# Quantifying the Land–Atmosphere Coupling Behavior in Modern Reanalysis Products over the U.S. Southern Great Plains

JOSEPH A. SANTANELLO JR. AND JOSHUA ROUNDY

*Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland*

PAUL A. DIRMEYER

*Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, Virginia*

(Manuscript received 9 October 2014, in final form 15 April 2015)

ABSTRACT

The coupling of the land with the planetary boundary layer (PBL) on diurnal time scales is critical to regulating the strength of the connection between soil moisture and precipitation. To improve understanding of land–atmosphere (L–A) interactions, recent studies have focused on the development of diagnostics to quantify the strength and accuracy of the land–PBL coupling at the process level. In this paper, the authors apply a suite of local land–atmosphere coupling (LoCo) metrics to modern reanalysis (RA) products and observations during a 17-yr period over the U.S. southern Great Plains. Specifically, a range of diagnostics exploring the links between soil moisture, evaporation, PBL height, temperature, humidity, and precipitation is applied to the summertime monthly mean diurnal cycles of the North American Regional Reanalysis (NARR), Modern-Era Retrospective Analysis for Research and Applications (MERRA), and Climate Forecast System Reanalysis (CFSR). Results show that CFSR is the driest and MERRA the wettest of the three RAs in terms of overall surface–PBL coupling. When compared against observations, CFSR has a significant dry bias that impacts all components of the land–PBL system. CFSR and NARR are more similar in terms of PBL dynamics and response to dry and wet extremes, while MERRA is more constrained in terms of evaporation and PBL variability. Each RA has a unique land–PBL coupling that has implications for downstream impacts on the diurnal cycle of PBL evolution, clouds, convection, and precipitation as well as representation of extremes and drought. As a result, caution should be used when treating RAs as truth in terms of their water and energy cycle processes.

## 1. Introduction

Land–atmosphere (L–A) interactions and coupling remain weak links in current approaches to understanding and improving predictions of the Earth–atmosphere system and its variability in a changing climate. However, recent community-based efforts (e.g., LandFlux; Mueller et al. 2013) have shown that current observational and model products have significant uncertainty and spread in surface [e.g., evapotranspiration (ET)] and planetary boundary layer (PBL) water and energy budget terms at global, continental, and regional scales (Rodell et al. 2015; T. L'Ecuyer et al. 2014, manuscript submitted to *J. Climate*).

In order for improvements to be made in the proper translation of land surface states (e.g., soil moisture) and anomalies (e.g., flood–drought) into atmospheric quantities (e.g., afternoon convection), a greater understanding of coupled model components and physics must be acquired (Betts and Barr 1996; Entekhabi et al. 1999; Guo et al. 2006; Jakob 2010). To this end, the influence of soil moisture on precipitation has been under community-wide investigation in a range of studies from local [via local L–A coupling (LoCo); Santanello 2011] to global (Koster et al. 2004) scales. Biases and errors in individual land and PBL variables can have far-reaching impact across the system, suggesting that understanding of the nature and accuracy of land–PBL coupling is paramount to assessing the full validity and limitations of a model and the subsequent pinpointing of areas for improvement.

Reanalysis (RA) products provide global, continuous, and long-term records of the climate system constructed by fusing together disparate observing networks with

*Corresponding author address:* Dr. Joseph A. Santanello Jr., NASA-GSFC, Code 617, Bldg. 22, Room G220, Greenbelt, MD 20771.
E-mail: joseph.a.santanello@nasa.gov

DOI: 10.1175/JCLI-D-14-00680.1

© 2015 American Meteorological Society

fundamentally sound model physics formulations. As a result, RA has become a core component of nearly all operational weather and climate prediction centers, each generating products with their own advantages and limitations in terms of observations, physics, and temporal and spatial scales. While it is tempting to use any such product as truth because of its complete coverage, it remains of foremost importance to understand the inherent limitations and strengths introduced by diverse model physics, observation types, and assimilation practices before reanalyses can be used to answer energy and water cycle questions over a range of short-term to climate scales.

A suite of modern-era RAs is now available to the public, generally covering the satellite era (1979–present) at global and continental scales. Three RA products that have been used in recent years are the National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR; Mesinger et al. 2006), NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011), and NCEP's Climate Forecast System Reanalysis (CFSR; Saha et al. 2010). Each of these RAs utilizes advanced yet distinct approaches toward integrating observations with the native model physics, as well as varying temporal and spatial characteristics. Fortunately (and in contrast to earlier RAs), the temporal resolution of each has been designed and/or updated to now allow the diurnal cycle to be resolved (and output) at 1–3-hourly increments. Such temporal sampling is critical for L–A coupling to be assessed on the actual time scales of land–PBL interactions.

Previous studies have focused on the accuracy of atmospheric components of RAs (e.g., Becker et al. 2009; Kennedy et al. 2011; Bosilovich 2013) and individual land surface variables (e.g., Decker et al. 2012; Mueller et al. 2013). While these have been useful as a "one at a time" type of evaluation approach and to identify inherent biases, what remains to be explored is how coupled L–A processes in RAs behave synchronously over the diurnal cycle and how they compare with observations of land and PBL variables. Further, it is important to assess whether trends in coupled components change over time or are sensitive to different regimes (e.g., dry versus wet), particularly if RA products are to be used for longer-term extremes (e.g., drought) and climate trend and predictability studies (Thorne and Vose 2010).

To address these issues, this paper quantifies the behavior of L–A coupling in NARR, MERRA, and CFSR products over the last two decades in the U.S. southern Great Plains (SGP). This work builds directly upon previous atmospheric and land-based evaluations of RA products by assessing the connections and feedbacks in the land–PBL system and identifying where diverse physics and assimilation practices in RAs might play a role. This is also the first such assessment of large-scale RA products using LoCo diagnostics, and the first to make use of new land and PBL datasets available from the SGP. Section 2 summarizes the most relevant recent RA intercomparisons, as well as the L–A coupling diagnostics to be employed in the current study. The RA systems and site description are presented in section 3, followed by the results in section 4 and application to ongoing drought investigations in section 5. Finally, conclusions and discussion of complicating issues and implications for future RA development follow in section 6.

## 2. Background

### a. RA intercomparison studies

#### 1) Atmospheric variables

Kennedy et al. (2011) focused on the atmospheric components of NARR and MERRA (viz., temperature, humidity, wind, and cloud fraction) and compared them against observations from radiosondes. The study region was centered over the Atmospheric Radiation Measurement Program (ARM) SGP Central Facility (CF), covering the 1999–2001 period. While their focus was on the full troposphere and did not include the land surface itself, a number of features of the RAs were found that are relevant to L–A interactions. Most notably, MERRA and NARR both exhibited a warm and dry bias near the surface, particularly in late spring and early summer. Both were dry in the PBL, where NARR radiation components were biased (high downwelling shortwave and low downwelling longwave) because of lower-than-observed cloud fraction. MERRA accumulated precipitation had a significant low bias (but was highly correlated with the timing of precipitation), while NARR was closer to observed because of its assimilation of precipitation (see section 3a).

In essence, the work of Kennedy et al. (2011) focused on atmosphere-only coupling processes (e.g., the relationship of relative humidity to vertical velocity and cloud fraction) with the goals of identifying parameterization behavior and skill present in the RAs. While overall temperature and humidity profiles in the troposphere behaved reasonably well, PBL characteristics diverged considerably during spring–summer. The results were also largely based on monthly or longer mean values over this period and did not account for diurnal cycle differences that tend to drive the L–A coupling (e.g., evaporation, PBL growth, and convective triggering). As a result, there are likely impacts of atmospheric biases found by Kennedy et al. (2011; e.g., downwelling shortwave in NARR and precipitation in

MERRA) on diurnal land surface states and fluxes that ultimately feed back upon the thermodynamics of the coupled system.

A similar study by Bosilovich (2013) focused on large-scale temperature and precipitation across the United States and found that the weakest performance (relative to observations) by RAs was found in regions where L–A interactions are likely influential, such as in the Great Plains. In particular, MERRA was found to have too low variance in temperature and precipitation during extremes (e.g., flood and drought years). The impact of L–A interactions was supported as a possible cause of variability and biases in these products as well, though this was hypothesized only through the modeled and observed behavior of temperature and precipitation alone.

#### 2) ATMOSPHERIC AND LAND VARIABLES

To date, there have only been limited assessments of land or PBL variables in RAs. The most detailed study of surface variables was performed by Decker et al. (2012), where six RA products were evaluated over multiple years (encompassing subsets of the 1991–2006 period) across 33 surface flux sites across the globe. Overall, their results showed that the RAs generally perform well with the diurnal cycle of temperature in the summer months (with strong radiative forcing), and even though it does not assimilate 2-m temperature, MERRA performs comparably to ERA-Interim (which does assimilate screen-level variables). However, nearly all RAs overestimate downward shortwave radiation, indicating too little cloud cover. CFSR and MERRA diverge considerably overall, despite utilizing very similar sets of observations in their assimilation. Not all RAs that incorporate observed precipitation perform well with moist processes (e.g., MERRA), but those that do (e.g., CFSR) also produce better latent heat flux as a result. At the same time, CFSR temperature and sensible heat flux perform worse than MERRA, so it is evident that the better precipitation and evaporation do not necessarily lead to better L–A coupling (where evaporative fraction might be a better diagnostic in this regard).

Another important finding of Decker et al. (2012) was that contributions to the overall errors in each RA were weighted much more heavily toward biases in the monthly mean data versus correlations in the 6-hourly data. This highlights the importance of assessing both the long-term–seasonal means and the diurnal cycle in terms of model evaluation and development practices.

Similar work performed by Yi et al. (2011) focused solely on MERRA atmospheric and land variables evaluated against global satellite products, models, and limited in situ data. Mixed results were found in terms of radiation, soil moisture, and surface temperature, with the most noticeable limitation of MERRA being found in its underrepresentation of the subgrid and diurnal cycle representation of precipitation and the subsequent impacts on soil moisture and evaporation. As a result, Reichle et al. (2011) have produced an offline land analysis called MERRA-Land that was driven by MERRA forcing with improved (observed) precipitation correction and land surface model parameter changes related to canopy interception.

