Understanding the Impacts of Soil Moisture Initial Conditions on NWP in the Context of Land-Atmosphere Coupling

Joseph A. Santanello, Jr.[1], Patricia Lawston[2,1], Sujay Kumar[1], and Eli Dennis[2,3]

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

[2] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD

[3] Cooperative Institute for Climate and Satellites - Maryland, College Park, MD

Corresponding Author: Dr. Joseph A. Santanello, Jr.

NASA-GSFC, Code 617, Bldg. 33, Room G220, Greenbelt, MD 20771 USA

Joseph.A.Santanello@nasa.gov

Revised for *Journal of Hydrometeorology*

ABSTRACT

The role of soil moisture in NWP has gained more attention in recent years, as studies have demonstrated impacts of land surface states on ambient weather from diurnal to seasonal scales. However, soil moisture initialization approaches in coupled models remain quite diverse in terms of their complexity and observational roots, while assessment using bulk forecast statistics can be simplistic and misleading. In this study, a suite of soil moisture initialization approaches is used to generate short-term coupled forecasts over the U.S. Southern Great Plains using NASA's Land Information System (LIS) and NASA Unified WRF (NU-WRF) modeling systems. This includes a wide range of currently used initialization approaches, including soil moisture derived from 'off-the-shelf' products such as atmospheric models and land data assimilation systems, high-resolution land surface model spinups, and satellite-based soil moisture products from SMAP. Results indicate that the spread across initialization approaches can be quite large in terms of soil moisture conditions and spatial resolution, and that SMAP performs well in terms of heterogeneity and temporal dynamics when compared against high resolution land surface model and in-situ soil moisture estimates. Case studies are analyzed using the local land-atmosphere coupling (LoCo) framework that relies on integrated assessment of soil moisture, surface flux, boundary layer, and ambient weather, with results highlighting the critical role of inherent model background biases. In addition, simultaneous assessment of land vs. atmospheric initial conditions in an integrated, process-level fashion can help address the question of whether improvements in traditional NWP verification statistics are achieved for the right reasons.

1. Introduction

The role of the land surface in numerical weather prediction (NWP) has been traditionally overlooked by the atmospheric modeling community (Santanello et al. 2018), who often employ primitive initialization approaches for soil moisture and temperature based on coarse atmospheric model products. These surface conditions have been treated simply as lower boundary conditions, with early LSM development driven by the atmospheric communities and little emphasis on the accuracy and observability of land surface states and processes (Dirmeyer and Halder 2016). However, recent studies have demonstrated the critical role of the land surface, and in particular soil moisture, in terms of impacts on precipitation (Welty and Zeng, 2018; Ford et al. 2015; Taylor et al. 2012; Koster et al. 2004; Findell and Eltahir 2003), temperature and humidity (Kala et al. 2015; Seneviratne et al. 2013; Mueller and Seneviratne 2012), and L-A coupling as a whole including the planetary boundary layer (PBL) (Johnson and Hitchens 2018; Dirmeyer and Halder 2016; Santanello et al. 2007, 2005). In particular, the initial condition (IC) of soil moisture has been shown to influence predictability on near-term (Dirmeyer and Halder 2016; Santanello et al. 2016, 2013a) and seasonal (Rajesh et al. 2018, Xiang et al. 2018; Hirsch et al. 2014; Koster et al. 2010) scales. Thus, quantification of the sensitivity of coupled models to the soil moisture initialization approach is an often overlooked, but potentially high-impact, exercise that should be performed by model evaluation and development communities.

From a process-level perspective, the connection of soil moisture to ambient weather and precipitation can be considered using the GEWEX local land-atmosphere coupling (LoCo; Santanello et al. 2018) paradigm and 'process chain', as follows:

$$\Delta SM \underset{\textbf{(a)}}{\longrightarrow} \Delta EF \underset{\textbf{(b)}}{\longrightarrow} \Delta PBL \underset{\textbf{(c)}}{\longrightarrow} \Delta Ent \underset{\textbf{(d)}}{\longrightarrow} \Delta T_{2m}, Q_{2m} \Rightarrow \Delta P, Cloud \qquad (1)$$

where the links (a-d) represent the sensitivities of: (a) evaporative fraction (EF; i.e. surface fluxes) to soil moisture (SM), (b) PBL evolution to surface fluxes, (c) entrainment (Ent) fluxes to PBL evolution, (d) the collective feedback of the free atmosphere on ambient weather, and the cumulative support of these links on cloud and precipitation formation. By parsing out the stepwise impact of soil moisture on surface fluxes and, likewise, the vertical coupling impacts of surface fluxes on PBL development and entrainment feedbacks, an understanding of the interaction of the coupled model sensitivities to soil moisture can be ascertained.

