Journal of Hydrometeorology

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

--Manuscript Draft--

Manuscript Number:	
Full Title:	Diagnosing the Nature of Land-Atmosphere Coupling: A Case Study of Dry/Wet Extremes in the U. S. Southern Great Plains
Article Type:	Article
Corresponding Author:	Joseph Anthony Santanello, Ph.D. NASA-GSFC Greenbelt, MD UNITED STATES
Corresponding Author's Institution:	NASA-GSFC
First Author:	Joseph Anthony Santanello, Ph.D.
Order of Authors:	Joseph Anthony Santanello, Ph.D.
	Christa Peters-Lidard
	Aaron Kennedy
	Sujay Kumar
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 are 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 testbed. 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 towards 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.
Suggested Reviewers:	Hsin-Yuan Huang hyhuang@ucla.edu Recent research in land-PBL coupling and interactions.
	Joseph Alfieri Joe.Alfieri@ARS.USDA.GOV Expertise in SGP region, observations and models (L-A interactions).

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.[1], Christa D. Peters-Lidard[1], Aaron Kennedy[2], and Sujay V. Kumar[3,1]

[1] NASA-GSFC Hydrological Sciences Branch, Greenbelt, MD

[2] University of North Dakota, Grand Forks, ND

[3] Science Applications International Corporation, McLean, VA

Submitted to *Journal of Hydrometeorology*

February 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 are 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 testbed. 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 towards 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 (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. 2009, 2011) in coupled models by diagnosing land-PBL interactions at the process level using a regional, high-resolution testbed. The methodology 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 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 and feedbacks act to support such events? 3) How well do large-scale, coarse resolution models represent LoCo during dry/wet extremes?

This comprehensive analysis extends from the work of Santanello et al. (2009, 2011; hereafter referred to as S09 and S11) by performing an evaluation using observations along with composite and regional (including reanalysis product) analyses. 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; Hurk et al. 2011), which coordinates a community working

group on studies related to L-A coupling (Santanello et al. 2011). 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, Hurk et al. 2010; Hirschi et al. 2011). The initial communication and all interaction between the land and atmosphere always occurs on local scales, however, which makes 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,

$$\underset{\textbf{(a)}}{\Delta SM \rightarrow \Delta EF_{sm}} \underset{\textbf{(b)}}{\rightarrow \Delta PBL} \underset{\textbf{(c)}}{\rightarrow \Delta ENT} \underset{\textbf{(d)}}{\rightarrow \Delta EF_{atm}} \blacktriangleright \Delta P/Clouds \qquad (1)$$

where EF is the evaporative fraction, defined as

$$EF = \frac{Qle_{sfc}}{Qh_{sfc} + Qle_{sfc}} \quad . \qquad (2)$$

and is a function of the sensible (Qh_{sfc}) and latent (Qle_{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: **a)** the surface fluxes (EF_{sm}) to soil moisture, **b)** PBL evolution to surface fluxes, **c)** entrainment fluxes at the PBL-top (ENT) to PBL evolution, and **d)** the collective feedback of the atmosphere (through the PBL) on surface fluxes (EF_{atm}) (Santanello et al. 2007, van Heerwaarden et al. 2009). Eq. 1 describes a complex set of dependent relationships, and as Siqueira (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). This power of this diagnostic lies in its ability to exploit the co-variance of 2-meter potential temperature (θ) and humidity (q) to quantify the components of the LoCo process chain, and 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 modelling system, each with a different LSM-PBL scheme combination. The results show that soil moisture anomalies (dry vs. 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 ratio (β_{sfc}, β_{ent}), and the entrainment ratio of heat and moisture (A_h, A_{le}) 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 that of the relationship between mean EF and PBL height (PBLH), which serve as bulk integrative measures of the state of the land surface and the PBL. How each is reflective and 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. While their results were applied to a few individual days, the focus was clearly on developing and demonstrating the methodology and associated diagnostics rather than specific or thorough evaluation of the schemes themselves that extend beyond a single diurnal cycle. 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 testbed and focused on climatologically and programmatically-relevant case studies over multi-day periods.

3. Model and Site Description

a. NU-WRF

The Advanced Research version of the Weather Research and Forecasting (WRF-ARW) model (Michalakes et al. 2001) is a state of the art mesoscale numerical weather prediction

system. Derived from the Fifth-Generation NCAR/Penn State Mesoscale Model (MM5; Anthes and Warner 1978), WRF-ARW has been designated as the community model for atmospheric research and operational prediction and is ideal for high-resolution (e.g. 1km) regional simulations on the order of 1-10 days. WRF-ARW has a Eulerian mass dynamical core and includes a wide array of radiation, microphysics, and PBL options as well as 2-way nesting and variational data assimilation capabilities.

Recently, work has been performed to develop a NASA-Unified WRF (NU-WRF) modeling system at NASA's Goddard Space Flight Center (GSFC). NU-WRF is built upon the WRF-ARW 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. (LIS-WRF; 2008). 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 testbed for L-A interaction studies and LoCo due to 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 V 6.x) and WRF-ARW (currently V 3.x) are coupled and tested, and are used exclusively for the 2006-7 simulations described in Section 3b.

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 version of the Tiled 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 vary in specific parameterizations and representation of soil and vegetation properties and physics. For example, Noah and HTESSEL solves 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 WRF-ARW Version 2.2. Noah is used operationally by the National Center 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 WRF-ARW releases), and it should be noted that CLM is an earlier version of the LSM for 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 GLACE experiments. 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 WRF-ARW, 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 non-local-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 non-local 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 (1st order to TKE) and application (single column to full 3-D) of those participating in the GEWEX Atmospheric Boundary Layer Study (GABLS; www.gewex.org/gabls.htm). We therefore expect the results to encapsulate much of the range of L-A coupling possible between LSMs and PBLs participating in PILPS and 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 x 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: 2006-7 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 International H_2O Project in June 2002 (IHOP-02; Weckworth et al. 2004), and evaluated using data from the Atmospheric and Radiation Measurement testbed located in the region (ARM-SGP). The ARM-SGP region has also recently been the focus of studies on extreme conditions observed during the 2006-7 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 2006-7 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. (2010). 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; NASA 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 and results of Dong et al. (2010), the summers (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-02 experiments in S09 and S11, each of the three LSMs in LIS were run offline (uncoupled) for an approximately 4-year period prior to the start time of the 2006 and 2007 case studies to create equilibrated, or 'spun-up', land surface states for initialization of LIS-WRF. Using these spun-up surface fields as initial surface conditions for the 2006-7 case studies, NU-WRF simulations were then performed over a single, high-resolution domain (500 x 500; see Fig. 2), centered over Oklahoma and Kansas with a horizontal resolution of 1 km and timestep 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 3 LSMs valid at 00Z 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. Overall, soil moisture in the ARM-SGP region varies significantly from dry and heterogeneous (generally < 25 percent volumetric) in 2006 to extremely wet (near saturation) and more uniform conditions 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 co-located soil moisture, net radiation, sensible, latent, and soil heat, along with co-located 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 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 was collected from 24 sites in the domain that measure surface fluxes using eddy correlation (ECOR) and Bowen ratio (EBBR) towers, along with the co-located surface meteorology, gravimetric soil moisture probes, and where available radiosonde profile data.

4. Results

In order 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: a) evaluation of the surface fluxes, b) application of LoCo diagnostics, and c) 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 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 1km 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 Wm-2), Qh (>50 Wm-2), and Qg (>40 Wm-2) exist in all LSM-PBL combinations. 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 the North American Land Data Assimilation System (NLDAS-2; Xia et al. 2011) dataset. 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. 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. Qg is relatively unbiased in all runs in both 2006 and 2007 due to the diurnal cycle of Qg (positive during the day, negative at night) that 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 towards 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 Wm-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 Wm-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. 4de), 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.

(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 Wm-2 daytime peak) of Rn versus that observed, both in domain average (Fig. 5ac) and at individual stations (Fig. 5de) and regardless of LSM or PBL choice. In the Noah runs (Fig. 5ad), 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 and is a significant contrast to the coupled runs overestimates. The reduced energy does not allow for evaporation to match that observed, though Qh and Qg are simulated well (though quite small relative to Qle). This comparison of offline (good forcing) vs. 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 (Fig. 5de) 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 towards Qh and Qg thereby producing an evaporative flux that is only slightly overestimated. HTESSEL + YSU exhibits the opposite effect of too much radiation. Due to its nearly saturated soil (as discussed above) HTESSEL produces evaporative fluxes that are extremely high, virtually matching the atmospheric 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 comprise 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, MSE, 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 7-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 co-evolution 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 also shows 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 generally follow suit.

The implications for different LSM-PBL coupling is also evident in the thermodynamics. Equivalent potential temperature (θ_e) is simulated quite well in all runs, but is much less important during this 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.

In order 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 and root mean squared error can be established as follows,

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

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

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 a) the MYJ produces the best coupling for all LSMs, and b) 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 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 non-local schemes during stable conditions (Shin et al. 2011, Steeneveld et al. 2008), as the YSU/MRF formulations for nighttime mixing resulting 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 a much 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 has having the best Qle and HTESSEL the worst (and overestimated), but here 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 underestimate the available energy (Rn - Qg), primarily as a result of overestimation 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 is largely consistent with the MAE results presented above, which is important for assessing LoCo as this is based on the flux components while TE was 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 towards 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 VS. PBL HEIGHT

A third diagnostic of the LoCo behavior is the relationship of evaporative fraction (EF) to PBL height (PBLH), which serves as integrated measures of the land and PBL condition, respectively. 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, due to 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 and land surface perspective, however, it is clear that this may be the right answer for the wrong reasons (e.g. surface Qle) and highlights the importance of not simply focusing on integrated properties of the system 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, 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 plotted for the two case studies and sites in Figs. 6-7. 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 stressed 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 towards 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 2245h. 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 NASA Energy and Water Cycle Study (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 testbed has been shown here to be useful and essential towards 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 MERRA and NARR fields 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 MERRA and NARR coupling and its variability over a semi-seasonal scale, rather than just the 1-week that

overlaps with the NU-WRF case studies. This better demonstrates how these products and their sensitivities in coupled components evolve on longer timescales(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 1/2 x 2/3 degree resolution. Figure 11 shows the mixing diagrams for MERRA's monthly mean diurnal cycles (June, July, 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 is also consistent. That the observations show wider diurnal range (T2m, Q2m) than MERRA during July 2006 is not surprising due to the smaller averaging (spatially and temporally) of the observations 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) North American Regional Reanalysis (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 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 NARR. 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; MERRA - Catchment).