L–A interactions themselves were the focus of an investigation by Dirmeyer (2013) in the context of CFSR and associated reforecasts of CFSv2. Distinct differences in the RA versus forecast model were found in terms of water cycle variable climatologies and the representation of the relationship between soil moisture and precipitation. Likewise, Shah and Mishra (2014) intercompared the ability of CFSR, MERRA, and ERA-Interim to represent drought during the Indian monsoon and found biases in temperature and mischaracterization of drought extent and area (i.e., precipitation patterns). Incorrect model translation of the land state to precipitation (i.e., L–A coupling) is likely a limiting factor in these RA applications, yet one that can be discerned only with more process-level analysis of the land–PBL model components.

As in Kennedy et al. (2011), the studies above are rather thorough but also represent a limited analysis of each variable without consideration of their coupled or diurnal evolution. They also urge caution on using RAs for subdaily time scales and applications (despite the recent availability of data on these scales). It is therefore quite timely now to drill down into a thorough evaluation of L–A coupling in RAs at diurnal time scales, focusing on a single well-observed location and metrics of the full land–PBL interaction.

### b. *L–A coupling and diagnostics*

The initial communication between the land and atmosphere occurs on local scales through the PBL. A community effort supported by the GEWEX Global Land–Atmosphere System Study (GLASS; van den Hurk et al. 2011) panel has therefore been ongoing to diagnose and quantify local L–A coupling in models. A thorough review of LoCo research and the related diagnostic approaches can be found in Santanello (2011) and Santanello et al. (2009, 2011, 2013b).

As discussed in Santanello et al. (2011), a full understanding and quantification of land–atmosphere interactions will be reached only through careful examination of a series of interactions and feedbacks (i.e., "links in the chain") between soil moisture (SM) and precipitation ($P$). These relationships depend on the sensitivities of (i)

surface fluxes of sensible and latent heat to soil moisture, (ii) PBL evolution to surface fluxes, (iii) entrainment fluxes to PBL evolution, and (iv) the collective feedback of the atmosphere (through the PBL) on surface fluxes (Santanello et al. 2007; van Heerwaarden et al. 2009). LoCo diagnostics focused on these interactions are therefore well suited to assess the fully coupled behavior of models in terms of the simultaneous evolution of land and PBL processes. Rather than single-variable evaluations, where compensating errors are often hidden and causality can be difficult to ascertain, coupled metrics can be used to learn about model differences and deficiencies in a systematic fashion.

Most notably, Santanello (2011) and Santanello et al. (2009, 2013b) developed a model intercomparison methodology based on Betts's (1992) ''mixing diagram'' theory. The power of this approach lies in its ability to exploit the covariance of 2-m potential temperature (T2) and humidity (Q2) to quantify the links in the chain of the soil moisture–precipitation relationship and in that it is based only on routine variables that can be applied to any model or observation product. How anomalies and/or errors in the surface fluxes computed by a particular model or scheme combination are translated into the atmospheric water and energy cycle can then be quantified using this approach. By extension, the relationship of evaporative fraction (EF; ratio of latent heat flux to the sum of latent and sensible heat fluxes) to PBL height (EF versus PBL Height; Santanello et al. 2009) and the lifting condensation level (LCL) deficit (Santanello 2011) are complementary metrics that further tease out the relationship between land and PBL processes.

Because they can be applied universally to any model or observation source, LoCo diagnostics are ideal metrics to evaluate the impact of varying physics and observations in RAs on L–A coupling.

Another branch of L–A diagnostics stems from the work of Findell and Eltahir (2003a,b) and utilizes measures of atmospheric stability [convective triggering potential (CTP)] and humidity [low-level humidity index (HI)] to classify coupling into regimes. The Findell and Eltahir (2003a) regime classification is based on the premise that convective precipitation is more likely for wet–dry soils as a function of the initial atmospheric state (CTP–HI space; i.e., wet–dry soil advantage regimes). Roundy et al. (2014) then extended this via a data-driven coupling classification called the coupling drought index (CDI) that allows for global application across models and datasets. The CDI is defined as the number of dry coupling days minus the wet coupling days divided by the total number of days for the evaluation period and ranges from $-1$ (all wet coupling) to $+1$ (all dry coupling). In addition, the CDI is based on all days (not just days with convective precipitation) in order to capture the overall wetting–drying feedback between the boundary layer and land surface, not just that associated with precipitation.

## 3. Reanalysis and site description

The three RAs chosen for this study have been used quite frequently in recent years, particularly for studies over the United States. Moreover, they share similarities in terms of using regular observations of temperature, humidity, pressure, and wind speed from radiosonde, as well as some variation of observed precipitation. It should be noted that none of the three RAs assimilate screen-level variables, in contrast to the ERA-Interim analysis approach using 2-m temperature that acts to confound assessment of the true (observable) coupling. A brief description of the RA products with relevant L–A coupling components will be presented here, with the reader being referred to the many core references of each for further information.

### a. NARR

NARR is an update of the former NCEP Reanalysis system, with a focus on producing improved precipitation through assimilation (Kennedy et al. 2011). To this end, NARR has shown improvement over the NCEP Reanalysis for a variety of variables, including precipitation, diabatic heating, and temperature (Mesinger et al. 2006). The core of NARR is the Eta atmospheric model, and the PBL and land surface model (LSM) physics employed are the Yonsei University (YSU; Hong et al. 2006) PBL and Noah LSM (Ek et al. 2003) schemes, respectively.

The NARR period of record is from 1979 to the present and is run at 32-km horizontal resolution with 45 layers in the vertical. Output is every 3 h at 29 levels, containing a full suite of state and flux variables across the L–A interface. One unique aspect of NARR is the continental domain centered over the United States, which enables higher resolution overall and more mesoscale features to be resolved relative to coarser RAs and GCMs. NARR has also been used frequently as initial/boundary conditions for community mesoscale models.

### b. MERRA

MERRA is based on NASA's Goddard Earth Observing System, version 5 (GEOS-5; Rienecker et al. 2011) and has global coverage at $0.5° \times 0.667°$ horizontal and 72-layer vertical resolution for the period 1979–present. MERRA was designed to optimally exploit and assimilate satellite-based datasets such as those from NASA's Atmospheric Infrared Sounder (AIRS). MERRA, as a result, assimilates numerous satellite data streams to

better constrain Earth's energy and hydrologic budgets (Bosilovich 2013), including that of instantaneous rain rates. The atmospheric core of MERRA is the GEOS-5 AGCM, and the land surface is represented by the Catchment LSM (Koster et al. 2000). The PBL scheme is the same as that used in GEOS-5: a combination of the first-order Louis et al. (1982) and Lock et al. (2000) turbulent schemes for stable and unstable cloud-topped PBLs. A unique aspect of MERRA is that it archives output on its native grid as opposed to other RAs whose outputs are typically reduced in spatial and temporal resolution before release. A key aspect for this study is that both 2D diagnostics and monthly mean diurnal cycles are produced hourly and include the variables required by the LoCo metrics described in section 2b.

### c. CFSR

The most recent RA to be developed of the three is CFSR, based on NCEP's CFSv2. CFSR is also global, with a horizontal resolution of T382 spectral truncation (0.313°) with 64 layers in the vertical, but the atmospheric multilevel data are archived at a 0.5° resolution with a record length from 1979 through March of 2011. Beginning April 2011, the CFSv2 was implemented into operations at which point an updated version of the CFSv2 (in terms of some of its physical parameterization) was used to produce the real-time CFSR product through the present (Saha et al. 2014). A key distinction of CFSR is its use of a coupled ocean model as opposed to specified sea surface temperatures in MERRA and NARR. CFSR also employs radiance and product-assimilation techniques similar to that of MERRA and at 6-h increments.

The PBL [Medium-Range Forecast Model (MRF); Hong and Pan 1996] and LSM (Noah) physics are similar to those in NARR, though with different versions of each. Observed precipitation, a combination of gridded and gauged data, is used to force the LSM as opposed to modeled precipitation, thus better constraining the surface energy balance and hydrology components. Output from CFSR is typically a combination of 6-hourly analysis cycle plus model forecasts at hourly intervals. Specifically for this study, monthly mean diurnal cycle output was produced by NCAR at 1-h resolution, and includes the full set of L–A variables needed for LoCo analysis (NCEP Environmental Modeling Center 2010).

### d. Observations at SGP

The SGP has been identified as a hot spot for L–A coupling in terms of the strength of interactions and potential impact of soil moisture anomalies on clouds and precipitation (e.g., Koster et al. 2004). Since the mid-1990s, the U.S. Department of Energy (DOE) has maintained a large, continuous record of observational data from ARM SGP (covering a large part of Oklahoma and Kansas). As a result of the unique wealth of data in this region, particularly for surface and atmospheric variables, the SGP has also been a hot spot of modeling and process studies (e.g., Santanello et al. 2007, 2013b; Zhang and Klein 2010; Phillips and Klein 2014).

The combination of long-term measurements of components of the LoCo process chain (Santanello et al. 2011) makes it a unique site to examine L–A interactions and to perform model evaluation and intercomparison studies. These data include soil-moisture, radiation, sensible, latent, and soil-heat fluxes, along with collocated surface meteorology data and profile data from radiosonde and lidar. Further, a new ARM best estimate (BE) product (ARMBE-Land; Xie et al. 2010, 2012) has recently been developed in conjunction with the LoCo community that synthesizes continuous, quality-controlled data for the 1996–2012 period at the ARM Central Facility. Combining the ARMBE-Land with the recently released ARMBE atmosphere (ARMBE-ATM) data thus now provides an hourly record of L–A observations for this period (McCoy and Xie 2012). [Note that surface flux data were unavailable during June, July, and August (JJA) of 2001, and therefore in Figs. 1 and 3 the 2001 observed surface fluxes were taken from the MERRA estimates as an approximation.]