To this end, the LoCo community has been developing metrics to quantify the links in the chain of Eq. (1) that can also be used to better understand traditional 'bulk' statistics of ambient weather (e.g. 2-meter temperature (T2m) and humidity (Q2m), RMSE and bias) commonly used by operational centers as benchmarks, in the context of the influence of soil moisture on model accuracy and development. Such approaches have been previously employed to assess the coupling behavior in modern global climate reanalysis products (Santanello et al. 2015), to quantify the impact of LSM calibration and assimilation on short-term coupled forecasts (Santanello et al. 2013a, 2015b), and to intercompare the coupled behavior of different parameterization combinations in regional NWP (Santanello et al. 2013b).

These studies assumed that the land surface IC was based on an offline, high-resolution, high-quality, long-term spinup of soil states from a land data assimilation system (LDAS) such as NASA's Land Information System (LIS; Kumar et al. 2008). However, despite their advantages, such spinup approaches are still not the norm outside of the land modeling community. In addition, with recent advances in satellite-based soil moisture retrievals such as those from SMOS and SMAP, and long-term in-situ networks such as those comprising the International Soil Moisture Network (Dorigo et al. 2011), there are now additional observationally-driven initialization approaches that need to be considered (Dirmeyer et al. 2018, 2016). Overall, there is a wide array of soil moisture IC approaches that are being used across the NWP and climate modeling communities (both research and operational), ranging in complexity, resolution, quality, and observability.

In this paper, we assess the impacts of soil moisture IC approaches in an NWP context using the NASA Unified WRF model (NU-WRF; Peters-Lidard et al. 2015), focused on an integrative, process-level assessment of L-A coupling and ambient weather implications. Specifically, we intercompare a suite of initializations of high resolution (1 km) short-term weather forecasts using 'off-the-shelf' soil moisture products from large scale atmospheric and land surface reanalysis products, high-resolution LIS spinups, and SMAP satellite retrievals which range in horizontal resolution from 1-33 km. Section 2 reviews the current suite of soil moisture IC approaches being used by the community. Section 3 describes the model and observation products used for initialization, and the LIS and NU-WRF modeling systems, along with case study and site descriptions for the coupled experiments. Section 4 presents an offline intercomparison of SMAP soil moisture with that of in-situ networks and LIS-based simulations over the domain of interest. Section 5 then presents the full suite of coupled NU-WRF experiments with varying ICs, and corresponding LoCo analysis and ambient weather evaluations. Discussion and conclusions then follow in Section 6.

2. Review of Soil Moisture Initialization Approaches

Figure 1 shows the suite of soil moisture initialization approaches for NWP and regional modeling commonly used by the operational and research communities. Despite their wide range in complexity, studies using each of these approaches have been published recently, and the typical usage by the community tends to favor the lower complexity approaches.

Going from least to most complex (in Fig. 1), while a uniform (homogeneous) soil moisture IC may have been commonly employed decades ago during the early years of atmospheric and L-A coupled model development, this approach lacks an accurate representation of surface heterogeneity and likely leads to poor NWP results in terms of surface flux partitioning and impact on PBL and atmospheric processes. There remain a few regional modeling applications (including operational) that employ homogeneous soil moisture ICs, such as those using RAMS (Gomez et al. 2016, 2015), where only recently has the sensitivity to varying ICs been evaluated in a systematic manner (Gomez et al. 2018). These studies affirm that there are significant impacts of soil moisture IC values on NWP, and point to the potential for satellite data to inform on improved, heterogeneous IC approaches. It should also be noted that homogeneous and idealized surface conditions are still the norm for the LES and cloud resolving model communities (focused on ~100m scales).