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.

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 is diagnosing L-A coupling have been applied to a high-resolution regional modeling testbed 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 comprise the land-PBL coupling (Eq. 1).

Key findings from the land surface energy budget analysis include:

• Significant errors exist in land surface energy balance simulations that depend on 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; CLM vice-versa.

• The differences in offline vs. coupled land surface fluxes are greater during the wet regime when 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 include:

• The sensitivity of L-A coupling is stronger towards 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 coupled overestimation of net radiation 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

highlighted an important aspect of LoCo diagnosis in models in terms of how errors are translated between components of the 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 best Qle and Qh fluxes (which are the only ones the atmosphere cares about and therefore produces the best 'coupling'). HTESSEL actually produces the best Qg fluxes, which then leads to the worst Qle and Qh fluxes (and therefore coupling as well). So, in essence HTESSEL could have a better soil thermal parameterization than CLM, but as a result of incorrect Rn forcing produces overestimates of the turbulent fluxes. CLM ends up producing the better coupling, but for the wrong reasons. It should also be noted that for longer timescales (e.g. seasonal), there will be a feedback of the soil heat flux error 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 (Balsamo, pers. communication, 2011). 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. Likewise, these results further highlight the need for improved PBL representation during stable 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 the surface fluxes are dominant, and due to limited PBL growth entrainment only acts to dry and warm the PBL slightly. During dry conditions and 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 that observations (e.g. profile data) to get at the biases in entrainment (and ratios).

Along these lines, a comprehensive study is being planned next that combines models (column, regional, 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. 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.

Acknowledgements. 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.

6. 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, 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,** 18523–18532.

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

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

Dong, X., et al., 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.

Ek M. B., 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, 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., S.I. Seneviratne, V. Alexandrov, F. Boberg, C. Boroneant, O.B. Christensen, H. Formayer, B. Orlowsky, and P. Stepanek, 2011: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience*, **4,** 17-21, doi:10.1038/ngeo1032.

Hong, S.Y., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Mon. Wea. Rev.*, **134**, 2318–2341.

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.

Hurk, B., M. Best, P. Dirmeyer, A. Pitman, J. Polcher, and J. A. Santanello, 2011: Acceleration of Land Surface Model Development over a Decade of Glass. *Bull. Amer. Meteor. Soc.*, **92**, 1593–1600.

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

Hurk, B. J. J. M. van den, 2011: Over a Decade of GLASS....BAMS

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

Koster , R. D., S. Mahanama, T. Yamada, G. Balsamo, A.A. Berg, M. Boisserie, P. Dirmeyer, F. Doblas-Reyes, G. Drewitt, C.T. Gordon, Z. Guo, J.H. Jeong, D.M. Lawrence, W.-S. Lee, Z. Li, L. Luo, S. Maleyshev, W.J. Merryfield, S.I. Seneviratne, T. Stanelle, B.J.J.M. van den Hurk, F. Vitart and E.F. Wood, 2010: Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment, *Geophys. Res. Lett.,* **37,** L02402, doi:10.1029/2009GL041677.

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, J. Chen, 2011: A Comparison of MERRA and NARR Reanalyses with the DOE ARM SGP Data. *J. Climate*, **24**, 4541–4557.

Koster R. D., Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Nature*, **306**, 1138–1140.

Kumar, S. V., C. D. Peters-Lidard, 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.,* **submitted**.

Kumar, S. V., C. D. Peters-Lidard, J. L. Eastman, and W.-K. Tao, 2008: An integrated high resolution hydrometeorological modeling testbed using LIS and WRF. *Environmental Modelling and Software*, **23,** 169-181.

Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E. F. Wood and J. Sheffield, 2006: Land

Information System - An Interoperable Framework for High Resolution Land Surface Modeling. *Environmental Modelling & Software,* **Vol. 21,** 1402-1415.

Matsui, T., W. Tao, H. Masunaga, C. D. Kummerow, W. S. Olson, N. Teruyuki, M. Sekiguchi, M. Chou, T. Y. Nakajima, X. Li, J. Chern, J. J. Shi, X. Zeng, D. J. Posselt, K. Suzuki, 2009: Goddard Satellite Data Simulation Unit: Multi-Sensor Satellite Simulators to Support Aerosol-Cloud-Precipitation Satellite Missions, *Eos Trans. AGU,* **90(52),** Fall Meet. 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. *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the use of high performance computing in meteorology*. Singapore, pp. 269-276.

NASA Energy- and Water-Cycle Study (NEWS) Science Integration Team, 2007: *A NASA Earth Science Implementation Plan for Energy and Water Cycle Research: "Predicting Energy and Water Cycle Consequences of Earth System Variability and Change"*, 89 pp.

Peters-Lidard, C. D., P. R. Houser, Y. Tian, S. V. Kumar, J. Geiger, S. Olden, L. Lighty, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E. F. Wood and J. Sheffield, 2007: High-performance Earth system modeling with NASA/GSFC's Land Information System. *Innovations in Systems and Software Engineering,* **Vol. 3(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., 2011: Results from Local Land-Atmosphere Coupling (LoCo) Project. *GEWEX Newsletter,* **21(4),** 7-9.

Santanello, J. A., C. D. Peters-Lidard, and S. V. Kumar, 2011: Diagnosing the Sensitivity of Local Land–Atmosphere Coupling via the Soil Moisture–Boundary Layer Interaction. *J. Hydrometeor*, **12**, 766–786.

Santanello, J. A., Christa D. Peters-Lidard, Sujay V. Kumar, Charles Alonge, Wei-Kuo Tao, 2009: A Modeling and Observational Framework for Diagnosing Local Land–Atmosphere Coupling on Diurnal Time Scales. *J. Hydrometeor*, **10**, 577–599.

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.

Siqueira, Mario, Gabriel Katul, Amilcare Porporato, 2009: Soil Moisture Feedbacks on Convection Triggers: The Role of Soil–Plant Hydrodynamics. *J. Hydrometeor*, **10**, 96– 112.

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. *Boundary-Layer Meteorol*, **139,** 261-281.

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 Meteorol. Clim.,* **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 Heerwaarden, C. C., Vilà-Guerau de Arellano, J., Moene, A. F. and Holtslag, A. A. M., 2009: Interactions between dry-air entrainment, surface evaporation and convective boundary-layer development. *Quarterly Journal of the Royal Meteorological Society*, **135,** 1277–1291. doi: 10.1002/qj.431.

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

Xia, Y., Ek, M., Wei, H. and Meng, J., 2012: Comparative analysis of relationships between NLDAS-2 forcings and model outputs. *Hydrol. Processes,* **26,** 467–474.

LIST OF TABLES

LIST OF FIGURES

Figure 4. Mean diurnal cycles of the land surface energy balance for the **14-20 July 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.

Figure 5. Mean diurnal cycles of the land surface energy balance for the **14-20 June 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.

Figure 6. Mixing Diagrams from the LIS-WRF simulations of Noah, TESSEL, and CLM composited over the 7-day case studies of 2006 (14-20 July) and 2007 (14-20 June), and overlain with lines of constant equivalent potential temperature (Theta-e) and saturation humidity (q*sat) deficit.

Figure 7. Same as Figure 6, but for the 14-20 June 2007 case.

Figure 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.

Figure 9. Mean daytime Evaporative Fraction vs. PBL Height for each simulation vs. observed, along with the diurnal standard deviation through the 7-day period

Figure 10. Hourly LCL Deficit (= PBL height - LCL; mb) calculated for each of the PBL-LSM couplings for the **a)** 2006 and **b)** 2007 case studies.

Figure 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.

Figure 12. Same as 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 Aaron Kennedy]

	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	11962.43	**8965.73**	11231.87	9407.22	**4401.29**	8563.57	**8149.11**	8356.46	9199.85
Total MAE	11252.89	**8719.16**	10495.39	8605.25	**4213.57**	7845.91	8017.51	**7973.80**	9061.74

Table 1. Statistics (Jkg-1) based on evolution of T2m and Q2m vs. observations, derived from the mixing diagrams in Fig. 6.

	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

Table 2. Same as Table1, but for the 14-20 June 2007 case.

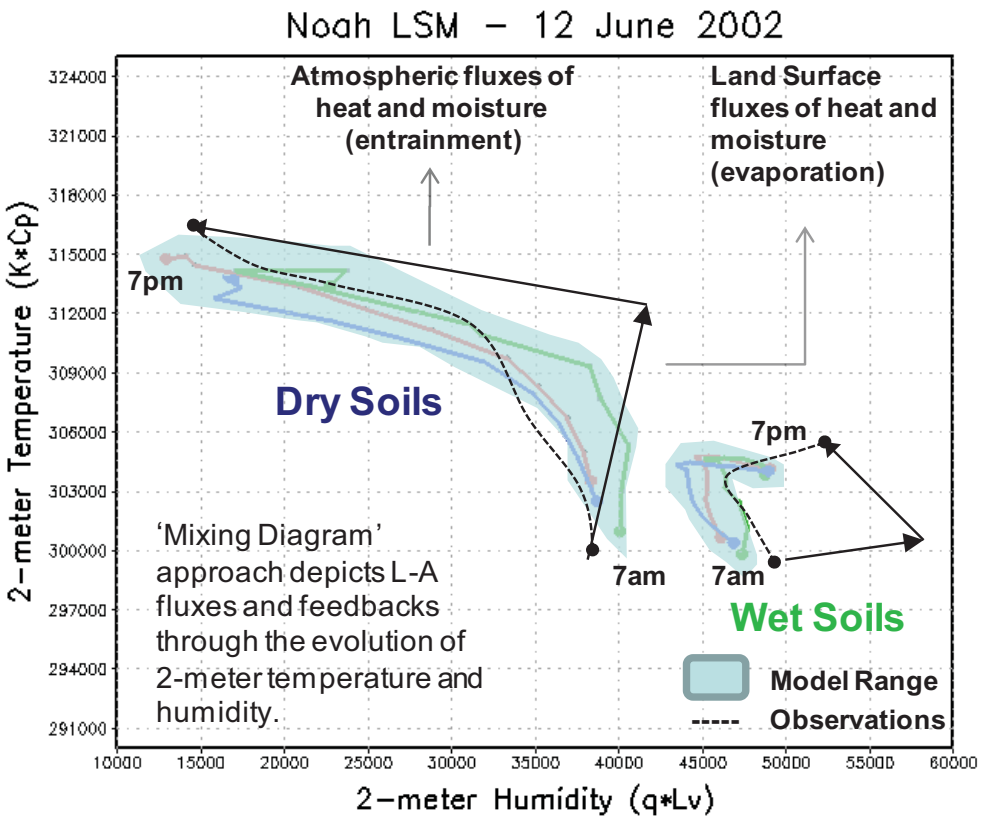


Figure 1: Mixing diagram showing the diurnal co-evolution (7am-7pm) of 2m-specific humidity and 2m-potential temperature on 12 June 2002 at a dry and wet soil location as simulated by a coupled mesoscale model (*derived from Figs. 2-5 in Santanello et al. 2009*). The shaded regions for each indicate the model range for different LSM-PBL scheme couplings (red, green, and blue) versus 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 versus those from the atmosphere due to entrainment, both of which are quantified using this approach.

