo processes spatially and temporally to the degree that the process-level dynamics are masked out. This is something to keep in mind for future observing systems and model evaluations employing PBLH or EF alone as surrogates for coupled processes.

4) LCL DEFICIT

It is therefore useful to return to the mixing diagram thermodynamic properties and their relation to PBL evolution during the diurnal cycle. S11 defined the LCL deficit as the difference between the actual PBLH and the level of the LCL, and is shown here (Fig. 10) for the two case studies and sites that were presented in Figs. 6 and 7. As expected, 2006 is sufficiently dry such that the PBL never reaches the LCL, and there is stronger grouping of the LCL deficit due to LSM choice (particularly in the afternoon). As was concluded by S09, however, it is important to recognize that each LSM–PBL coupling does have an impact on the coupled system and LCL deficit regardless of whether or not the LCL is reached and moist processes take over. This effect is missed by integrated studies examining only soil moisture and precipitation.

The wet regime in 2007 shows a considerable shift toward negative values overall, and much greater spread throughout the day due to the choice of PBL as well as LSM. This is again consistent with the analyses presented earlier. Noah struggles to reach the LCL, while CLM and HTESSEL produce PBL growth that exceeds the LCL for a good part of the afternoon and are more in line with observations at 2245 h. Spatial analysis of cloud liquid water (not shown) confirm that clouds were produced and sustained in this portion of the domain for HTESSEL and CLM on this afternoon, while Noah remained clear. Overall this diagnostic approach is another step in quantifying how the choice of LSM can impact not only the surface fluxes and condition, but also can propagate through the PBL and support cloud formation.

c. *LoCo representation in large-scale models*

Recent NEWS and other community-based efforts (e.g., GEWEX's Landflux) have shown that current data and model products have significant uncertainty and spread in surface flux and other water and energy budget terms across global, continental, and regional scales. Although limited in scope, the LoCo diagnostic approach using the NU-WRF test bed has been shown here to be useful and essential toward understanding scheme behavior and coupling. It is therefore hoped that by applying LoCo diagnostics to community products and models at coarser and global scales, improvements can be made in the proper translation of land surface states and anomalies (e.g., flood/drought) into atmospheric quantities (e.g., afternoon convection).

As a first look, we will build upon the in-depth analysis of reanalysis products evaluated against in situ observations from the ARM-SGP site that have been performed as part of the NEWS study by Kennedy et al. (2011). In addition to demonstrating that LoCo diagnostics can be applied to large-scale models, this section aims to characterize the coupling and its variability in reanalysis data over a semiseasonal scale, rather than just the one week that overlaps with the NU-WRF case studies. This better demonstrates how these products and their sensitivities in coupled components evolve on longer time scales (i.e., longer than 7 days).

1) MERRA

NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011) is an assimilation system based on the Goddard Earth Observing System Data Analysis System, version 5 (GEOS-5) model that covers the period 1979–present with global coverage at $\frac{1}{2}° \times \frac{2}{3}°$ resolution. Figure 11 shows the mixing diagrams for MERRA's monthly mean diurnal cycles (June, July, and August) for 2006 and 2007, along with the observations from the 7-day case studies at site E13. MERRA performs quite well in terms of its land–PBL coupling relative to detailed in situ observations. Even at monthly mean scales, characteristics of the surface and PBL fluxes and states are captured well by MERRA during both regimes. Both the July (2006) and June (2007) results match closely with the corresponding month of the 7-day case study in the observations, indicating that the seasonal evolution