Next, from an observational perspective, in-situ soil moisture measurements are direct measurements at fixed depths while satellite measurements are indirect and only sensitive to a thin layer near the surface typically less than a few centimeters. The in-situ measurements are typically considered truth in the development and parameterization of the soil models themselves. Therefore, soil moisture ICs based on a dense in-situ network that covers the model domain would be an ideal approach. Such networks are rarely dense enough to meet NWP application requirements. One example is the DOE's ARM Southern Great Plains (ARM-SGP) observatory covering OK and KS, where ~20 sites measuring soil moisture are available. Even in this dense network, 20 observations across a domain with 250,000 grid cells is hardly representative of land surface and soil type heterogeneity, and interpolation procedures and assumptions do not readily apply to obtain distributed estimates. Recent examples of utilizing a dense in-situ network in the context of NWP ICs do exist (e.g. Massey et al. 2016), but remain limited due to the heterogeneous nature of soil properties.

The next grouping of ICs in terms of complexity is what are deemed 'off-the-shelf' products, where soil moisture is extracted from existing LDAS or atmospheric modeling systems. Atmospheric-based products are simply the soil moisture fields derived from the land surface component of the commonly used initial/boundary condition datasets for NWP. These atmospheric models include GFS, ECWMF, NARR, NAM, and others, and the advantage of using these land ICs is that they are inherently consistent (spatially and temporally, as well as climatologically) with the atmospheric ICs. The disadvantage is that these atmospheric models are quite coarse spatially (~25–40 km) relative to the grid size of the NWP or WRF application (e.g. 1–3 km in this case). Thus, initializing a domain with 1-km horizontal grid spacing with data that has 30-km resolution will miss some crucial heterogeneity and likely does not capture the true nature of the local L-A coupling. Another caveat to this approach is when the LSM used in the NWP application differs from that in the atmospheric product, as soil moisture climatologies differ across LSMs and thus are not easily transferable. In addition, any inherent biases (e.g. precipitation) in the atmospheric model will be reflected in the land states (e.g. soil moisture) as well.

From an LDAS perspective, the GLDAS (Rodell et al. 2004) and the NLDAS Version 2 (NLDAS-2; Xia et al. 2012) are examples of uncoupled environments, run routinely at 25 km and 12.5 km, respectively. Both are offline LSMs driven by high-quality, ground-observation based atmospheric forcing and parameters, though there is still a significant gap in resolution compared with the target model (1 km vs. 12.5–25 km). The land ICs from LDAS systems are likely to be more accurate and heterogeneous versus atmospheric model-based ICs described above, however there may be inconsistency issues with the LSM used in the LDAS and that used in the NWP application (e.g. soil or vegetation parameters). These ‘canned’ spinups have been largely underutilized by the NWP community and represent a spinup shortcut that does not require a multi-year spinup or access to forcing data. A recent study from Dillon et al. (2016) compared the impacts of using GLDAS versus GFS soil moisture ICs on short-term WRF forecasts, and neither approach consistently outperformed the other over South America. Gomez et al. (2018) updated RAMS with spatially distributed soil ICs from GLDAS, and found significant improvement in ambient weather forecasts. Jacobs et al. (2017) used an Australian LDAS (Australian Water Availability Project (AWAP); http://www.csiro.au/awap/) to generate a 5–km gridded soil moisture IC product for WRF heatwave simulations, and found that forecasts improved significantly over those using ERA-I-based soil moisture ICs, as AWAP corrected for the cool, wet bias inherent in ERA-I. Lin and Cheng (2016) also compared the impacts of GLDAS versus GFS-based ICs on WRF forecasts over Taiwan, and showed improvements over certain regions where GFS was biased and where soil exerts more control over surface fluxes. Overall, these studies demonstrate general improvement from LDAS-based ICs over those of coarser atmospheric-based products.

In terms of model-observation fusion, LSM spinup approaches, as described earlier, can generate high resolution, accurate ICs at the resolution of the target model using observed forcing and parameter data. LSM spinups are necessary to generate soil moisture and temperature profiles that have equilibrated over time, and thus are often on the order of a few years to decades in length leading up to the time of coupled model initialization (Rodell et al. 2005). Spinups are facilitated by systems such as LIS and the High Resolution LDAS (HRLDAS; Chen et al. 2017), and require multi-year offline simulations driven by high quality forcing data. As a result, the advantages are in resolution, representativeness and quality (including consistency of LSM states with observed meteorology), but can be limited by the availability of accurate forcing data and computational demand. As compared to LDAS systems, LSM spinups can further resolve spatial heterogeneity down to 1km or less, and in coupled systems such as LIS/NU-WRF, will ensure identical LSM settings in both the offline spinup and coupled simulations. To date, LSM spinups have been employed in numerous regional (WRF) modeling studies, but are typically only employed by those in the land (LIS and HRLDAS) communities (Santanello et al. 2013ab, Case et al. 2011, 2008, Kumar et al. 2008, Rajesh et al. 2017, Hirsch et al. 2014), while atmospheric modelers have been much less inclined to invest in this approach.