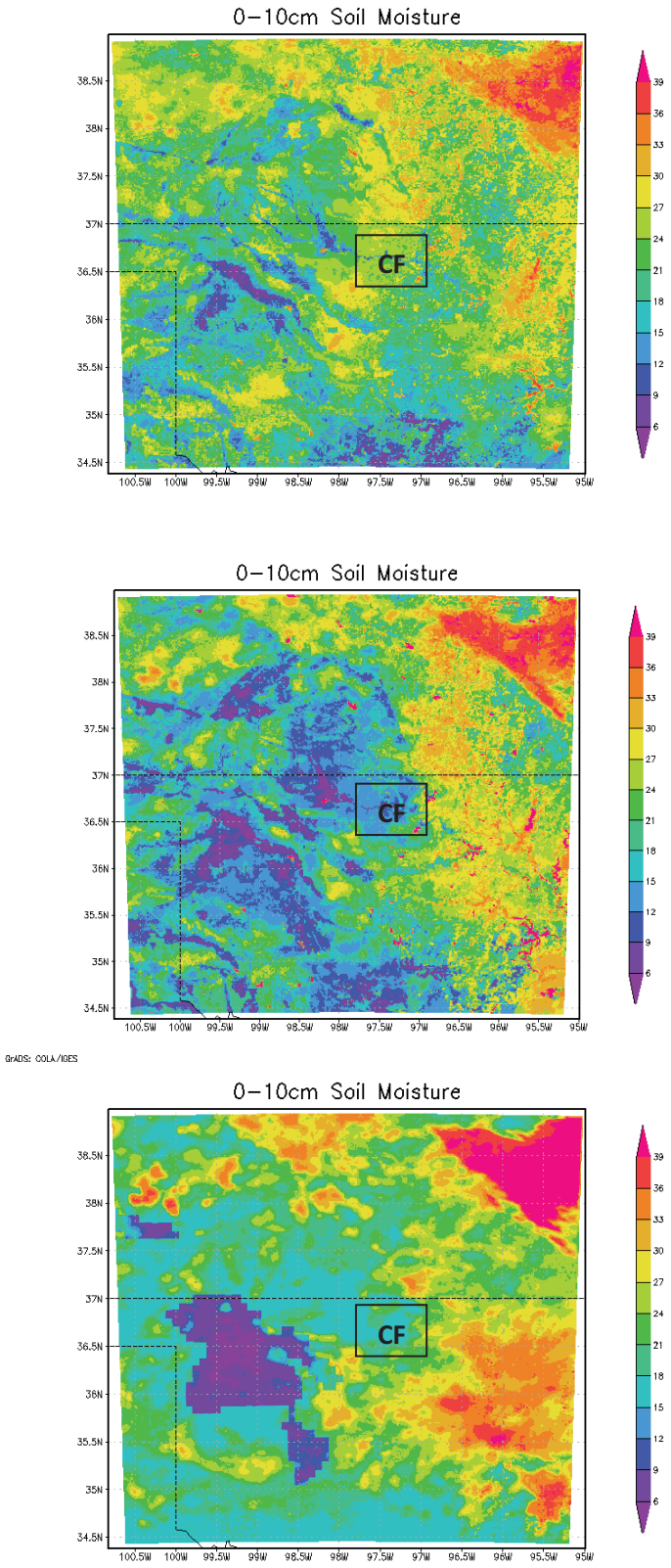


Figure 2a: Soil moisture (m^3m^{-3}*100) in the upper 0-10cm layer valid at 00Z on 14 July 2006 as simulated from a 3.5 year spinup of the a) Noah, b) CLM, and c) HTESSEL model over the 1-km LIS-WRF domain in the SGP. The ARM-SGP Central Facility (CF) at Lamont, OK is also shown.

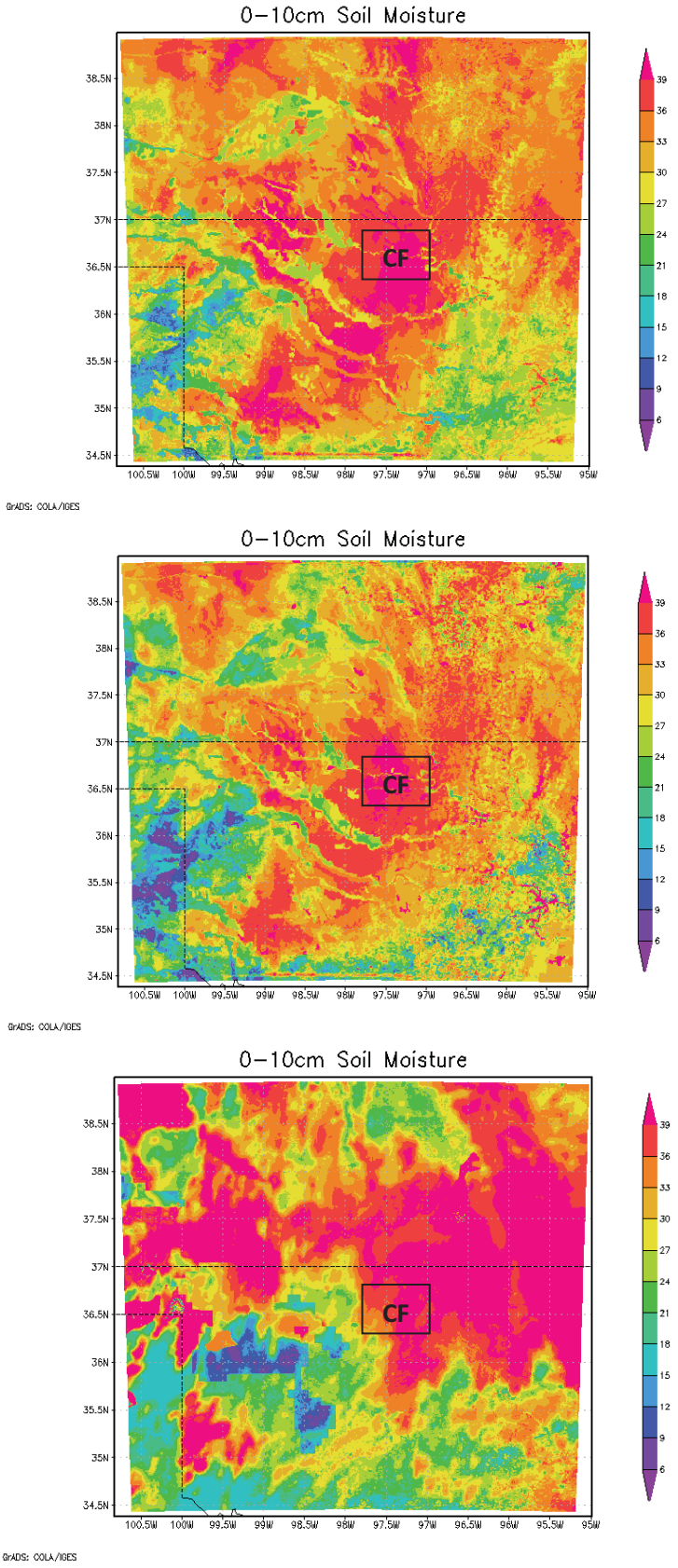


Figure 2b: Same as Fig. 2a, but valid at 00Z on 14 June 2007.