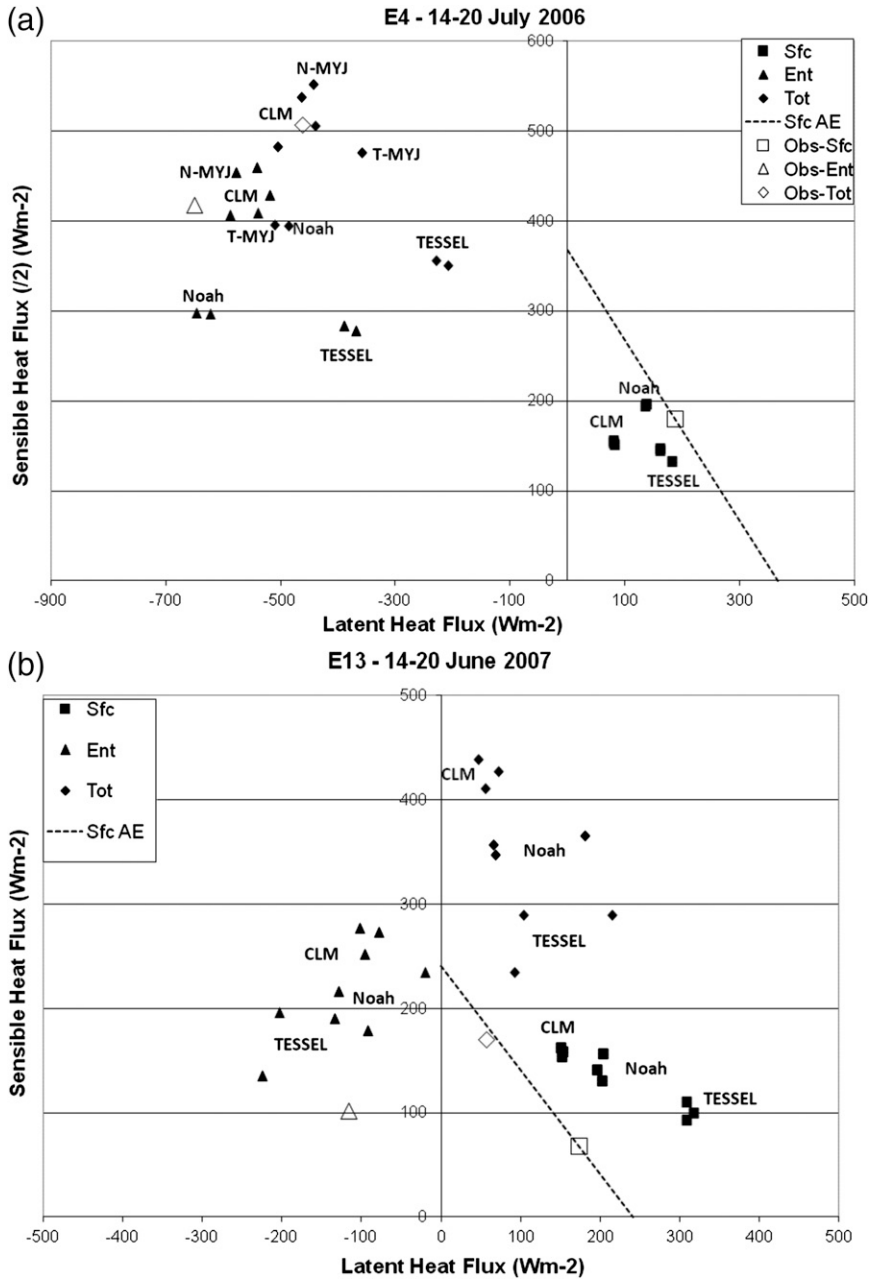


FIG. 8. Heat and moisture budgets (SFC, ENT, and TOTAL) from the NU-WRF simulations vs observed, derived from the mixing diagrams in Figs. 6 and 7 for the (a) 2006 and (b) 2007 case studies. Note that for 2006, the sensible heat flux has been scaled by 0.5 for the entrainment and total fluxes.

is also consistent. That the observations show wider diurnal range (T2m, Q2m) than MERRA during July 2006 is not surprising because the smaller averaging (spatially and temporally) of the observations is more representative of localized conditions (e.g., that produces particularly large PBL height and entrainment during a dry-down period) than the monthly mean and coarse resolution of MERRA.

2) NARR

The National Centers for Environmental Prediction (NCEP) NARR (Mesinger et al. 2006) also covers the 1979–present period, but with 32-km horizontal resolution and only over North America. Figure 12 shows the mixing diagrams for the summer monthly mean diurnal cycles for 2006 and 2007. The overall monthly patterns