Recent advances in satellite retrieval of land surface states now allow for soil moisture ICs to be fully or partly derived from satellite data. Satellites such as SMAP can provide gridded products comparable in spatial resolution to that of the off-the-shelf products described above (e.g. SMAP 9- and 36-km products). However, using satellite products directly as ICs is not currently advisable, due to differing climatologies and biases inherent in satellite versus LSM-based soil moisture. As satellite soil moisture becomes more accurate (as is the case with SMAP), and LSM soil moisture becomes more 'observable', opportunities to directly employ satellite products as ICs will become apparent. The traditional method to incorporate satellite-based soil moisture into ICs has been through data assimilation after utilizing bias correction techniques such as CDF matching to account for the satellite versus LSM biases (Reichle and Koster 2004). Assimilation incorporates some of the satellite signal that the LSM may miss, but impacts are typically muted due to the low random error present in satellite and LSM products, and high accuracy of atmospheric (i.e. precipitation) forcing of the LSM. Recent efforts at NCEP and Environment Canada have employed SMAP data assimilation during LSM spinup to improve soil moisture ICs for NWP, with results showing inconsistent improvements across large continental domains. These approaches and results will be discussed in more detail in Sections 5 and 6 in the context of the results presented in this study.

An additional IC approach that does not fall neatly into the categories in Fig. 1 is that of ‘self-spinup’, as described by Angevine et al. (2014) and performed by Dy and Fung (2016). In self-spinup, the NWP or regional model of interest can be initialized with coarse atmospheric-based soil moisture ICs, then cycled over months or years at a time (resources permitting), which improves the spatial resolution as it responds to precipitation and other forcing at the finer target model resolution. A disadvantage is that the accuracy of resulting soil moisture is entirely dependent on the free running model performance and precipitation accuracy, which likely is less than that of the observationally-constrained atmospheric models. As with other IC approaches, there are tradeoffs, and in this case that is gaining spatial heterogeneity but perhaps losing the spatial accuracy of soil moisture anomalies.

Also not considered in the suite of ICs in Fig. 1 are basic sensitivity studies that vary the soil moisture IC, often uniformly, in ‘brute force’ fashion in order to assess impacts on NWP (e.g. Kalverla et al. 2016, Ament et al. 2006, Daniels et al. 2016, Collow et al. 2014). There have been a host of studies in this regard, each emphasizing a different aspect of coupled impacts, typically focusing on singular impacts on temperature or precipitation. Overall, these sensitivity studies highlight the importance of soil moisture ICs, and underscore the need for high-quality and high-resolution ICs (preferably from satellite or LSM spinup).

3. Experimental Design

The suite of soil moisture IC approaches in Fig. 1 will be intercompared in this study using

SMAP for satellite observations, LIS for LSM spinup, NU-WRF as the coupled forecast model, along with the standard off-the-shelf products from NARR, GFS, and NLDAS-2 and in-situ evaluation data from the ARM-SGP network.

a. SMAP Soil Moisture

NASA's SMAP mission was launched in January 2015, and has been providing passive microwave retrievals of soil moisture from April 2015 to present. SMAP soil moisture has performed well to date, reaching the mission target of +/- 0.04 m3 m-3 accuracy over most regions (Chan et al. 2018), and with higher temporal consistency (i.e. less noise) and overall information content than other passive microwave-based soil moisture products from missions such as SMOS, AMSR-E, and ASCAT (Kumar et al. 2017). SMAP soil moisture also exhibits wetting and drydown responses that are consistent with those modeled by LSMs (Shellito et al. 2017), and has been useful in detecting the timing and spatial extent of irrigation (Lawston et al. 2017). In addition, SMAP has been used successfully in data assimilation studies by operational and research centers (e.g. Fang et al. 2018, Carerra et al. 2018). In this study, we use the SMAP L3 enhanced soil moisture retrieval, based on the 33-km retrieval algorithm but posted at 9-km spatial resolution after utilizing the oversampling of the SMAP footprint. Section 4 presents a comparison of SMAP products against in-situ and modeled soil moisture in order to assess any relative biases or observability issues amongst these products before they are infused into offline or coupled modeling applications.

b. LIS and NU-WRF Modeling Systems

As mentioned in Section 1, we now have the ability to generate high resolution, high-quality, long-term integrations in offline land data assimilation systems such as LIS, which incorporate high resolution, observed forcing and satellite parameter and state datasets. LIS, with its choice of LSMs, parameter and forcing data, and assimilation and calibration modules, can be run in offline (uncoupled) mode for multiyear spinups that can then be used to initialize coupled models such as NU-WRF.