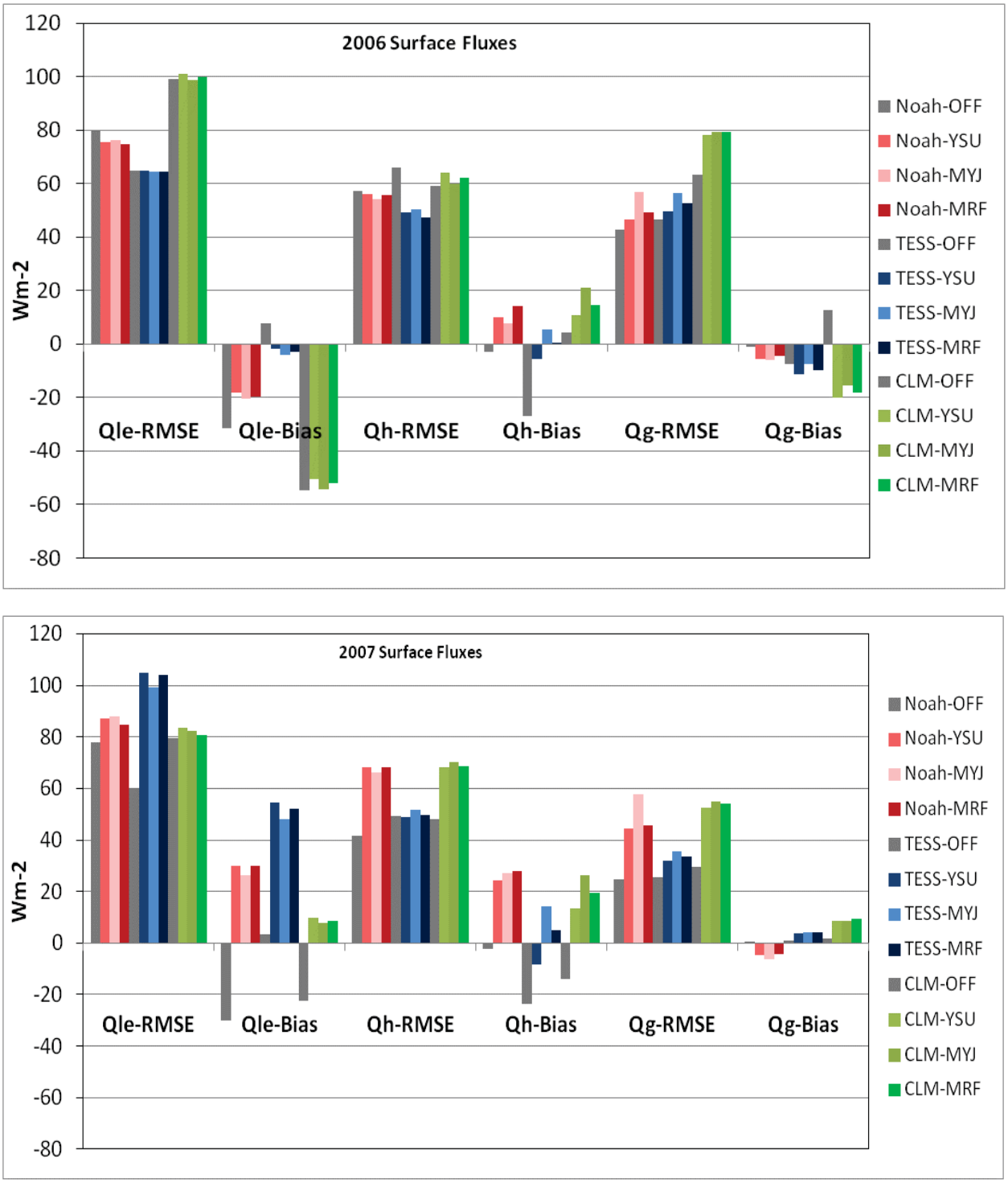
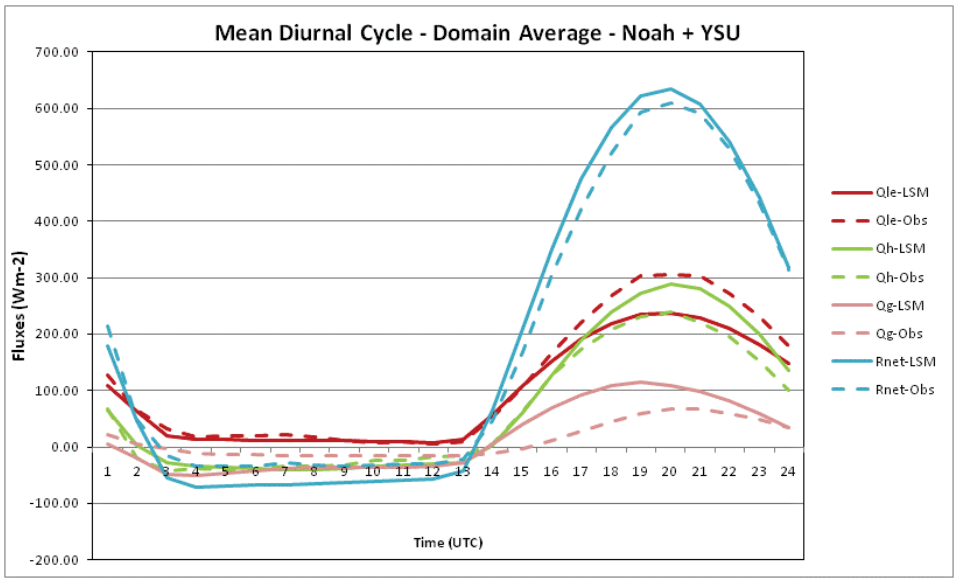
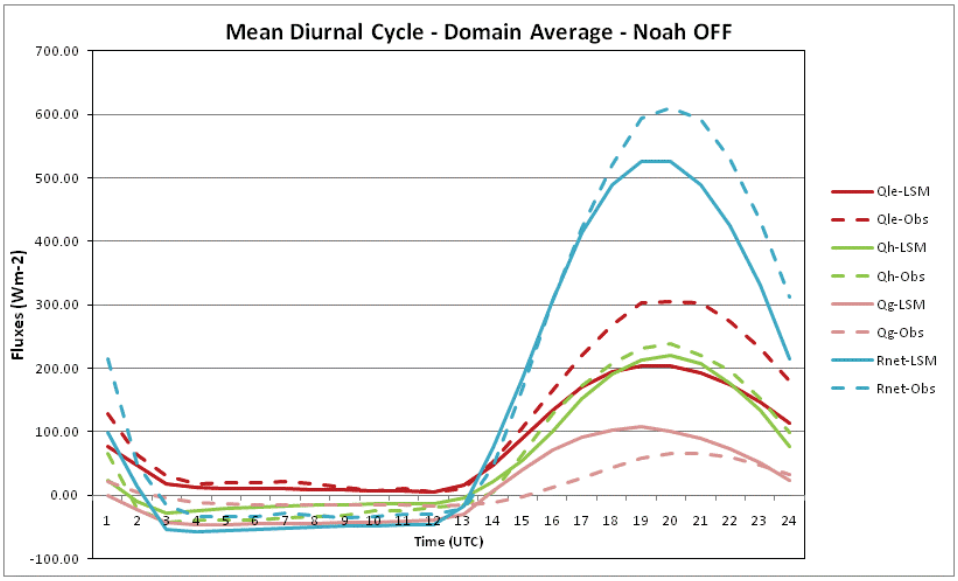
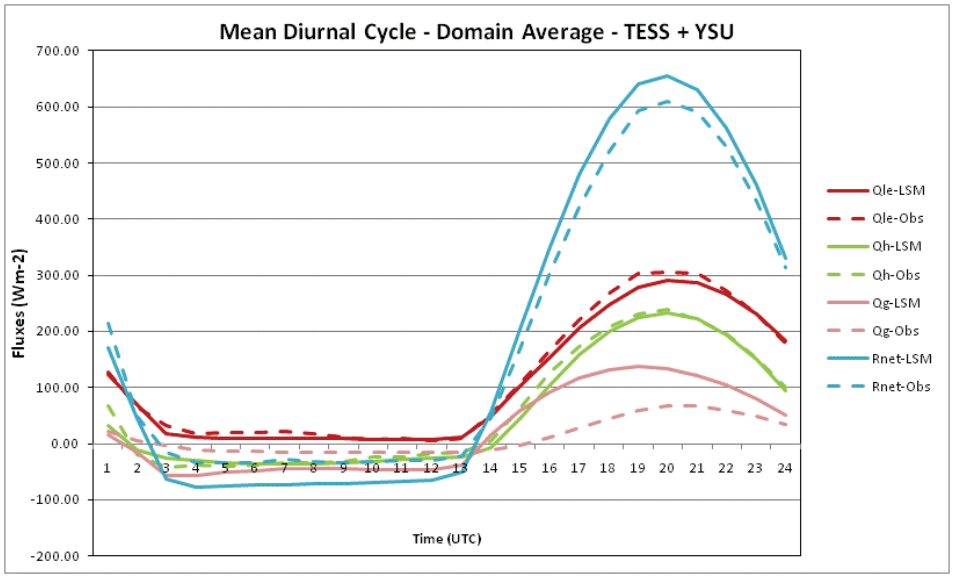


Figure 3: Domain-average RMSE and Bias statistics for the land surface fluxes of latent (Qle), sensible (Qh), and ground (Qg) heat as simulated by the LSM-PBL couplings in LIS-WRF over the a) 14-20 July 2006 and b) 14-20 June 2007 periods. Also shown are the offline simulations of each LSM spinup (OFF)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


Mean Diurnal Cycle - Domain Average - Noah + YSU
Fluxes (Wm-2)
Time (UTC)
Qle-LSM
Qle-Obs
Qh-LSM
Qh-Obs
Qg-LSM
Qg-Obs
Rnet-LSM
Rnet-Obs


Mean Diurnal Cycle - Domain Average - Noah OFF
Fluxes (Wm-2)
Time (UTC)
Qle-LSM
Qle-Obs
Qh-LSM
Qh-Obs
Qg-LSM
Qg-Obs
Rnet-LSM
Rnet-Obs


Mean Diurnal Cycle - Domain Average - TESS + YSU
Fluxes (Wm-2)
Time (UTC)
Qle-LSM
Qle-Obs
Qh-LSM
Qh-Obs
Qg-LSM
Qg-Obs
Rnet-LSM
Rnet-Obs