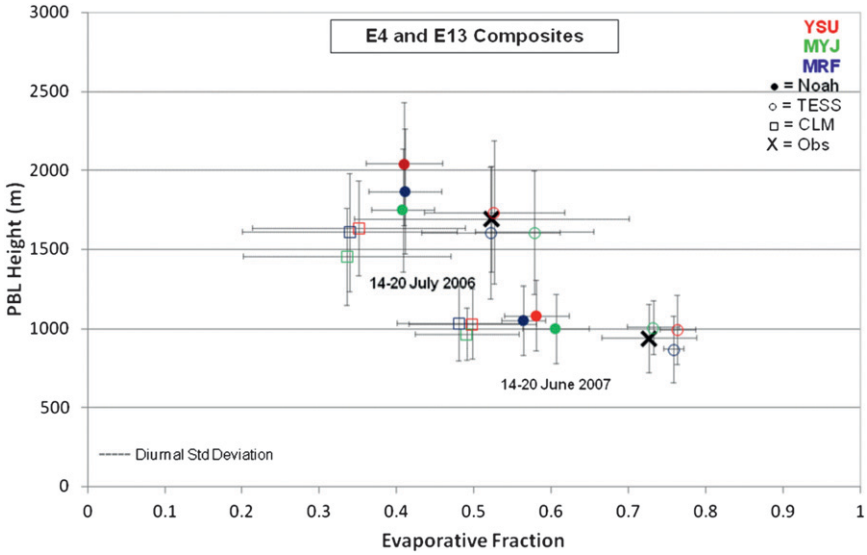


FIG. 9. Mean daytime evaporative fraction vs PBL height for each simulation vs observed, along with the diurnal standard deviation through the 7-day period.

and evolution of T2m and Q2m are similar to that seen in MERRA, but in the dry regime NARR exhibits a cooler and damped dynamic range relative to observations and MERRA. The major difference appears to be in lower PBL growth and entrainment rates (particularly dry air). In 2007, the coupled fluxes and states are remarkably similar to that seen in both MERRA and the observations (for June), including the dynamic range.

Overall, NARR produces a slightly wetter dry regime and drier wet regime than is observed or produced by MERRA. This is a moderation of extremes, in a sense, and is likely due to differences in reanalysis treatment of the large-scale averaging over monthly diurnal cycles, the representation of PBL height (and vertical levels), and the soil moisture (and evaporative sensitivity) dynamic range and drying thresholds as controlled by the respective land surface schemes (NARR–Noah LSM; MERRA–Catchment LSM).

When comparing against the high-resolution NU-WRF runs with detailed land surface initialization, we see somewhat comparable results in the reanalysis products. As expected, the severity of the extremes can be captured well by NU-WRF and lead to a larger response by the PBL, but in the case of CLM and HTESSEL the dry and wet extremes (respectively) were overestimated to a degree, which negatively impacted their LoCo components. It should also be noted that the surface Bowen ratio is overestimated by NARR (which uses the Noah LSM) in both regimes, indicating a dry bias that is also consistent with the Noah results in the NU-WRF experiments above.

Lastly, while a detailed evaluation of these model products and physics is beyond the scope here, these results do show how LoCo diagnostics can be applied across a range of scales and models to gain insight on their relative and absolute behavior in terms of land–PBL coupling components. Total energy metrics, PBL heat and moisture budgets, EF versus PBLH, and LCL deficit analyses would yield further insight into these models, and is being planned as part of a comprehensive intercomparison of models, locations, and metrics in a future LoCo study.

5. Discussion and conclusions

In this study, recent advances in diagnosing L–A coupling have been applied to a high-resolution regional modeling test bed during case studies consisting of consecutive dry and wet extreme conditions in the SGP. Results demonstrate both the accuracy and sensitivity of LSM and PBL scheme components and their coupling during these regimes, focusing on the process level and the interactions and feedbacks that constitute the land–PBL coupling [Eq. (1)].

Key findings from the land surface energy budget analysis inclued the following:

- Significant errors exist in land surface energy balance simulations that depend primarily on choice of LSM and dry/wet regime.
- In terms of evaporative fluxes, HTESSEL performs best in the dry regime and worst in the wet regime, and vice versa for CLM.
- The differences in offline versus coupled land surface fluxes are greater during the wet regime when