The LoCo analysis was performed using collocated surface meteorology and flux tower observations in the ARMBE products. Fluxes were collected from Bowen ratio (EBBR) and eddy correlation (ECOR) flux tower measurements, whose maximum error ranges for latent and sensible heat flux are ~10% for EBBR with perfect closure (by definition) and ~(5%–6%) for ECOR with 75%–90% closure (Wilson et al. 2002). For PBL height estimates, the CF radiosonde data was used along with a new hourly profile product available from ARM. This new product (MERGESONDE; Troyan 2012) is a combination of radiosonde, profiler, and model (ECMWF) data, but heavily weighted toward in situ (i.e., radiosonde) measurements while allowing for hourly profiles to be produced. This is a significant advancement in our ability to describe the diurnal cycle of the PBL and to implement LoCo metrics at their process level. MERGESONDE data have recently been improved in representation of PBL profiles based on feedback from this study, and a new version released in September 2014 is used herein.

### e. Experimental design

Monthly mean diurnal cycles from the RA products were derived from the highest native temporal and spatial resolution available, as described above. This was

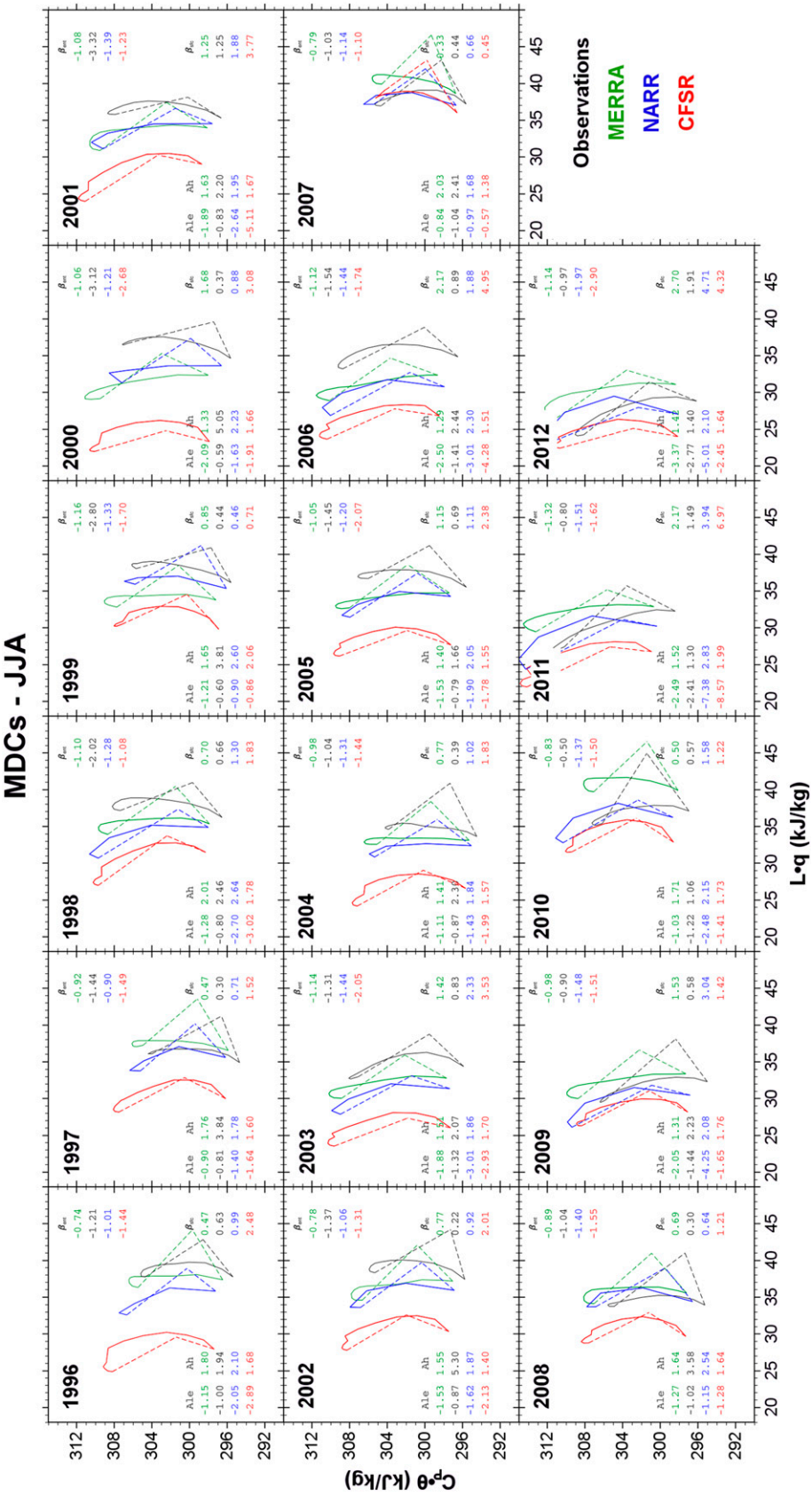


FIG. 1. Mixing diagrams calculated from JJA monthly mean diurnal cycles of MERRA (green), NARR (blue), and CFSR (red), along with observations (black) from the ARM SGP CF over the 1996–2012 period. Dashed lines indicate the surface and atmospheric flux vectors. Also included are the surface ($\beta_{sfc}$) and atmospheric ($\beta_{atm}$) Bowen ratios, and sensible ($A_h = Qh_{atm}/Qh_{sfc}$) and latent heat ($Ale = Qle_{atm}/Qle_{sfc}$) flux ratios. See Santanello et al. (2009).

1 hourly for MERRA and CFSR and 3 hourly for NARR. The grid cell of each RA closest to the ARM CF (36.605°N, 97.485°W) was then found. To correspond with the ARMBE record, the period 1996–2012 was analyzed and composited over the summer months (JJA) such that there is a single diurnal cycle from each RA that is representative of each summer. The ARMBE data was then averaged up to match that of the JJA monthly mean diurnal cycles of the RA products. This required averaging up to mean monthly and mean JJA values for 2-m temperature, humidity, latent and sensible heat fluxes (all from ARMBE), and PBL height estimates (via a critical bulk Richardson number approach) from the MERGESONDE data. Finally, LoCo metrics were applied to each RA and the observations in order to produce the analysis. Note that for the mixing diagram analysis, we have followed Santanello et al. (2013a) in redefining the residual vector in the diagrams (formerly the ''entrainment vector'' as in Santanello et al. 2009) as the ''atmospheric response vector'' ($\mathbf{V}_{atm}$) to more precisely reflect the inherent assumptions in this approach.

As this study represents the first time that LoCo metrics have been used to diagnose coupling at seasonal time scales, the LoCo results are also compared with the CDI, which has been previously used to diagnose coupling in reanalysis and forecast models at seasonal time scales for continental domains (Roundy et al. 2013, 2014) and provides a unique comparison for the ARM CF that lends spatial perspective to the findings.

## 4. Results

The coupled behavior of the RA products can be assessed using the suite of LoCo diagnostics that focus on the diurnal cycle of land–PBL variables. To this end, this represents the first application of the LoCo approach to seasonal and interannual cycles and to large-scale models and RA products and makes use of new land and PBL observation products from the SGP.

### *a. Mixing diagrams*

The annual (1996–2012) summertime diurnal evolution of T2, Q2, surface fluxes of latent ($Qle_{sfc}$) and sensible ($Qh_{sfc}$) heat, and atmospheric fluxes of latent ($Qle_{atm}$) and sensible ($Qh_{atm}$) heat can be seen in the mixing diagram analysis of Fig. 1. The interannual spread and variability among RAs and observations are evident, and a number of patterns emerge. CFSR is distinguished by being consistently the least humid in terms of Q2 (with the exception of 2007). This is a result of a drier surface condition in CFSR, which leads to large surface Bowen ratios ($\beta_{sfc}$) that are often double those from the other RAs and more than 3 times those observed. MERRA and NARR are more similar in terms of T2 and Q2, but NARR tends to be a bit drier overall with correspondingly higher surface Bowen ratios. However, beginning with the wet year of 2007, MERRA becomes most humid while NARR is closer to observations (with CFSR remaining dry, but less so).

All RAs tend to have similar ranges in T2 and an overall diurnal signature consistent with observations. In terms of Q2, CFSR has a noticeable curvature in the diurnal evolution relative to the other RAs, indicating that there is a more demonstrable moistening of the PBL in the morning hours due to evaporation (capped by a slowly growing PBL) that evolves into stronger drying in the afternoon corresponding with more rapid PBL growth and entrainment. Without the new hourly RA output available from CFSR, this signature of the interplay of evaporation, PBL growth, and entrainment here would not be evident. NARR and MERRA tend to be more linear in Q2 evolution in most years, as supported by more rapid morning PBL growth and less buildup of PBL moisture (not shown). The afternoon drying of the PBL through entrainment becomes more evident in all RAs and observations in extremely dry years (e.g., 2011 and 2012).

Analogously, during wet regimes (e.g., 2007) the wet surface and limited growth of the PBL produce a much smaller diurnal range in T2 and Q2 with very little afternoon drying. In addition, the RAs collapse on each other in terms of state and flux components to values very near the observations. That they behave similarly (and close to observed) during this wettest year of the period suggests that the atmosphere-limited regime is represented well in the RA products, including the dominance of clouds and precipitation over land surface and PBL forcing. This also supports the idea that land surface model impacts are largest during dry regimes (e.g., 1996, 2000, and 2006), where the spread across RAs becomes larger because of different surface and PBL physics allowing for thermodynamic divergence in the coupled land–PBL system (Santanello et al. 2013b).