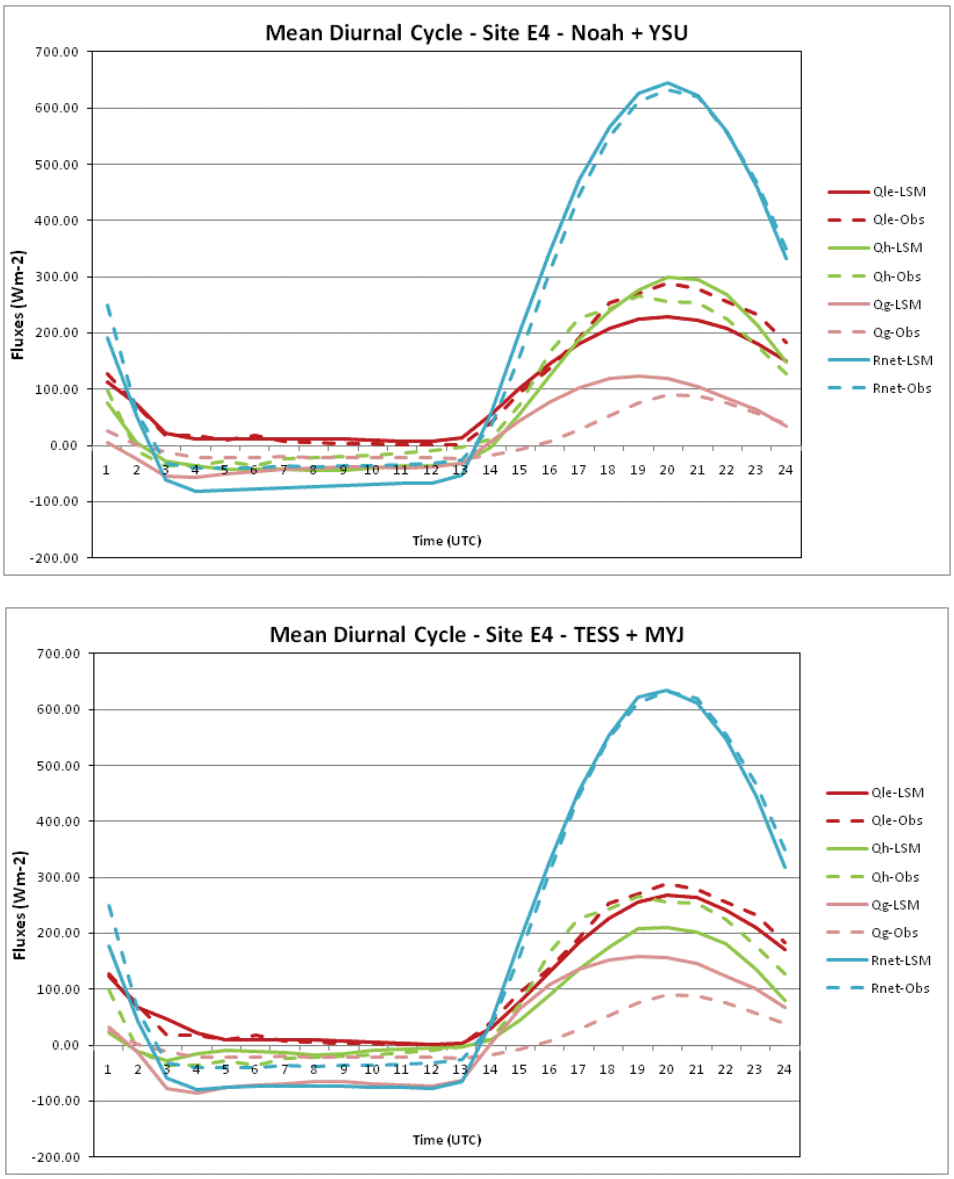
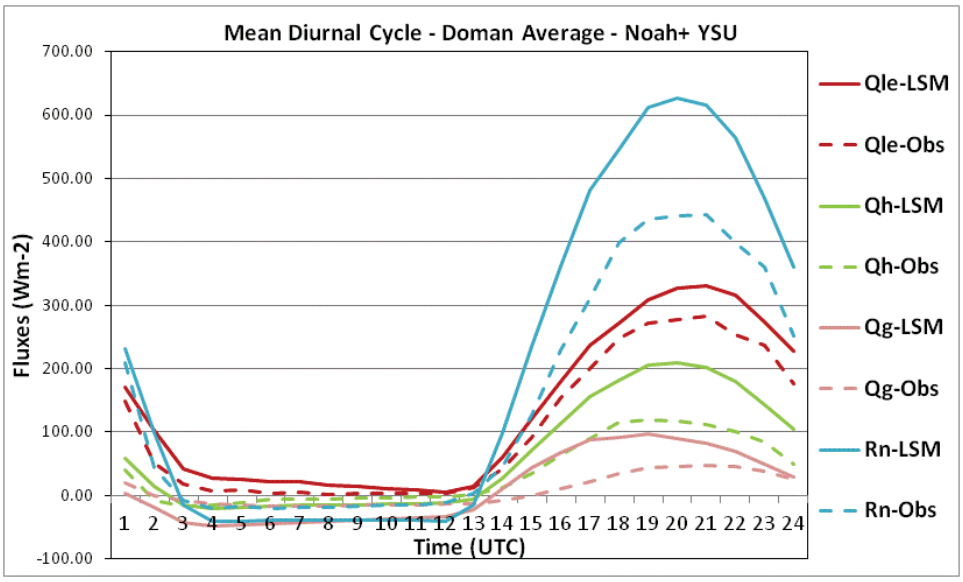
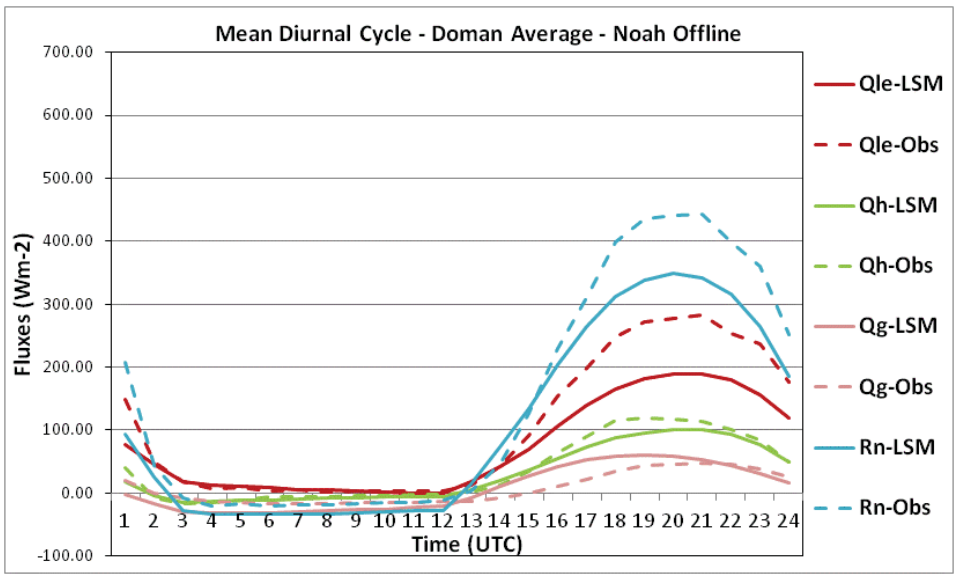
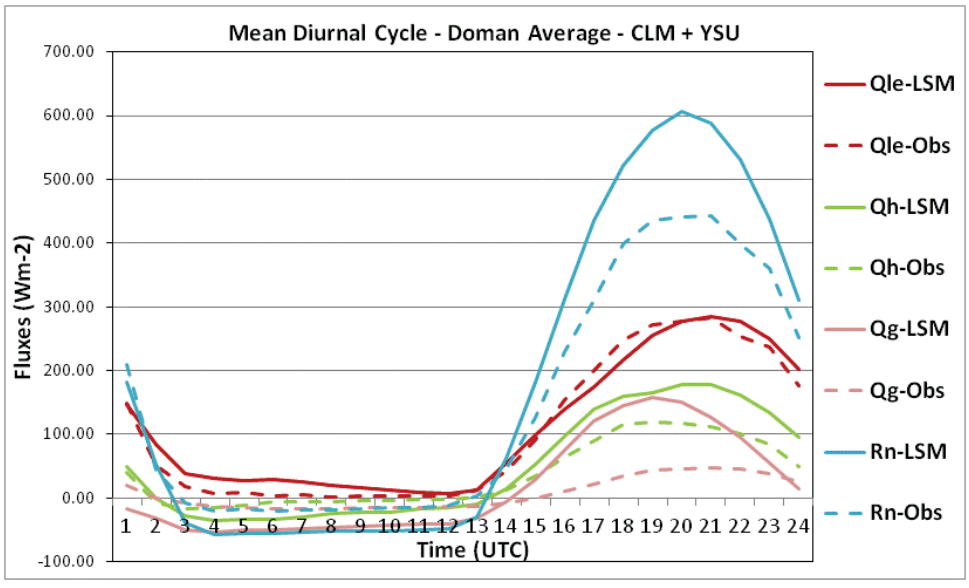
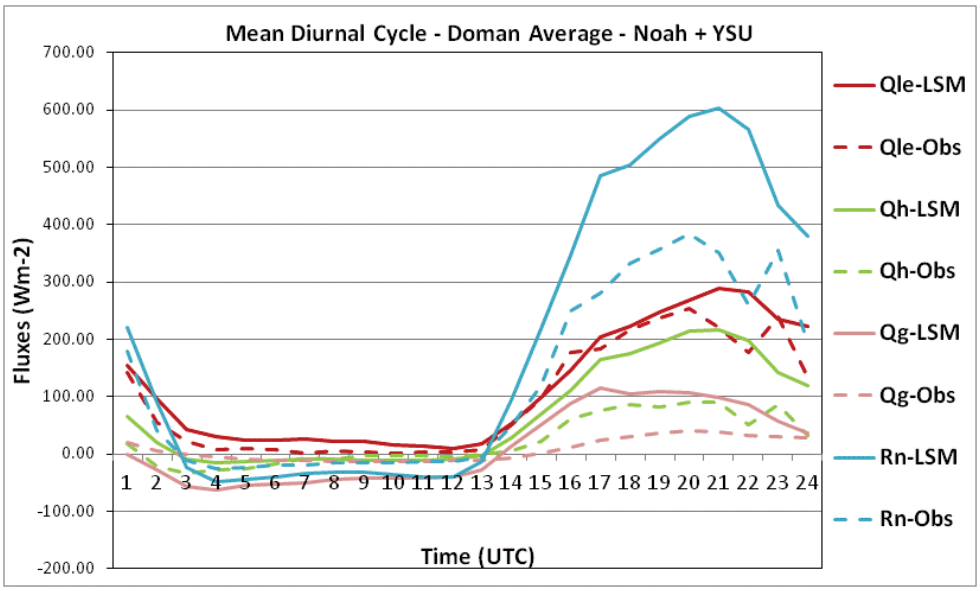


Figure 4. Mean diurnal cycles of the land surface energy balance for the **14-20 July 2006** period as simulated and observed for domain-averaged a) Noah + YSU, b) Noah offline, and c) TESSEL + YSU and Site E4 d) Noah + YSU and e) TESSEL + MYJ.