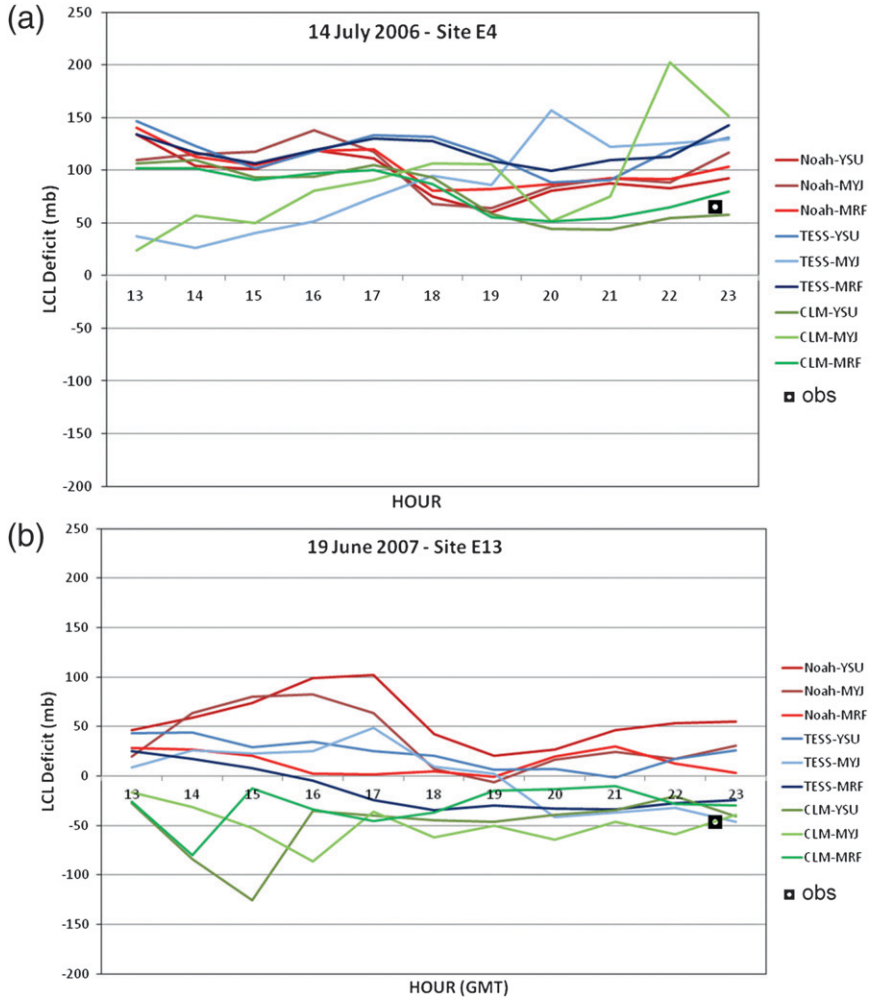


FIG. 10. Hourly LCL deficit (=PBL height − LCL; hPa) calculated for each of the PBL–LSM couplings for the (a) 2006 and (b) 2007 case studies.

simulated radiation can deviate significantly from observed forcing.

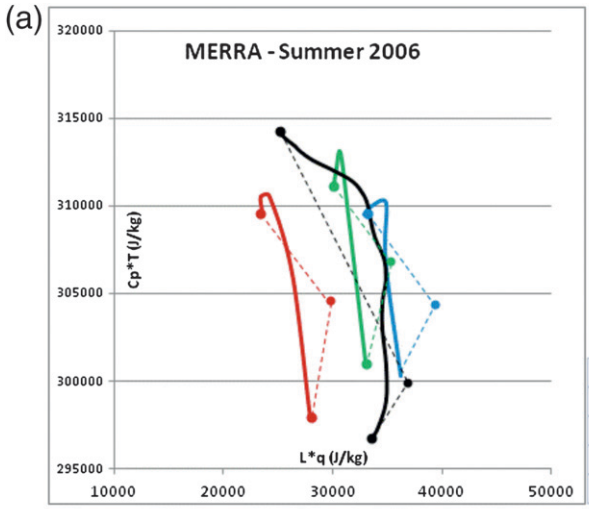
- A key factor in LoCo is the degree to which each scheme partitions energy (and input radiation biases) into the soil heat flux.