The energy space representation of mixing diagrams (as depicted in Fig. 1) allows for the calculation of summary statistics in terms of the T2 and Q2 evolution in RAs. Figure 2 shows the root-mean-squared error (RMSE) and bias statistics for each RA over the 1996–2012 period, as calculated from hourly differences in T2 and Q2 between RA and observations (i.e., the cumulative difference in the curves in Fig. 1 in energy units). All three RAs show a warm, dry bias throughout, with the exception of post-2007 MERRA. CFSR stands out in terms of having the largest RMSE (Fig. 2a) values, particularly prior to 2007. The source of the error lies primarily in the dry bias of CFSR (Fig. 2c), as discussed above, relative to the smaller

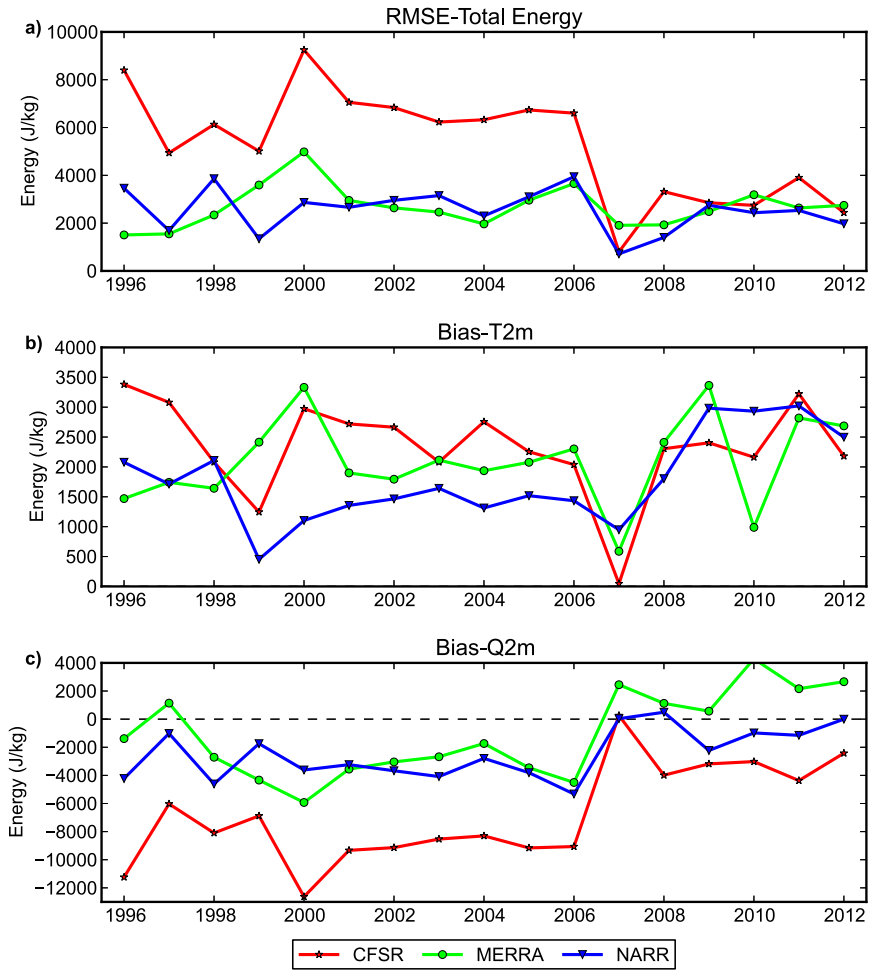


FIG. 2. Summary statistics of (a) RMSE of total energy (J kg$^{-1}$) and bias of (b) 2-m temperature (T2m; J kg$^{-1}$) and (c) 2-m specific humidity (Q2m; J kg$^{-1}$), as derived from the mixing diagrams in Fig. 1 for the RAs over the 1996–2012 period.

magnitude of the warm bias (Fig. 2b). NARR performs slightly better than MERRA overall, but all RAs see a post-2007 shift to lower Q2 bias and higher T2 bias (particularly in NARR).

This reduction in Q2 bias is likely due to a combination of a drier climatological regime setting up over the SGP after 2007 along with changes to the RA systems themselves. Analysis of observed Q2 time series over the period (not shown) shows a distinct shift to rapidly drying regimes post-2007 (evident in the mixing diagrams as well), which results in the observations becoming closer to the climatologically drier RAs. The larger impact on CFSR is also likely a result of changes to the assimilation in this time period that allowed for improved AIRS field of view and, more importantly, the addition of Infrared Atmospheric Sounding Interferometer (IASI) measurements into the CFSR assimilation (Saha et al. 2014).

### b. *Evaporative fraction versus PBL height*

The bulk behavior of the land–PBL coupling can be summarized in the relationship of summertime mean EF to maximum PBL height (PBLH; Fig. 3). There is often a large stratification of both EF and PBLH across the RAs, with the dry bias of CFSR most evident in the form of the low evaporation and large PBL growth. As discussed above, MERRA is the wettest of the three RAs as reflected in high EF and low PBLH, but all three RAs tend to be drier than observations. In fact, MERRA is characterized by significantly lower dynamic range in PBLH than CFSR and NARR throughout, even in very dry regimes (e.g., 2011) as MERRA is near 2 km while the others are closer to 4 km. MERRA still responds consistently in terms of relatively higher PBLH during dry years and vice versa, however. Interestingly, NARR produces high PBLH values that are often comparable

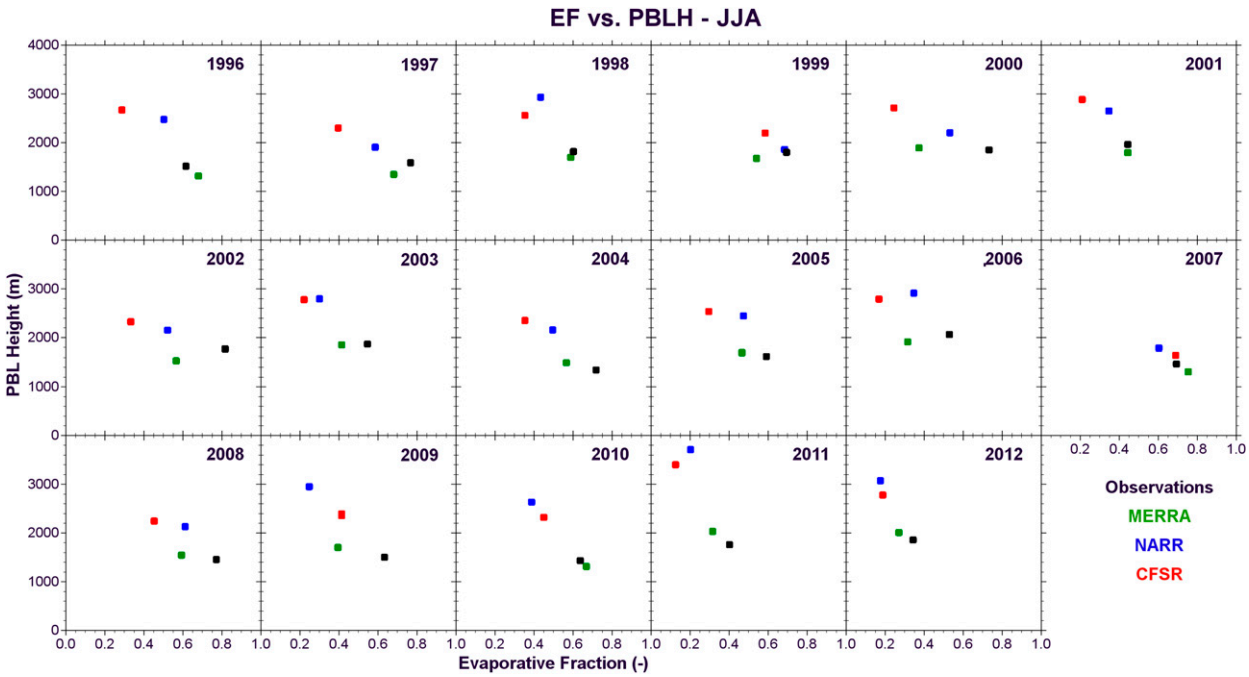


FIG. 3. The relationship of JJA mean daytime mean evaporative fraction (EF) vs JJA maximum PBL height (PBLH; m) for the three RA products—MERRA (green), NARR (blue), and CFSR (red)—and observations (black) at the ARM SGP CF over the 1996–2012 period.

to or exceeding that of CFSR, despite NARR being more moist in terms of Q2 and EF. That indicates higher sensitivity to evaporation in NARR that allows for deeper PBL growth at intermediate soil moisture values (and soil-limited regimes).

To further parse out the EF–PBLH relationship in the RAs, the individual monthly (June, July, and August) values from their full period of record (1979–2012) are plotted for each in Fig. 4a. The overall lower PBLH is seen in MERRA, as well as many more points above 0.5 EF compared to NARR (near 0.5) and CFSR (mostly below 0.5). The larger PBLH in NARR at intermediate EF is also more apparent in the monthly values, particularly in the 0.4–0.5 EF range. As a result, the correlation ($R^2$) of the EF–PBLH relationship is lowest for CFSR (0.72), while MERRA is considerably higher (0.87). The slope of each exhibits the opposite behavior, with the more tightly constrained MERRA having a lower slope than NARR and CFSR due to its lower dynamic range in PBLH. A nonlinear relationship of SM with PBLH was found in Santanello et al. (2005, 2007), particularly for dry values, suggesting CFSR is able to resolve these anomalies better than the other RAs (and that the monthly means do lose some of the nonlinear signal).

Monthly anomalies in EF and PBLH (not shown) are consistent with the trends seen in Fig. 4a in the monthly means. As expected, there are larger negative EF anomalies in MERRA and PBLH in CFSR because their typical values reside at the high end of the spectrum in each. The tendency toward larger dry anomalies is also evident in all three RAs and supports that there were more frequent drought years (e.g., 2003, 2006, 2011, and 2012) in the SGP region over the last 17 years than anomalously wet years (e.g., 2007).