Mean Diurnal Cycle - Doman Average - Noah+ YSU
Fluxes (Wm-2)
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
-100.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Time (UTC)
Qle-LSM
Qle-Obs
Qh-LSM
Qh-Obs
Qg-LSM
Qg-Obs
Rn-LSM
Rn-Obs


Mean Diurnal Cycle - Doman Average - Noah Offline
Fluxes (Wm-2)
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
-100.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Time (UTC)
Qle-LSM
Qle-Obs
Qh-LSM
Qh-Obs
Qg-LSM
Qg-Obs
Rn-LSM
Rn-Obs


Mean Diurnal Cycle - Doman Average - CLM + YSU
Fluxes (Wm-2)
700.00
600.00
500.00
400.00
300.00
200.00
100.00
0.00
-100.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Time (UTC)
Qle-LSM
Qle-Obs
Qh-LSM
Qh-Obs
Qg-LSM
Qg-Obs
Rn-LSM
Rn-Obs



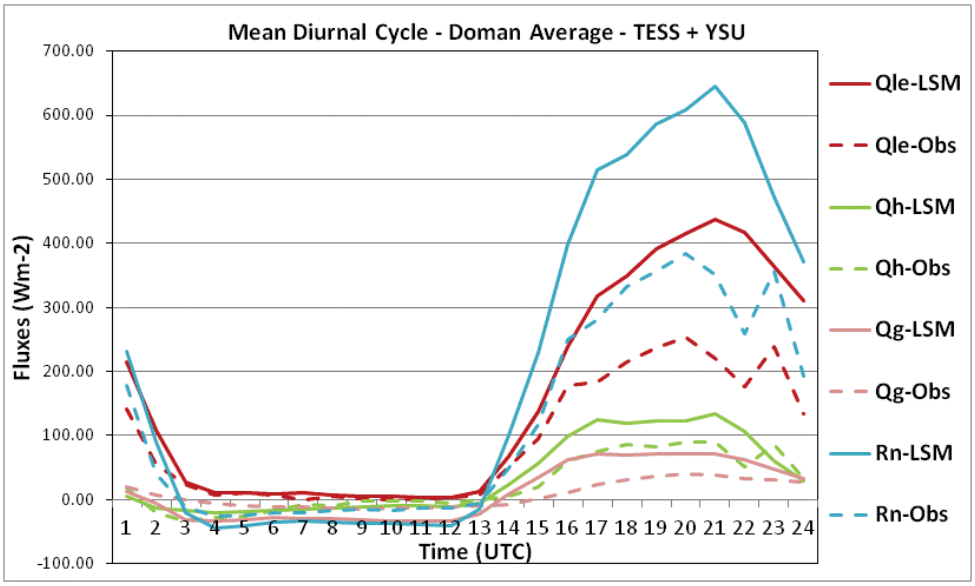


Figure 5. Mean diurnal cycles of the land surface energy balance for the **14-20 June 2007** period as simulated and observed for domain-averaged a) Noah + YSU, b) Noah offline, and c) CLM + YSU and Site E13 d) Noah + YSU and e) TESSEL + YSU.

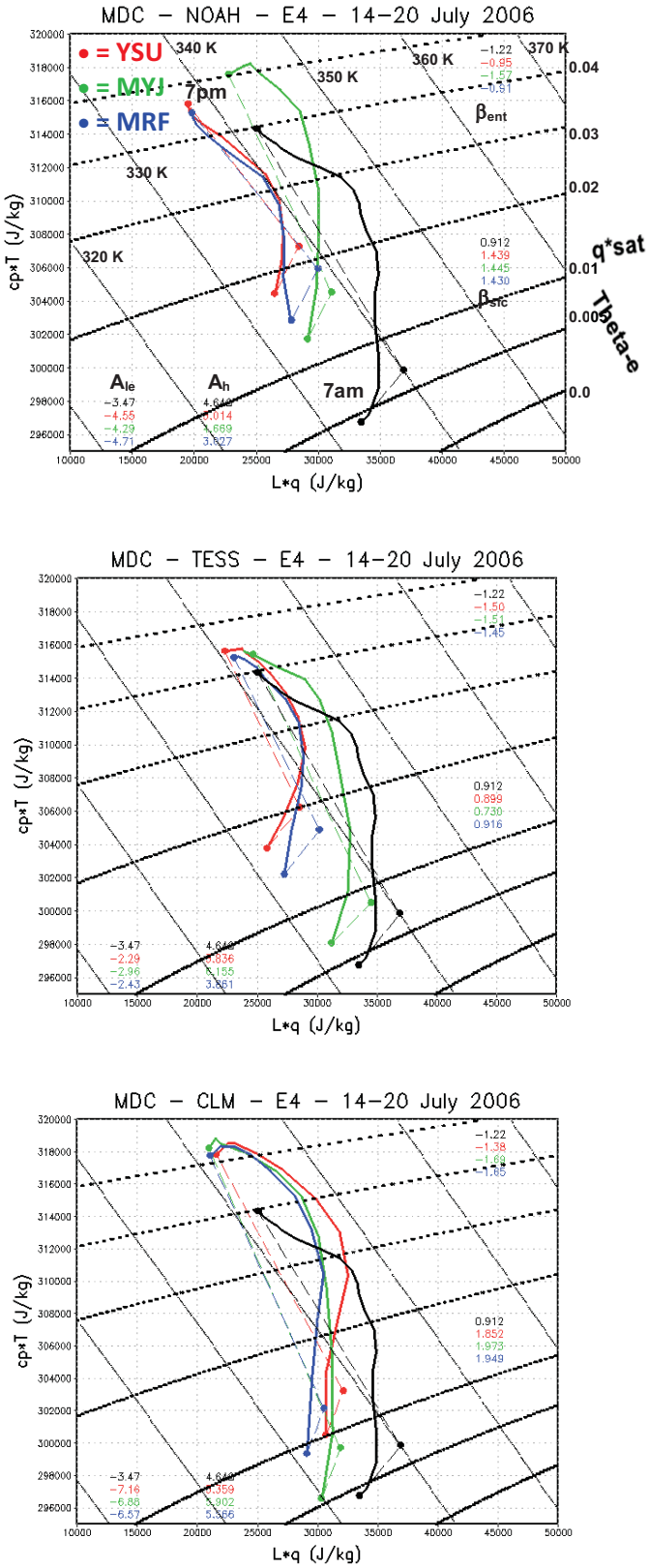
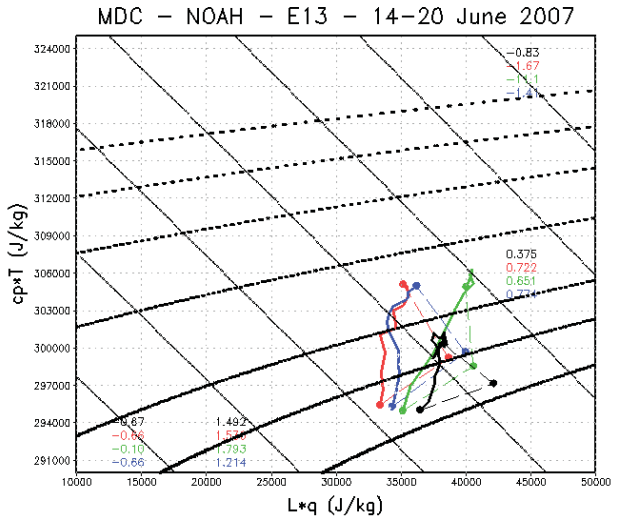
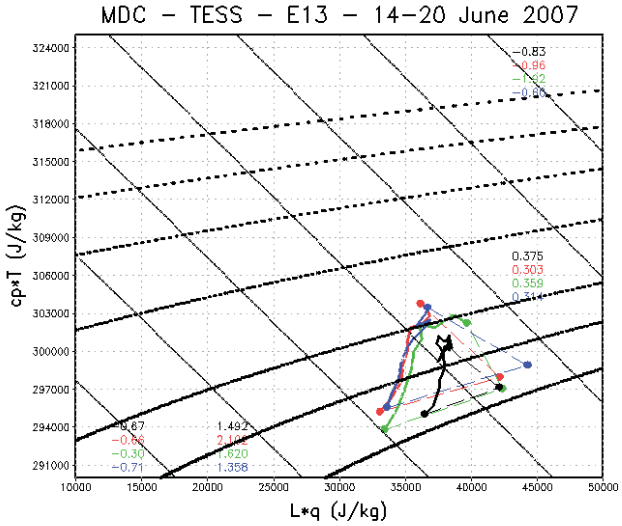


Figure 6. Mixing Diagrams from the LIS-WRF simulations of Noah, TESSEL, and CLM composited over the 7-day case studies of 2006 (14-20 July) and 2007 (14-20 June), and overlain with lines of constant equivalent potential temperature (Theta-e) and saturation humidity (q*sat) deficit.





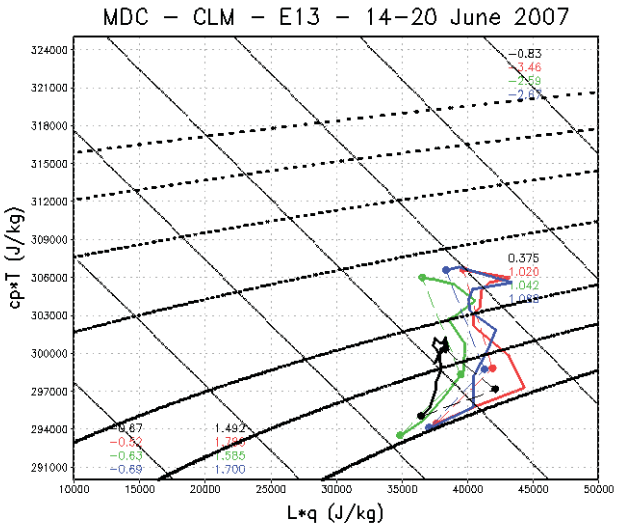
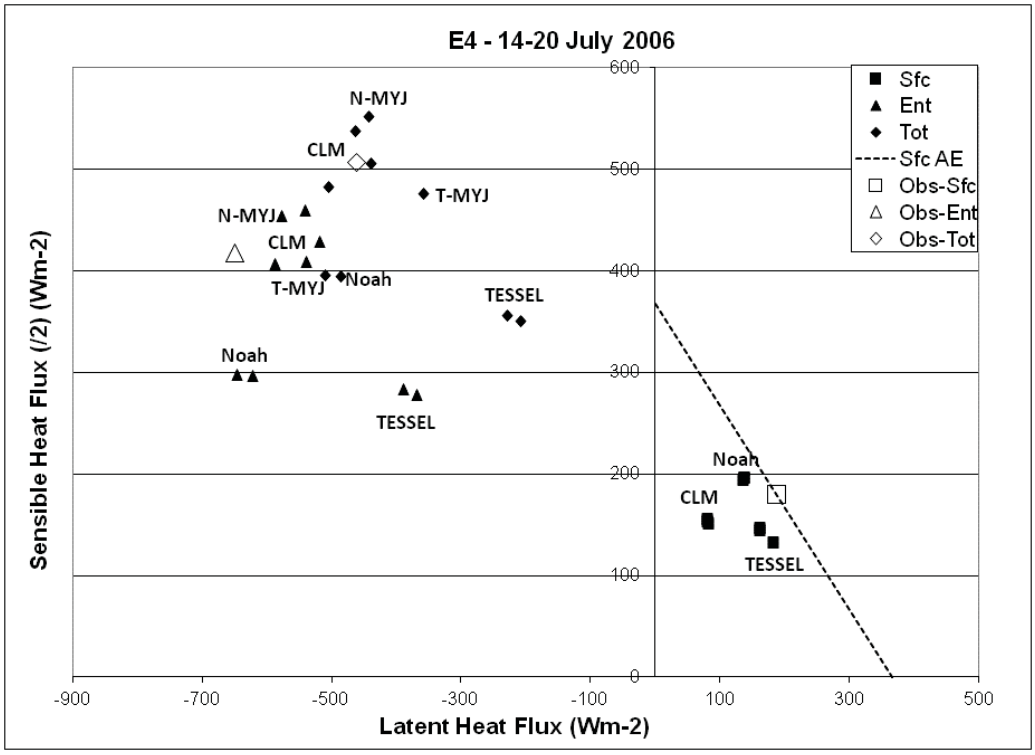