Key findings from the LoCo analysis inclued the following:

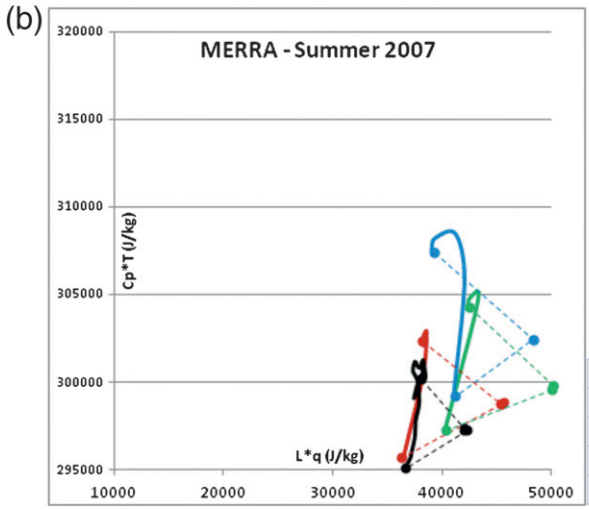
- The sensitivity of L–A coupling is stronger toward the LSM during dry conditions, while both the LSM and PBL choice are comparable during wet conditions.
- The MYJ scheme produces best MAE and heat and moisture budgets in both the dry and wet regimes.
- Overall, HTESSEL produces the best overall coupling metrics (MAE, PBL budgets, and EF/PBLH).
- Large-scale reanalysis products generally perform well in representing land–PBL coupling at monthly mean scales and are sensitive to the dry/wet regimes.

While the overestimation of net radiation in NU-WRF during the wet regime was unexpected, it confirms the importance of offline LDAS driven by observed forcing in providing the best estimates of land surface states for hydrometeorological applications. The 2007 results also highlight an important aspect of LoCo diagnosis in terms of how errors are translated between components of the modeling system. In particular, the contrast between HTESSEL and CLM allows for an interesting (and often ignored) feature to become evident regarding the soil heat flux. CLM tends to bury much of the extra Rn into the soil heat flux, thereby allowing for the most accurate Qle and Qh fluxes (but for the wrong reasons), while HTESSEL actually produces the best Qg and in the process overestimates Qle and Qh.

In essence, the atmospheric model only cares about the fluxes of Qh and Qle coming from the land surface and, therefore, from a land coupling perspective, CLM



	Jun	Jul	Aug	Obs
Bent	-0.80105	-0.86682	-0.79514	-1.22199
Bsfc	3.589199	2.322341	1.215952	0.912545
Ah	0.741255	0.798984	1.293633	4.64802
Ale	-3.3213	-2.1406	-1.97827	-3.47101

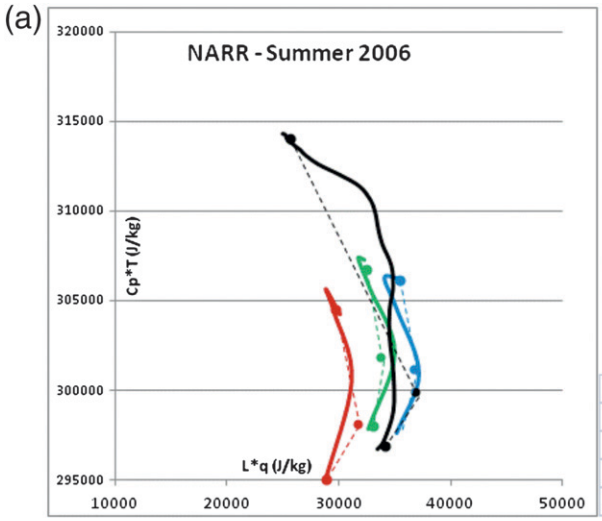


	Jun	Jul	Aug	Obs
Bent	-0.44614	-0.573	-0.54234	-0.83047
Bsfc	0.340416	0.246756	0.426687	0.37522
Ah	1.049836	1.817806	1.629709	1.493067
Ale	-0.80105	-0.78282	-1.28218	-0.67459