### c. *LCL deficit*

The locally forced impact of the land–PBL coupling on the potential formation of clouds and precipitation can be summarized using the diurnal cycle of the LCL deficit. If the PBL reaches the LCL (indicated by a negative LCL deficit in millibars), then the potential for condensation and cloud development exists. Santanello et al. (2011, 2013b) have also demonstrated that a negative LCL deficit is correlated closely in space and time with cloud cover in a mesoscale (1-km resolution) model.

Because of its dry characteristics, CFSR rarely approaches the zero level of LCL deficit (again, with the exception of 2007), and exhibits a distinct diurnal cycle of high values in the morning rapidly decreasing in the midday[~(1400–1800) UTC] period and then flattening out in the afternoon. NARR sees a brief morning increase in LCL deficit (1400 UTC) followed by a prolonged decrease through the rest of the day, often reaching negative values in the late afternoon (2200 UTC). MERRA shows a much more consistent LCL deficit throughout the day, with the actual value sensitive to the severity of the dry

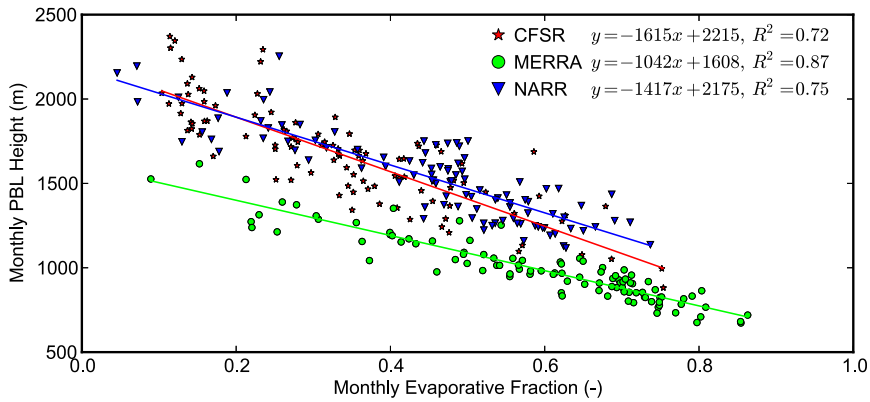


FIG. 4. Monthly (for June, July, and August) mean values of daytime evaporative fraction (EF) vs maximum PBL height (PBLH; m) over the 1979–2012 period: MERRA (green), NARR (blue), and CFSR (red).

(high LCL deficit; e.g., 2006) or wet (low LCL deficit; e.g., 2007) regime. CFSR is often highest of the RAs in the morning, and overall MERRA is consistently higher than NARR (which is the lowest of the three especially in the afternoon).

The mechanisms behind these patterns can be explained by the interplay of the humidity and PBL growth inherent in the RAs. In CFSR, the dry condition leads to generally higher LCLs than the other RAs in the morning, but the combination of the morning moistening of the PBL (seen in Fig. 1) followed by more rapidly growing PBL during midday quickly reduces the LCL deficit. The switch to a dry air entrainment–dominated regime in the afternoon then ensures that the LCL and PBLH balance each other out in the afternoon. The slow decay of the LCL deficit in NARR is consistent with the linear T2, Q2, PBLH, and entrainment shown in Fig. 1 for NARR, where the rather deep PBL growth seen in Fig. 4 reaches progressively closer to the LCL (which is consistent throughout) over time. MERRA's flat line signature is typical of its linear T2 and Q2 evolution combined with more limited PBL growth.

Based on these results, one would expect a strong diurnal signal of clouds and potentially convection in NARR as opposed to MERRA or CFSR. Yet again, 2007 is an exception as it is an atmospherically controlled wet regime, which is almost consistently saturated near the top of the PBL during the entire daytime period in all three RAs. Note that the absence of negative LCL deficit values does not preclude the possibility of clouds and precipitation on individual days on the JJA averaging period. That the RAs are close to zero (e.g., 2010) is enough to suggest moist processes dominated a good portion of the JJA period. Likewise, LCL deficits near or over about 100 mb (1 mb = 1 hPa) (e.g., 2011 and 2012) clearly indicate a dry–drought regime and clear skies for most days of those years.

Observed LCL deficit in Fig. 5 is driven principally by low LCL values due to high Q2 and low T2 that are easily exceeded by even the moderate PBLH estimates from the MERGESONDE data. The diurnal signature is reflective of the T2 and Q2 evolution in Fig. 1 (morning moistening and afternoon drying) as well. After the SGP shifts to a drier regime in 2008, LCL deficit values are more comparable to those from the RAs. It should also be noted that the flux, T2, Q2, and PBLH observations (each from distinct instrumentation and data streams) all agree that the SGP was more humid than the RAs for the bulk of the 17-yr period.

### d. *Diurnal precipitation*

Figure 6 shows the total hourly precipitation generated by each RA over the daytime period. MERRA shows a distinct midday maximum of precipitation near 1800 UTC that is persistent even in the dry extremes (2011 and 2012), consistent with earlier findings from MERRA and other GCMs. The peak of MERRA is typically much higher and sharper than those of NARR and CFSR, which exhibit less consistent patterns in precipitation. NARR is more linear throughout the day but does show indications of an afternoon rise in precipitation after 2000 UTC. CFSR is much more variable and lower magnitude because of its dry condition, especially in late afternoon when it is often near zero. However, CFSR does occasionally show a signal of early morning (e.g., 1999) and midafternoon (e.g., 2002) precipitation. Observations show a generally bimodal behavior of precipitation with maxima in the morning and afternoon (e.g., 2000 and 2002) and are not reproduced well by any of the RAs. In particular, MERRA overestimates precipitation in both magnitude and frequency.

The more nuanced diurnal cycle of clouds and convection can be seen in NARR and MERRA using this analysis

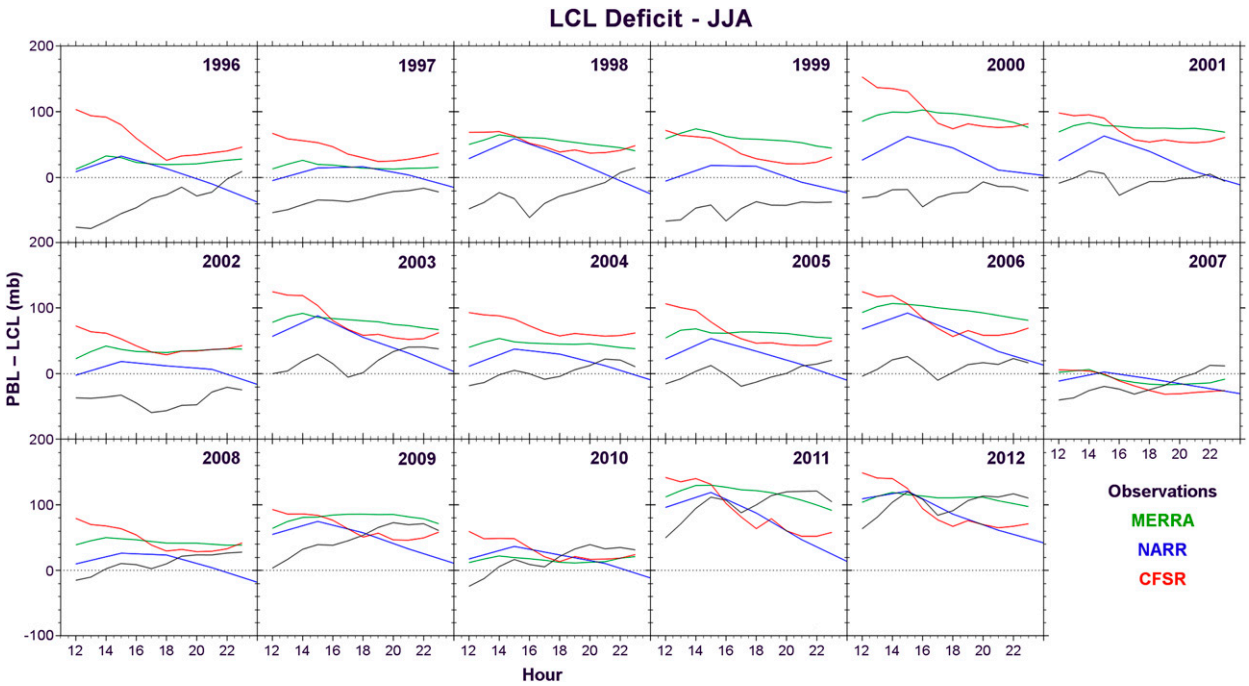


FIG. 5. Average hourly (UTC) daytime cycle of the lifting condensation level deficit (LCLdef; mb) from the three RA products—MERRA (green), NARR (blue), and CFSR (red)—and observations (black) at ARM SGP CF, as calculated from JJA monthly mean diurnal cycles over the 1996–2012 period. LCLdef is defined as the difference between the height of the PBL and the LCL; negative values indicate the LCL has been reached at that time.

approach. However, these results suggest that LCL deficit may not be a good indicator for the driving mechanisms of the diurnal cycle of precipitation in a model such as MERRA that exhibits weaker land–PBL sensitivity (e.g., Fig. 4) and is relatively wetter climatologically (versus NARR and CFSR) such that atmospheric controls are likely more dominant (as opposed to L–A influence).

### *e. L–A feedbacks*

The combination of diagnostics in Figs. 1–6 provides insight into possible feedbacks between the land and PBL that support the overall climatology of the RAs. In MERRA, a more humid surface condition leads to lowering of the LCL but is outweighed by the more limited PBL growth in producing a largely positive LCL deficit throughout the day. In this context, the results suggest a negative feedback of soil moisture on precipitation. In NARR, there is ample moisture that keeps the LCL lower, but a larger sensitivity and deeper PBL growth allows for consistent negative LCL deficits in the afternoon. This would support a positive feedback of soil moisture on precipitation. In contrast, the dry condition of CFSR produces very high LCL values that outweigh those of the deeper PBL growth, thereby supporting positive LCL deficits and limiting the formation of clouds. Because of the dry surface condition, this would also be considered a positive feedback of soil moisture on precipitation. Likewise, in 2007 the wet regime supports a positive soil moisture anomaly in CFSR correlated with more precipitation. Overall, yearly anomalies tend to lead to positive feedbacks in the RAs (e.g., CFSR—wet anomalies; MERRA—dry anomalies) and suggest the climatology of NARR for this region is situated near the "sweet spot" of having enough surface moisture along with enough potential for PBL growth to support the potential for clouds and precipitation.