Figure 7. Same as Figure 6, but for the 14-20 June 2007 case.



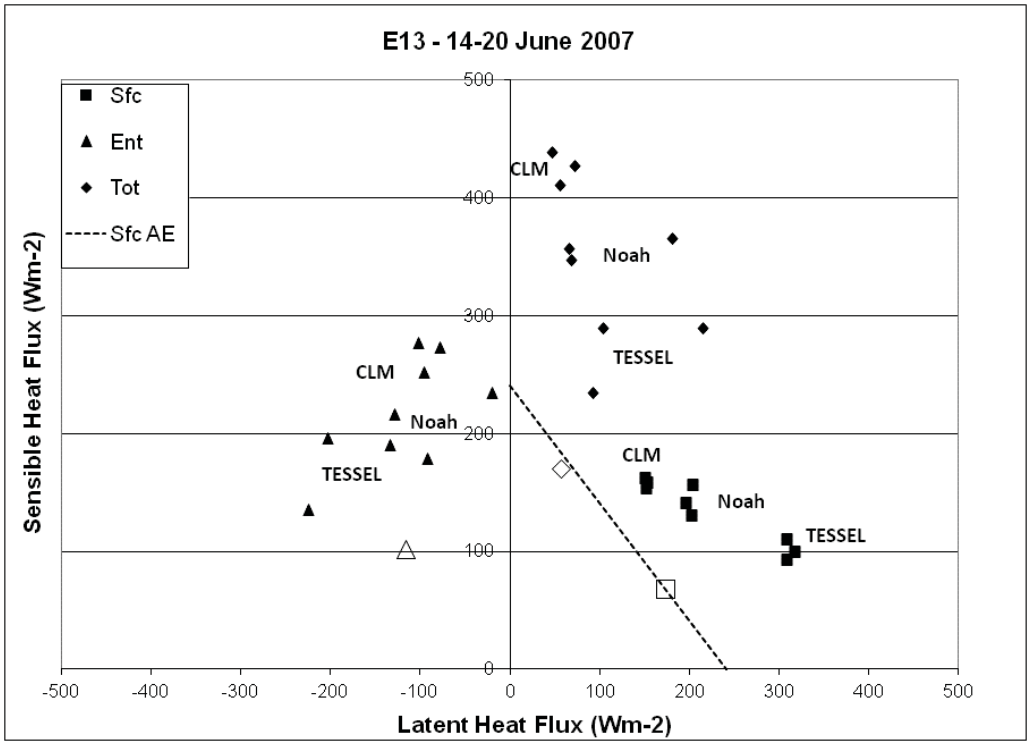


Figure 8. Heat and moisture budgets (SFC, ENT, and TOTAL) from the NU-WRF simulations vs. observed, derived from the mixing diagrams in Figs. 6 and 7 for the **a)** 2006 and **b)** 2007 case studies. Note that for 2006, the sensible heat flux has been scaled by 0.5 for the entrainment and total fluxes.

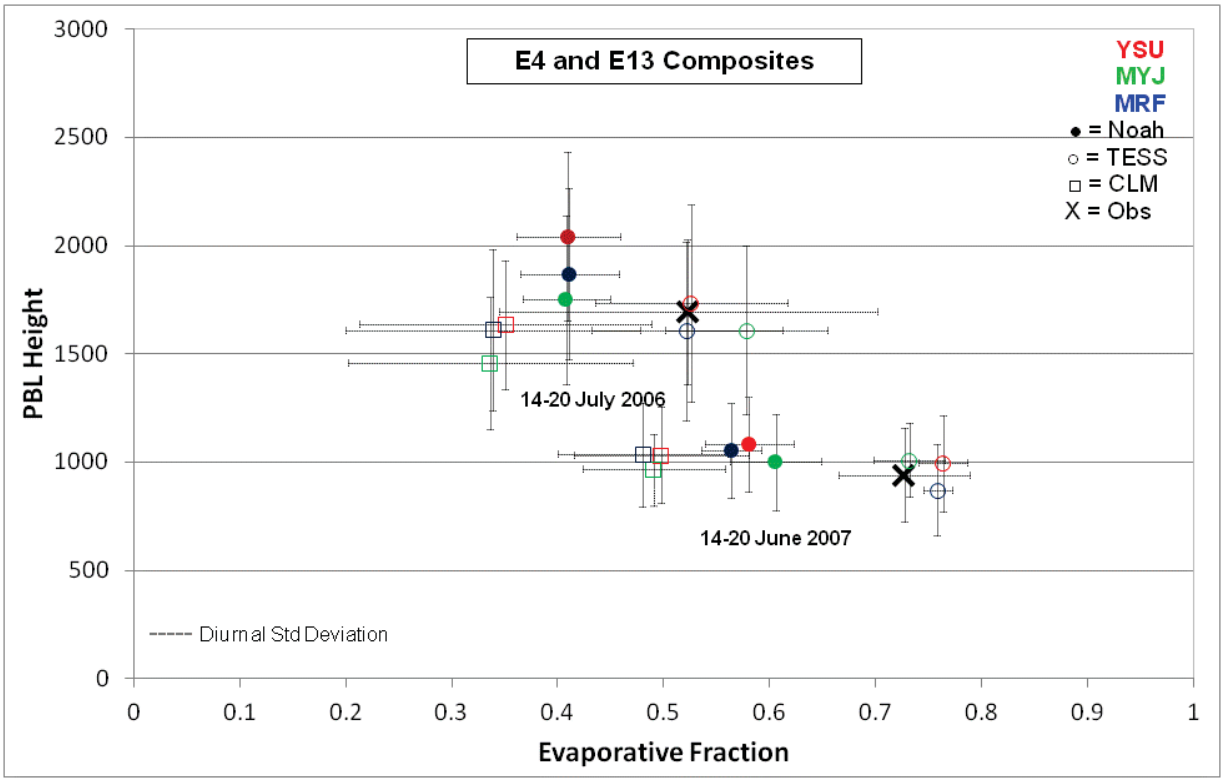
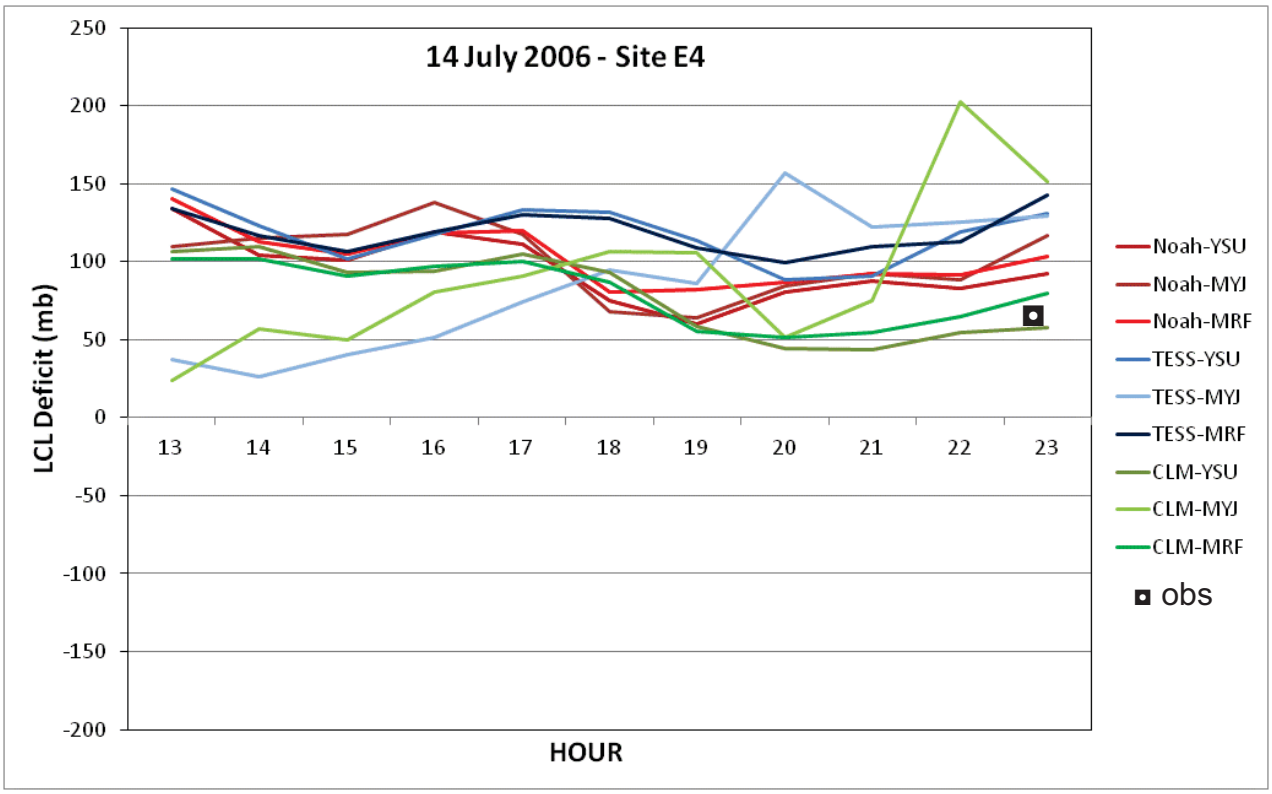