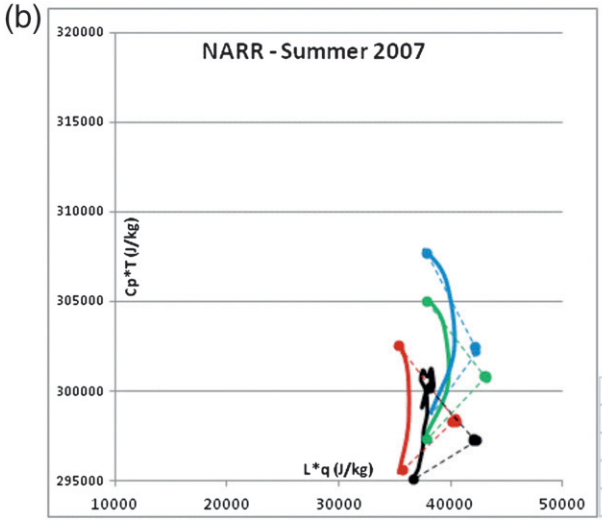
FIG. 11. Mixing diagrams derived from MERRA monthly mean diurnal cycles (red—June; green—July; blue—August) and the 7-day composite observations for the (a) 2006 and (b) 2007 case studies.

should perform best in 2007. Instead, we see the opposite occur in that HTESSEL produces the best coupling metrics [mixing diagrams, MAE, PBL budgets (entrainment and total), and EF versus PBLH] despite CLM providing the better land boundary conditions. This again supports that during a wet regime the land influence becomes diminished relative to the PBL (and presumably convective and microphysics) components. It should also be noted that for longer time scales (e.g., seasonal), there will be a feedback of the Qg bias in CLM on the coupled system in terms of evolving heat and moisture states.

The results are supportive of those from Kato et al. (2006) in that the choice of LSM does have substantial impact on simulated water and energy fluxes and states. Likewise, it is hoped that this type of analysis can pinpoint strengths and deficiencies in schemes (offline and coupled) that lead to model development. For example, that HTESSEL performs poorly (too much evaporation) in the wet extreme may be due to the dew deposition representation in the model that can lead to supersaturation at the surface during very wet conditions (G. Balsamo 2011, personal communication). This will be investigated by the developers of HTESSEL at ECMWF, with modifications being tested both in their offline configuration and NU-WRF as performed in this study. The partitioning of excess radiation into Qg (versus Qh and Qle) is another result that should lead to increased scrutiny of energy balance calculations within and across LSMs, as the impact on coupled simulations has been shown here to be important. Likewise, these results further highlight the need for improved PBL representation during stable



	Jun	Jul	Aug	Obs
Bent	-3.77169	-3.53481	-3.50328	-1.22199
Bsfc	1.105404	3.311369	2.153849	0.912545
Ah	1.763765	1.386364	1.466334	4.64802
Ale	-0.51692	-1.29873	-0.90151	-3.47101



	Jun	Jul	Aug	Obs
Bent	-0.78031	-0.81074	-1.15509	-0.83047
Bsfc	0.577685	0.59349	0.832655	0.37522
Ah	1.275572	1.214228	1.560154	1.493067
Ale	-0.94434	-0.88886	-1.12465	-0.67459

FIG. 12. As in Fig. 11, but for NARR. Mixing diagrams derived from NARR monthly mean diurnal cycles (red—June; green—July; blue—August) and the 7-day composite observations for the (a) 2006 and (b) 2007 case studies. (Courtesy of A. Kennedy.)

conditions, as there are implications for subsequent daytime coupling components and performance.

In terms of the PBL budget analysis, it is interesting that during the wet regime the observed total heat and moisture budget is approximately equal to the available energy at the surface (Rn − Qg), though at a much higher Bowen ratio. This suggests that, because of limited PBL growth, entrainment only acts to dry and warm the PBL slightly. During the dry regime with large PBL growth and entrainment, the total heat and moisture budget is considerably larger in magnitude than the surface available energy. Investigating the relative impacts of entrainment versus surface energy and their accuracies in the coupled schemes is a next step in this research as well. The main limitation to date is the availability of routine observations (e.g., profile data) to get at the biases in entrainment (and ratios).

Likewise, a detailed assessment of the impact of land–PBL coupling on precipitation is planned during the next phase of NU-WRF research. Buoyed by the analysis of the ''links in the chain'' of LoCo presented here, a comprehensive analysis that extends to the complex interactions with the convective and microphysics schemes in WRF will be performed. Preliminary results have shown sensitivity to LSM–PBL coupling, but also systematic biases in rainfall across all scheme combinations

that are likely due to synoptically driven features. Another interesting question to address in this regard will be the potential impact of the Rn bias on precipitation amount and timing.