NU-WRF is NASA-GSFC's version of the community WRF-ARW model, and is essentially a superset of the ARW model that includes unique NASA assets and physics capabilities including radiation, microphysics, chemistry, and land surface (via LIS). In addition to providing the soil ICs via LSM spinup, LIS is also coupled to NU-WRF and can be used as the LSM during fully coupled simulations. This is advantageous in terms of utilizing the identical model, grid, and configuration in the offline spinup as during the coupled experiment. The LIS/NU-WRF coupling (Kumar et al. 2008) has been used extensively in research focused on quantifying forecast impacts of different land cover (Case et al. 2011, 2008), irrigation (Lawston et al. 2015), soil condition (Zaitchik et al. 2013), atmospheric forcing (Santanello et al. 2016), LSM calibration (Santanello et al. 2013a), and land data assimilation formulation (Santanello et al. 2016, Feng et al. 2018, Carrera et al. 2018), and serves as an ideal testbed to examine the sensitivity to soil moisture initialization approaches. For this study, LIS version 7 (LISv7.0; lis.gsfc.nasa.gov) is employed with the Noah LSM, version 3.3 LSM (Ek et al. 2003) and coupled to NU-WRF, version 8 patch 4 (https://nuwrf.gsfc.nasa.gov/).

c. *Off-the-Shelf Products*

The Global Forecast System (GFS; Environmental Modeling Center 2003) is an operational, global spectral model driven by the Global Data Assimilation System (GDAS), which incorporates satellite, surface, aircraft, and other observations from across the globe into a gridded, model space. The land component of GFS was upgraded to the Noah LSM version 2.7.1 in the mid-2000's, reducing prominent biases in snow pack, evaporation, and precipitation. The GFS analyses are generated at 6-hourly intervals and gridded at 0.25° spatial resolution (available via rda.ucar.edu/datasets/ds084.1). The North American Regional Reanalysis (NARR; Mesinger et al. 2006) uses the Eta model, the Noah LSM, and advances in data assimilation to create long term, consistent weather data at 32-km spatial resolution and 3-hourly intervals (available via rda.ucar.edu/datasets/ds608.0). NARR was the first reanalysis to include precipitation assimilation and shows considerable improvement over the previous NCEP reanalysis system (Kennedy et al. 2011). Finally, NLDAS-2 provides near-real time, 1/8° (~12-km) resolution, quality-controlled datasets of atmospheric forcing needed to run LSMs, as well as LSM output from four different models driven by these data. We use both the NLDAS-2 meteorological forcing (to drive LIS offline simulations) and the NLDAS-2 model output from the Noah LSM (version 2.8) for land initial conditions, discussed further in the experimental design.

d. Experimental Design

An extensive survey of potential coupled case study dates was performed over a regional modeling domain over the U.S. SGP region (NE, KS, OK, and TX). The initial time of the case studies was limited to those that had full coverage from SMAP at 6am local time, and due to the geographical location and SMAP orbital pattern, this occurred every ~6 days. A relatively clear-sky morning, with weak synoptic flow and potential locally-induced convection in the afternoon was desirable. This would allow local land effects due to surface and soil moisture heterogeneity to be maximized. In addition, contrasts in soil moisture across the domain were deemed as advantageous to the goals of this study in terms of highlighting differences in ICs captured by approaches in Fig. 1. Lastly, the Enhanced Soundings for Local Coupling Studies (ESLCS; Ferguson et al. 2016) campaign took place in Summer 2015, which was comprised of 12 IOP days with hourly radiosondes launches during the daytime that were deemed useful for model validation. Taking all of these factors into account, 11 July 2015 was chosen as the primary coupled case study for this study, with 10 June 2015 as a secondary case to support any conclusions made from the July case.