While these results suggest the presence of feedbacks, it should also be cautioned that the LCL deficit itself might be better suited for individual diurnal cycles (rather than seasonal averages) when the different controlling processes can be quantified. It is likely that the atmospheric (convective, radiation, microphysics) schemes dominate most aspects of precipitation generation, so the signal of the LSM–PBL connection is easily washed out in longer-term (JJA) averaging.

## 5. Applicability to drought investigations

Further LoCo analysis of the full record of RAs (1979–2012) can also yield important results for drought and predictability studies. As part of a collaborative investigation of the 2012 drought over the central United

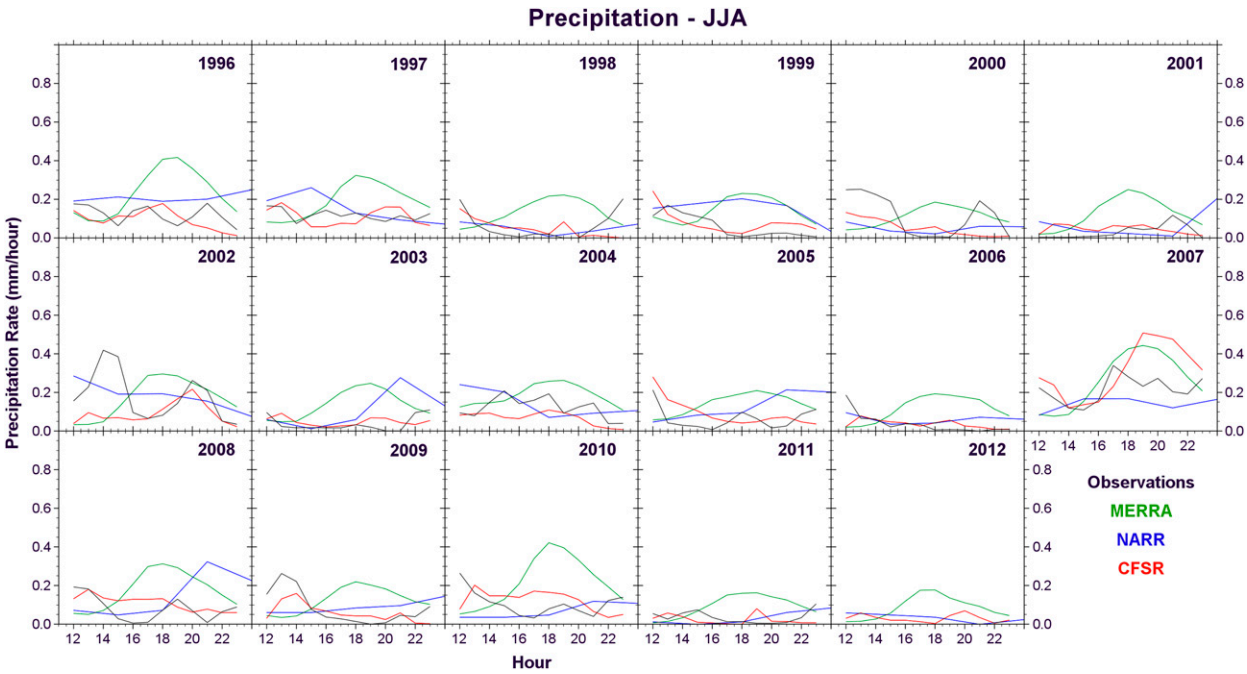


FIG. 6. The daytime cycle of precipitation (mm $h^{-1}$) calculated from JJA monthly mean diurnal cycles of the three RA products— MERRA (green), NARR (blue), and CFSR (red)—and observations (black) at ARM SGP CF over the 1996–2012 period.

States for the NASA Energy and Water Cycle Study (NEWS), the mechanisms that differentiate 2012 from other drought years in the same region were investigated. It was hypothesized that the rapid onset of the 2012 drought was the result of the absence of large-scale influences such as sea surface temperature anomalies, which allowed for L–A interactions to play a larger role in forcing and deepening the drought. By comparing the LoCo analyses from 2011 and 2012, it can be seen that a positive July minus June anomaly in PBLH exists in 2012 that is significantly larger than that of 2011 and one of the largest in the period of record (Fig. 7a). EF (not shown) shows only a slight negative value of July minus June in 2012, indicating that the land surface was preconditioned to be dry in June before the impacts were felt in the PBL in July. The onset of the drought was then supported by significantly increased PBL height in July and entrainment and residual layer (positive) feedbacks that further dried the soil and supported the rapid deepening of the drought [as described in Santanello et al. (2007)]. The large increase in mean PBLH from June to July in 2012 can also be seen in Fig. 7b. Further details of this interdisciplinary study can be found in the paper by Wang et al. (2014) and demonstrate how LoCo diagnostics can be used to apply RAs to understand drought phenomena in a consistent and physically sound framework.

The CDI, another LoCo metric, combines the relationship of soil moisture to the overlying PBL, stability, and humidity and has been used previously to assess the dry–wet climatology of CFSR as well as identifying errors in the forecast system. In this study, the CDI is used to compare all the RAs at the SGP CF and for the whole United States from 1979 to 2012 for the JJA period. The JJA CDI time series for the grid cell over the SGP CF (Fig. 8) confirms that the lower values in MERRA support more wet coupling (higher soil moisture promoting more precipitation) while the larger positive values in CFSR support more dry coupling (almost uniformly over the period). As seen above, the RAs are equally responsive to dry and wet regimes (e.g., 2006 versus 2007), indicating that their potential biases do not preclude them from responding (and in the wet regime case, responding accurately) to specific anomalies. In comparison, there is a large spread in the CDI from the RAs for 2009, which was neither an extreme wet nor dry year.

The spatial and temporal patterns of the CDI (Fig. 9) show that in the SGP region (and over most of the United States) MERRA produces fewer dry coupling events than NARR and CFSR both in location and in magnitude. This is consistent with the results presented in Figs. 1–6 in demonstrating the dry nature of CFSR, which leads to more dry regimes and positive feedbacks (dry soil leading to less precipitation). The results also show 2007 as a year in which the RAs collapse and agree with each other over the SGP (as in Fig. 1), with all RAs indicating a wet regime; however, there is a large

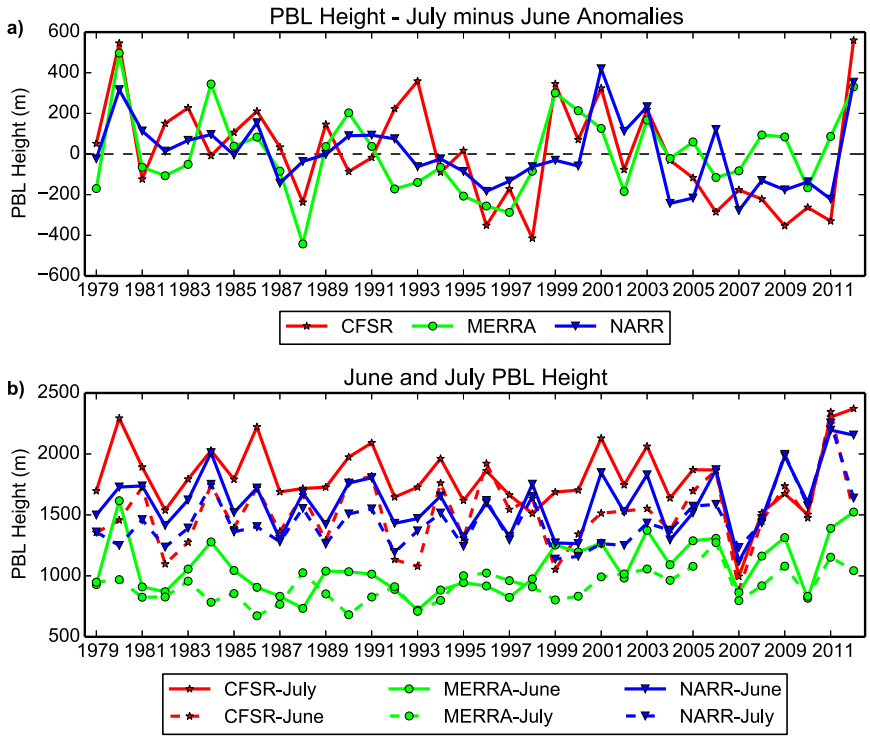


FIG. 7. (a) July minus June anomalies in daytime mean PBLH (m) and (b) June and July mean values of PBLH (m) from each RA product over the 1979–2012 period.

difference in the RAs in terms of the extent of the dry regime over the eastern portion of the United States. This further indicates that the biases and inherent differences in the RAs are more clearly identified during dry regimes. Furthermore, in 2009 it is evident that none of the RAs indicate an extreme dry or wet regime compared to that of 2006 or 2007. However, NARR is much wetter and CFSR drier over most of the United States, with the MERRA somewhere in between. This further indicates that in a year without large-scale anomalies, the individual RAs can have large disagreements. This suggests that not only in dry regimes, but also in relatively neutral regimes, models and assimilation techniques play a larger role in the characterization of the local coupling.