Lastly, a community-wide LoCo effort is also being planned next that combines models (column, regional, and global), sites and regimes, and satellite observations of surface, near-surface, and PBL states as a benchmark from which to intercompare products. In this project, the LoCo methodology will being repeated for other sites, regions, and case studies in order to further understand the coupling strength and behavior in MERRA versus that of high-resolution regional models (e.g., LIS–WRF), other reanalysis products (e.g., NARR), and remotely sensed observations [e.g., Atmospheric Infrared Sounder (AIRS)]. Understanding how these models and their components perform both coupled and offline remains a critical challenge (e.g., NARR; Fan et al. 2011) from which the ultimate improvement of water and energy cycle representation in models of all scales relies.

Acknowledgments. This work was supported and motivated by the NASA Energy and Water Cycle Study (NEWS; PM: Jared Entin) and the Modeling and Extremes Working Groups. The NU-WRF team was also instrumental in providing support related to LIS–WRF coupling and a stable and updated version of the system. We also appreciate the past and ongoing collaboration with the LoCo community and working group that has stimulated this work, in particular Michael Ek, Cor Jacobs, Obbe Tuinenburg, Chiel van Heerwaarden, Bart van den Hurk, and Martin Best.

REFERENCES

Anthes, R. A., and T. T. Warner, 1978: Development of hydrodynamic models suitable for air pollution and other mesometerological studies. *Mon. Wea. Rev.,* **106,** 1045–1078.

Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. K. Betts, 2009: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeor.,* **10,** 623–643.

Betts, A. K., 1992: FIFE atmospheric boundary layer budget methods. *J. Geophys. Res.,* **97** (D17), 18 523–18 532.

Chin, M., R. B. Rood, S.-J. Lin, J. F. Muller, and A. M. Thomspon, 2000: Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties. *J. Geophys. Res.,* **105** (D20), 24 671–24 687.

Dai, Y. J., and Coauthors, 2003: The Common Land Model. *Bull. Amer. Meteor. Soc.,* **84,** 1013–1023.

Dong, X., and Coauthors, 2011: Investigation of the 2006 drought and 2007 flood extremes at the Southern Great Plains through an integrative analysis of observations. *J. Geophys. Res.,* **116,** D03204, doi:10.1029/2010JD014776.

ECMWF, cited 2011: Integrated Forecast System documentation. [Available online at http://www.ecmwf.int/research/ifsdocs/.]

Ek, M. B., and A. A. M. Holtslag, 2004: Influence of soil moisture on boundary layer cloud development. *J. Hydrometeor.,* **5,** 86–99.

——, K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta Model. *J. Geophys. Res.,* **108,** 8851, doi:10.1029/2002JD003296.

Fan, Y., H. M. van den Dool, and W. Wu, 2011: Verification and intercomparison of multimodel simulated land surface hydrological datasets over the United States. *J. Hydrometeor.,* **12,** 531–555.

Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate,* **24,** 4973–4991.

Hirschi, M., and Coauthors, 2011: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nat. Geosci.,* **4,** 17–21, doi:10.1038/ngeo1032.

Hong, S. Y., and H. L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124,** 2322–2339.

——, Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.,* **134,** 2318–2341.

Janjic, Z. I., 2001: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP mesoscale model. National Centers for Environmental Prediction Office Tech. Rep. 437, 61 pp.

Kato, H., M. Rodell, F. Beyrich, H. Cleugh, E. van Gorsel, H. Liu, and T. P. Meyers, 2006: Sensitivity of land surface simulations to model physics, parameters, and forcings, at four CEOP sites. *J. Meteor. Soc. Japan,* **85A,** 187–204.

Kennedy, A. D., X. Dong, B. Xi, S. Xie, Y. Zhang, and J. Chen, 2011: A comparison of MERRA and NARR reanalyses with the DOE ARM SGP data. *J. Climate,* **24,** 4541–4557.

Koster, R. D., and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science,* **305,** 1138–1140.

——, and Coauthors, 2010: Contribution of land surface initialization to subseasonal forecast skill: First results from a multimodel experiment. *Geophys. Res. Lett.,* **37,** L02402, doi:10.1029/2009GL041677.