Figure 9. Mean daytime Evaporative Fraction vs. PBL Height for each simulation vs. observed, along with the diurnal standard deviation through the 7-day period



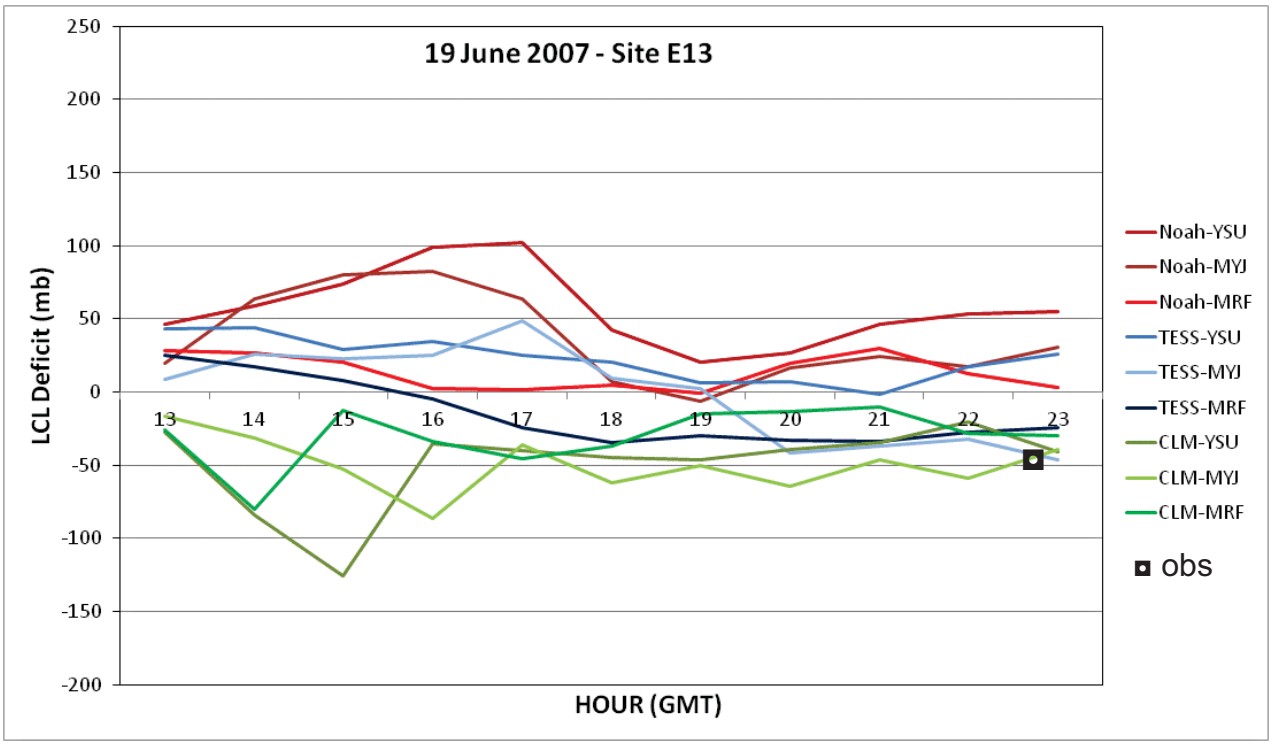
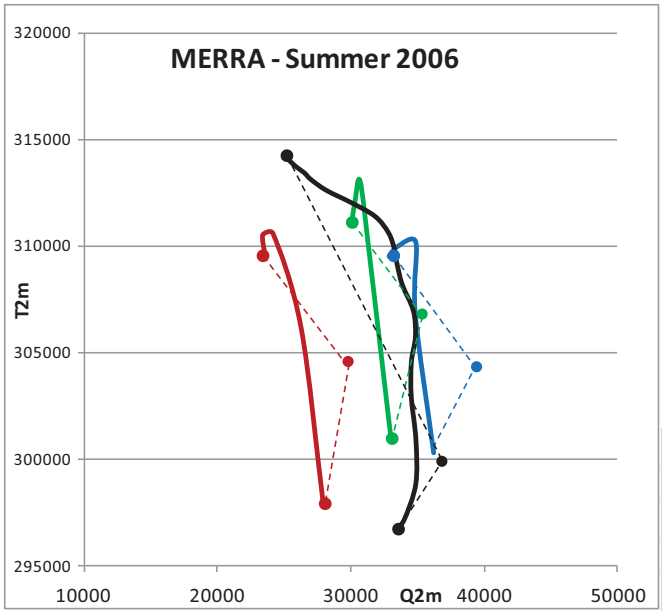
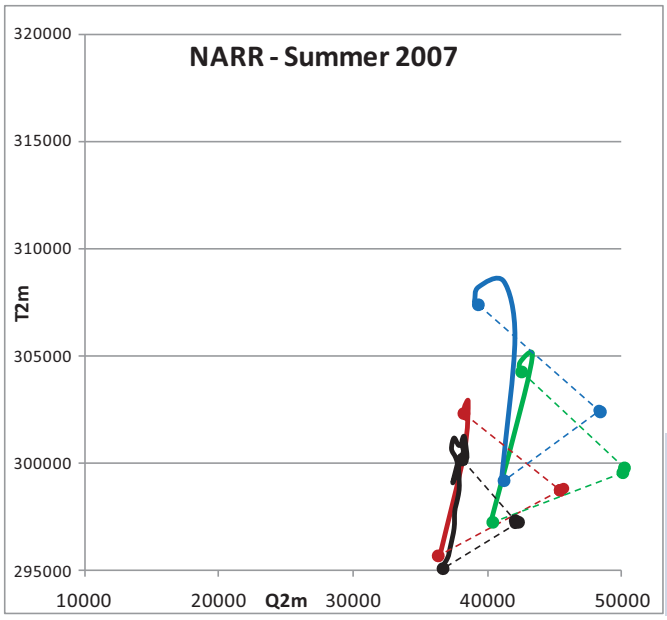


Figure 10. Hourly LCL Deficit (= PBL height - LCL; mb) calculated for each of the PBL-LSM couplings for the **a)** 2006 and **b)** 2007 case studies.

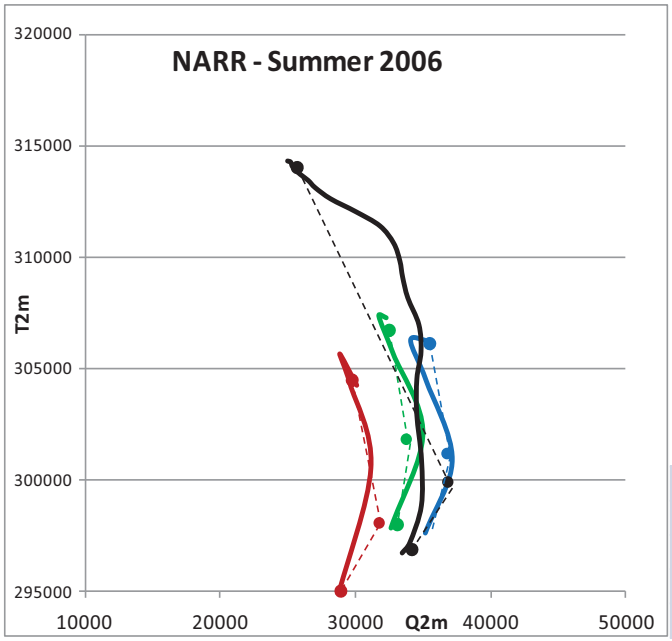


	Jun	Jul	Aug	Obs
Bent	-0.80105	-0.86682	-0.79514	-1.22199
Bsfc	3.589199	2.322341	1.215952	0.912545
Ah	0.741255	0.798984	1.293633	4.64802
Ale	-3.3213	-2.1406	-1.97827	-3.47101

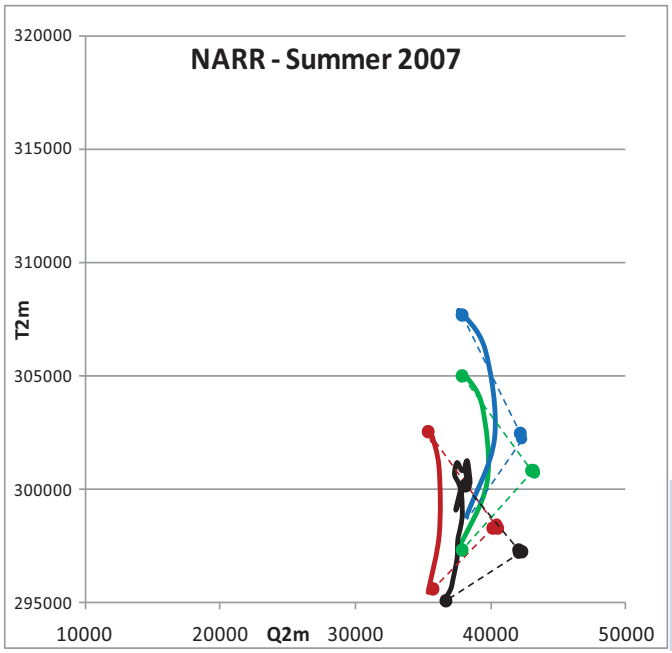


	Jun	Jul	Aug	Obs
Bent	-0.44614	-0.573	-0.54234	-0.83047
Bsfc	0.340416	0.246756	0.426687	0.37522
Ah	1.049836	1.817806	1.629709	1.493067
Ale	-0.80105	-0.78282	-1.28218	-0.67459

Figure 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.



	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

Figure 12. 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 Aaron Kennedy]