## 6. Discussion and conclusions

This paper has presented a comprehensive analysis and evaluation of the summertime behavior of L–A coupling in three modern RA products over the U.S. SGP. From a combination of mixing diagram, EF–PBLH, and LCL deficit analyses, it was found that CFSR, NARR, and MERRA all differ substantially from one another and from observations during much of the 1996–2012 period of investigation. Most notably, the RAs tended to be drier and warmer than observations in terms of T2 and Q2, where CFSR was the driest and MERRA the wettest of the three. The surface moisture conditions (reflected in EF) thus led to corresponding PBL development and coupled feedbacks that are

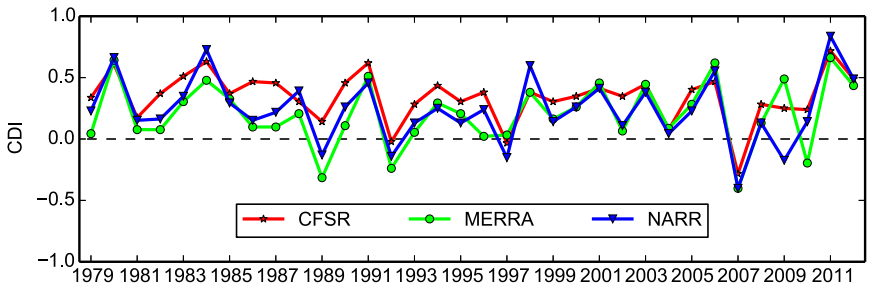


FIG. 8. The JJA CDI for each RA product—MERRA (green), NARR (blue), and CFSR (red)—for the grid cell closest to the ARM SGP CF over the 1979–2012 period.

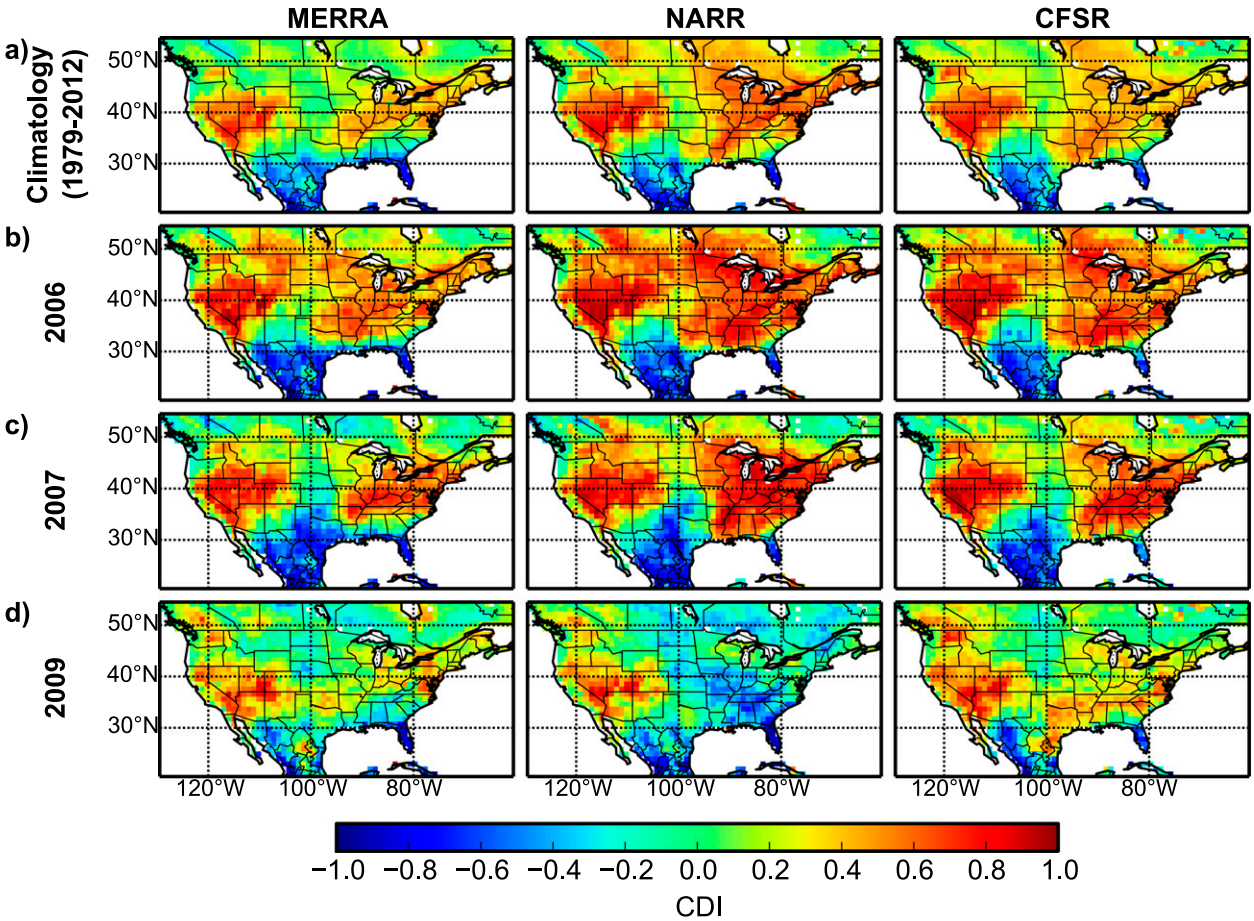


FIG. 9. The JJA CDI over CONUS for (left) MERRA, (center) NARR, and (right) CFSR: (a) the 34-yr climatology for the period 1979–2012, (b) 2006, (c) 2007, and (d) 2009.

unique in each RA product and ultimately help support their climatology in terms of land surface, PBL, and moist processes (cloud and precipitation).

Rather than a detailed validation, the focus of this study was on intercomparison and process understanding. This is the first comprehensive assessment of the land–PBL coupling in community RA products using the LoCo metric suite, and the first to make use of new high-quality and continuous land and PBL datasets available from ARM SGP. Though site specific and focused on a single region, this is a critical step toward understanding the process level of these RAs in the vertical direction (from the soil through the PBL). This approach has also been shown to be valuable in understanding the L–A drivers of specific year anomalies (e.g., drought) as well as understanding why spatial indices depict the SGP region as they do (e.g., CDI). Further, the correlation of changes in RA assimilation data streams to improvements in hydrometeorological variables (e.g., CFSR and Q2) supports that process- and site-level analysis can still inform on RA performance. This also highlights the potential value of atmospheric sounders (such as AIRS and IASI) for L–A coupling studies, which has largely been ignored to date.

The results from traditional evaluation of RA products and individual land–atmosphere variables are generally consistent with those found here. Overall, the higher spatial resolution of NARR did not produce results that were superior to the slightly coarser CFSR and MERRA products. For example, Urankar et al. (2012) found that PBLH was considerably higher in CFSR than in MERRA. The results of Bosilovich (2013) found a smaller standard deviation in MERRA precipitation and T2 than in CFSR or observations over the SGP region, particularly during extremes. This is in agreement with the more humid and less variable MERRA behavior seen here (e.g., the lower variability in EF versus PBLH shown in Fig. 4). The proper representation of the diurnal range of T2 in the summertime was also suggested by Decker et al. (2012) for CFSR and MERRA. Becker et al. (2009) investigated NARR and showed a high bias in incoming shortwave radiation and low cloud fraction that have

been confirmed by Kennedy et al. (2011) and others using the eta model and also found in CFSR and MERRA by Decker et al. (2012). This is consistent with the dry and warm bias in the RAs seen here. The Noah LSM (employed in NARR and CFSR) has also been shown to have a significant dry bias, particularly over the SGP (Santanello et al. 2013a), and highlights the importance of the land surface component of RAs in terms of impacting the broader L–A coupling.

An ongoing issue in the land modeling community is that of the differing climatologies of soil moisture across LSMs, which make intercomparison and assimilation difficult. It may be the case that PBL schemes also have different climatologies across regimes (e.g., stable versus convective; wet versus dry) and model resolutions. This is evident in the unique relationship of EF–PBLH across the spectrum in each RA product (as a function of its LSM and PBL scheme coupling). Analogous to the importance of the soil moisture–evaporation relationship in LSMs (e.g., Koster and Mahanama 2012), the EF–PBLH relationship may demonstrate the ability of soil moisture variations to impact the atmosphere and may serve as an identifiable metric of L–A coupling.

The LoCo analysis presented here has enabled these individual targeted (land or atmosphere) study results to be linked together in the context of the coupled land–PBL system and feedbacks. It was also shown here that even in the absence of observations of land or PBL variables, much could still be learned from an RA intercomparison in the LoCo context. Because PBL observations are often scarce on diurnal time scales, it should be noted that readily available 2-m data (T2 and Q2) on their own can add considerable value to the analysis (e.g., Figs. 1 and 2) and integrate much of the PBL feedback in terms of the RA climatology and diurnal cycle assessment. The coupling tendency of each can be quantified and compared to the other products, along with (most importantly) the implications of this coupling on the broader RA components and overall representation of atmospheric reality. Future work should also consider application of coupling metrics for RA products that assimilate screen-level variables [such as ERA-Interim and the Japanese 55-year Reanalysis Project (JRA-55)], in particular to address the implications of these approaches on the representation of land–PBL processes.

Finally, the focus of this study was on the land–PBL processes on seasonal scales and the interplay of surface and PBL variables (rather than clouds and precipitation). This is the necessary precursor to fully understanding the soil moisture–precipitation relationship, feedbacks, and causality. Future work will look in detail at the diurnal processes and interactions with cloud cover, height, convection, and precipitation types. This also requires a strategy to reconcile point observations with large gridcell averages of precipitation and clouds that are often much more heterogeneous than quantities like temperature, humidity, and the fetch of the PBL. Such analyses will go a long way to bridging the gap between local–process-oriented metrics of L–A coupling and larger-scale assessments of the soil moisture–precipitation relationship (Findell et al. 2011).

*Acknowledgments.* This work was supported by NASA's Energy and Water Cycle Study (NEWS) and their Modeling, Analysis, and Prediction (MAP) Program. In addition, Joshua Roundy was supported by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by Oak Ridge Associated Universities through a contract with NASA. Many thanks go to Bob Dattore at UCAR for producing the monthly mean diurnal cycle data and variables from CFSR specifically for this study. We also thank David Troyan at Brookhaven National Laboratory for the reprocessing of ARM SGP MERGESONDE data to better incorporate radiosonde data in the PBL. Finally, the guidance of Michael Bosilovich was invaluable in terms of the MERRA data processing and interpretation.