Kumar, S. V., and Coauthors, 2006: Land information system: An interoperable framework for high resolution land surface modeling. *Environ. Modell. Software,* **21,** 1402–1415.

——, C. D. Peters-Lidard, J. L. Eastman, and W.-K. Tao, 2008: An integrated high resolution hydrometeorological modeling testbed using LIS and WRF. *Environ. Modell. Software,* **23,** 169–181.

——, ——, J. A. Santanello, K. Harrison, Y. Liu, and M. Shaw, 2012: Land surface Verification Toolkit (LVT)—A generalized framework for land surface model evaluation. *Geosci. Model Dev.,* **5,** 869–886.

Matsui, T., and Coauthors, 2009: Goddard Satellite Data Simulation Unit: Multi-sensor satellite simulators to support aerosol-cloud-precipitation satellite missions. *Eos, Trans. Amer. Geophys. Union,* **90** (Fall Meeting Suppl.), Abstract A21D-0268.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.,* **20,** 851–875.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.,* **87,** 343–360.

Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock, 2001: Development of a next generation regional weather research and forecast model. *Proc. Ninth Workshop on the Use of High Performance Computing in Meteorology,* Singapore, ECMWF, 269–276.

NEWS Science Integration Team, 2007: Predicting energy and water cycle consequences of earth system variability and change. NASA Earth Science Implementation Plan for Energy and Water Cycle Research, 89 pp.

Peters-Lidard, C. D., and Coauthors, 2007: High-performance Earth system modeling with NASA/GSFC's Land Information System. *Innovations Syst. Software Eng.,* **3,** 157–165.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* **24,** 3624–3648.

Santanello, J. A., M. A. Friedl, and M. Ek, 2007: Convective planetary boundary layer interactions with the land surface at diurnal time scales: Diagnostics and feedbacks. *J. Hydrometeor.,* **8,** 1082–1097.

——, C. D. Peters-Lidard, S. V. Kumar, C. Alonge, and W.-K. Tao, 2009: A modeling and observational framework for diagnosing local land–atmosphere coupling on diurnal time scales. *J. Hydrometeor.*, **10,** 577–599.

——, ——, and ——, 2011: Diagnosing the sensitivity of local land–atmosphere coupling via the soil moisture–boundary layer interaction. *J. Hydrometeor.,* **12,** 766–786.

Shin, H. H., and S.-Y. Hong, 2011: Intercomparison of planetary boundary-layer parametrizations in the WRF model for a single day from CASES-99. *Bound.-Layer Meteor.,* **139,** 261–281.

Siqueira, M., G. Katul, and A. Porporato, 2009: Soil moisture feedbacks on convection triggers: The role of soil–plant hydrodynamics. *J. Hydrometeor.,* **10,** 96–112.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp. [Available online at http://box.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf.]

Steeneveld, G. J., T. Mauritsen, E. I. F. DeBruijn, J. V.-G. DeArellano, G. Svensson, and A. A. M. Holtslag, 2008: Evaluation of limited-area models for the representation of the diurnal cycle and contrasting nights in CASES-99. *J. Appl. Meteor. Climatol.,* **47,** 869–887.

Troen, I., and L. Mahrt, 1986: A simple model of the atmospheric boundary layer: Sensitivity to surface evaporation. *Bound.-Layer Meteor.,* **37,** 129–148.

van den Hurk, B., F. Doblas-Reyes, G. Balsamo, and R. D. Koster, S. I. Seneviratne, and H. Camargo Jr., 2010: Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dyn.,* **38,** 349–362, doi:10.1007/s00382-010-0956-2.

——, M. Best, P. Dirmeyer, A. Pitman, J. Polcher, and J. Santanello Jr., 2011: Acceleration of land surface model development over a decade of glass. *Bull. Amer. Meteor. Soc.,* **92,** 1593–1600.

van Heerwaarden, C. C., J. Vilà-Guerau de Arellano, A. F. Moene, and A. A. M. Holtslag, 2009: Interactions between dry-air entrainment, surface evaporation and convective boundary-layer development. *Quart. J. Roy. Meteor. Soc.,* **135,** 1277–1291, doi:10.1002/qj.431.

Weckworth, T. M., and Coauthors, 2004: An overview of the International H_2O Project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.,* **85,** 253–277.