## REFERENCES

Becker, E. J., E. H. Berbery, and R. W. Higgins, 2009: Understanding the characteristics of daily precipitation over the United States using the North American Regional Reanalysis. *J. Climate,* **22,** 6268–6286, doi:10.1175/2009JCLI2838.1.

Betts, A. K., 1992: FIFE atmospheric boundary layer budget methods. *J. Geophys. Res.,* **97,** 18 523–18 531, doi:10.1029/91JD03172.

——, and A. G. Barr, 1996: First International Satellite Land Surface Climatology Field Experiment 1987 sonde budget revisited. *J. Geophys. Res.,* **101,** 23 285–23 288, doi:10.1029/96JD02247.

Bosilovich, M. G., 2013: Regional climate and variability of NASA MERRA and recent reanalyses: U.S. summertime precipitation and temperature. *J. Appl. Meteor. Climatol.,* **52,** 1939–1951, doi:10.1175/JAMC-D-12-0291.1.

Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. Zeng, and M. G. Bosilovich, 2012: Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *J. Climate,* **25,** 1916–1944, doi:10.1175/JCLI-D-11-00004.1.

Dirmeyer, P. A., 2013: Characteristics of the water cycle and land–atmosphere interactions from a comprehensive reforecast and reanalysis data set: CFSv2. *Climate Dyn.,* **41,** 1083–1097, doi:10.1007/s00382-013-1866-x.

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.

Entekhabi, D., and Coauthors, 1999: An agenda for land surface hydrology research and a call for the second International Hydrological Decade. *Bull. Amer. Meteor. Soc.,* **80,** 2043–2058, doi:10.1175/1520-0477(1999)080<2043:AAFLSH>2.0.CO;2.

Findell, K. L., and E. A. B. Eltahir, 2003a: Atmospheric controls on soil moisture–boundary layer interactions. Part I: Framework development. *J. Hydrometeor.,* **4,** 552–569, doi:10.1175/1525-7541(2003)004<0552:ACOSML>2.0.CO;2.

——, and ——, 2003b: Atmospheric controls on soil moisture–boundary layer interactions. Part II: Feedbacks within the Continental United States. *J. Hydrometeor.,* **4,** 570–583, doi:10.1175/1525-7541(2003)004<0570:ACOSML>2.0.CO;2.

——, P. Gentine, B. R. Lintner, and C. Kerr, 2011: Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation. *Nat. Geosci.,* **4,** 434–439, doi:10.1038/ngeo1174.

Guo, Z., and Coauthors, 2006: GLACE: The Global Land–Atmosphere Coupling Experiment. Part II: Analysis. *J. Hydrometeor.,* **7,** 611–625, doi:10.1175/JHM511.1.

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, doi:10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

——, Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.,* **134,** 2318–2341, doi:10.1175/MWR3199.1.

Jakob, C., 2010: Accelerating progress in global atmospheric model development through improved parameterizations: Challenges, opportunities, and strategies. *Bull. Amer. Meteor. Soc.,* **91,** 869–875, doi:10.1175/2009BAMS2898.1.

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, doi:10.1175/2011JCLI3978.1.

Koster, R. D., and S. P. P. Mahanama, 2012: Land surface controls on hydroclimatic means and variability. *J. Hydrometeor.,* **13,** 1604–1620, doi:10.1175/JHM-D-12-050.1.

——, M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.,* **105,** 24 809–24 822, doi:10.1029/2000JD900327.

——, and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science,* **305,** 1138–1140, doi:10.1126/science.1100217.

Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.,* **128,** 3187–3199, doi:10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2.

Louis, J., M. Tiedtke, and J. Geleyn, 1982: A short history of the PBL parameterization at ECMWF. *Proc. ECMWF Workshop on Planetary Boundary Layer Parameterization,* Reading, United Kingdom, European Centre for Medium-Range Weather Forecasts, 59–80.

McCoy, R., and S. Xie, 2012: ARM Best Estimate Atmospheric Measurements (ARMBE-ATM) Southern Great Plains Central Facility (C1) dataset. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive, accessed 10 February 2014, doi:10.5439/1039931.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.,* **87,** 343–360, doi:10.1175/BAMS-87-3-343.

Mueller, B., and Coauthors, 2013: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-dataset synthesis. *Hydrol. Earth Syst. Sci. Discuss.,* **10,** 769–805, doi:10.5194/hessd-10-769-2013.

NCEP Environmental Modeling Center, 2010: NCEP Climate Forecast System Reanalysis (CFSR) monthly products, January 1979 to December 2010. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, accessed July 2014. [Available online at http://rda.ucar.edu/datasets/ds093.2.]

Phillips, T. J., and S. A. Klein, 2014: Land–atmosphere coupling manifested in warm-season observations on the U.S. southern great plains. *J. Geophys. Res. Atmos.,* **119,** 509–528, doi:10.1002/2013JD020492.

Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. Touré, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate,* **24,** 6322–6338, doi:10.1175/JCLI-D-10-05033.1.

Rienecker, M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* **24,** 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Rodell, M., and Coauthors, 2015: The observed state of the water cycle in the early 21st century. *J. Climate,* doi:10.1175/JCLI-D-14-00555.1, in press.

Roundy, J. K., C. R. Ferguson, and E. F. Wood, 2013: Temporal variability of land–atmosphere coupling and its implications for drought over the southeast United States. *J. Hydrometeor.,* **14,** 622–635, doi:10.1175/JHM-D-12-090.1.

——, ——, and ——, 2014: Impact of land-atmospheric coupling in CFSv2 on drought prediction. *Climate Dyn.,* **43,** 421–434, doi:10.1007/s00382-013-1982-7.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.,* **91,** 1015–1057, doi:10.1175/2010BAMS3001.1.

——, and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate,* **27,** 2185–2208, doi:10.1175/JCLI-D-12-00823.1.

Santanello, J. A., 2011: Results from Local Land–Atmosphere Coupling (LoCo) Project. *GEWEX News,* No. 21, International GEWEX Project Office, Silver Spring, MD, 7–9.

——, M. A. Friedl, and W. P. Kustas, 2005: An empirical investigation of convective planetary boundary layer evolution and its relationship with the land surface. *J. Appl. Meteor.,* **44,** 917–932, doi:10.1175/JAM2240.1.

——, ——, and M. B. Ek, 2007: Convective planetary boundary layer interactions with the land surface at diurnal time scales: Diagnostics and feedbacks. *J. Hydrometeor.,* **8,** 1082–1097, doi:10.1175/JHM614.1.

——, 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, doi:10.1175/2009JHM1066.1.

——, ——, and ——, 2011: Diagnosing the sensitivity of local land–atmosphere coupling via the soil moisture–boundary layer interaction. *J. Hydrometeor.,* **12,** 766–786, doi:10.1175/JHM-D-10-05014.1.

——, S. V. Kumar, C. D. Peters-Lidard, K. Harrison, and S. Zhou, 2013a: Impact of land model calibration on coupled land–atmosphere prediction. *J. Hydrometeor.,* **14,** 1373–1400, doi:10.1175/JHM-D-12-0127.1.

——, C. D. Peters-Lidard, A. Kennedy, and S. Kumar, 2013b: Diagnosing the nature of land–atmosphere coupling: A case study of dry/wet extremes in the U.S. southern Great Plains. *J. Hydrometeor.,* **14,** 1373–1400, doi:10.1175/JHM-D-12-0127.1.

Shah, R., and V. Mishra, 2014: Evaluation of the reanalysis products for the monsoon season droughts in India. *J. Hydrometeor.,* **15,** 1575–1591, doi:10.1175/JHM-D-13-0103.1.

Thorne, P. W., and R. S. Vose, 2010: Reanalyses suitable for characterizing long-term trends. *Bull. Amer. Meteor. Soc.,* **91,** 353–361, doi:10.1175/2009BAMS2858.1.

Troyan, D., 2012: Merged sounding value-added product. U.S. Department of Energy Tech. Rep. DOE/SC-ARM/TR-087, 13 pp. [Available online at https://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-087.pdf?id=51.]

Urankar, G., T. V. Prabha, G. Pandithurai, P. Pallavi, D. Achuthavarier, and B. N. Goswami, 2012: Aerosol and cloud feedbacks on surface energy balance over selected regions of the Indian subcontinent. *J. Geophys. Res.,* **117,** D04210, doi:10.1029/2011JD016363.

van den Hurk, B., 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, doi:10.1175/BAMS-D-11-00007.1.

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.

Wang, S., and Coauthors, 2014: Could the 2012 drought have been anticipated—A review of NASA Working Group research. *J. Earth Sci. Eng.,* **4,** 428–437.

Wilson, K., and Coauthors, 2002: Energy balance closure at FLUXNET sites. *Agric. For. Meteor.,* **113,** 223–243, doi:10.1016/S0168-1923(02)00109-0.

Xie, S., and Coauthors, 2010: ARM climate modeling best estimate data: A new data product for climate studies. *Bull. Amer. Meteor. Soc.,* **91,** 13–20, doi:10.1175/2009BAMS2891.1.

——, and Coauthors, 2012: ARM Best Estimate (ARMBE-Land) Southern Great Plains Central Facility (C1) dataset. Atmospheric Radiation Measurement (ARM) Climate Research Facility Data Archive, accessed 10 February 2014. [Available online at http://www.arm.gov/data/eval/78.]

Yi, Y., J. S. Kimball, L. A. Jones, R. H. Reichle, and K. C. McDonald, 2011: Evaluation of MERRA land surface estimates in preparation for the soil moisture active passive mission. *J. Climate,* **24,** 3797–3816, doi:10.1175/2011JCLI4034.1.

Zhang, Y., and S. A. Klein, 2010: Mechanisms affecting the transition from shallow to deep convection over land: Inferences from observations of the diurnal cycle collected at the ARM Southern Greeat Plains site. *J. Atmos. Sci.,* **67,** 2943–2959, doi:10.1175/2010JAS3366.1.