Thus, LIS and NU-WRF are run on a single 750-km x 1100-km domain over the SGP at 1-km spatial resolution (Fig. 2) using a 3-s time step, GSFC microphysics, GSFC long- and shortwave radiation, Mellor–Yamada–Nakanishi–Niino (MYNN) PBL scheme, and Monin–Obukhov surface layer scheme. NARR and GFS data were used for atmospheric initialization for different simulations that will be discussed below, with 3-hourly lateral boundary condition nudging, and 61 vertical levels. Simulations were initialized at 12 UTC on the morning of 11 July, and run for 24 hours.

Each of the soil moisture IC approaches were implemented as in Table 1 for a total of eight coupled simulations. The ‘off-the-shelf ICs from NLDAS-2, NARR, and GFS were performed by using the soil moisture (and temperature) profile ICs from those atmospheric model products. Three LIS spinups were performed beginning on 1 January 2010 through 31 December 2016, the domain of which along with input land cover and soil type data as well as ARM-SGP site locations are shown in Fig. 2. The LIS-Control run used NLDAS-2 atmospheric forcing along with default climatological greenness vegetation fraction (GVF) data from NCEP. Two permutations of LIS spinup were also performed, one using GDAS atmospheric forcing data (LIS-GDAS) instead of NLDAS-2, and the other with real-time GVF data from the VIIRS satellite (LIS-VIIRS) instead of climatological GVF. The goal of the LIS suite of runs was to create a mini-ensemble of the range of IC spread that would be generated from different LSM spinup approaches and forcing/parameter data quality. Atmospheric forcing was varied to provide spread based on uncertainty in precipitation forcing, while GVF was varied to account for uncertainty in soil moisture due to vegetation amount and evaporation. For the three spinup runs, NU-WRF was then run coupled to LIS throughout the 24h simulation, thus ensuring consistency from spinup through coupled forecast in terms of the LSM configuration. The off-the-shelf ICs were taken as described above from the NLDAS-2, GFS, and NARR products which provided the 4 layers of soil moisture and temperature data to the Noah LSM in NU-WRF. Each of these runs employed climatological GVF during the coupled NU-WRF.

The model-data fusion approaches to ICs were performed using SMAP data and direct insertion. The SMAP overpass provided nearly complete spatial coverage of the domain, but where necessary, a nearest-neighbor approach was used to interpolate for missing values. For these runs, SMAP was used as the top 5-cm soil moisture data on top of existing NARR and NLDAS-2 soil moisture profiles (identical to those taken off-the-shelf above), which were used for the remaining 3 soil layers (see Table 1 for layer specifications). While direct insertion is certainly not an advisable practice for operational purposes due to the relative biases of SMAP and LSM soil moisture climatologies, it serves a distinct purpose here to provide an upper bounds on what could be expected from data assimilation (where increments would ultimately be much smaller than what is seen here), as well as to see if the biases and noise of SMAP are indeed small enough to begin to consider such approaches as direct insertion. It is possible that introduction of SMAP on top of modeled profiles will result in a shock to the system and cause issues with equilibrium of the soil profile and associated fluxes and states, which also can be examined here. It should be noted that assimilation of SMAP into LIS spinups is an area of active research, and one that deserves independent treatment in future studies. Nonetheless, based on prior soil moisture assimilation experiments we would expect that such a spinup would fall somewhere near or within the spread of the three existing LIS spinups produced here.

e. In-Situ/Evaluation Data

For the offline evaluation of LoCo metric application over the SGP domain, data is acquired from the ARM-SGP network of sites and instruments at the Central Facility (CF) in Lamont, OK, the Plains Elevated Convection at Night (PECAN; Geerts et al. 2017) site at Ellis, KS, and 16 ARM Extended Facilities (EFs) across OK and KS (see Fig. 2ab) These include high-quality, nearly continuous meteorological, surface flux, and atmospheric profile measurements going back to the mid-1990's. Specifically, soil moisture from the recently installed Soil Temperature and Moisture Profile (STAMP) in-situ probes at the CF and 16 EFs are used, and represent a major improvement in ARM-SGP measurements of soil moisture quality and ancillary data (Cook 2018). Surface sensible and latent heat flux data from the Eddy Correlation Flux Measurement System (ECOR) towers at the CF and 7 EFs are also used. Temperature and humidity data at 2-meters is taken from the meteorological sites at CF and EFs. Vertical profiles of temperature and humidity are acquired at the CF and Ellis, KS sites. CF typically provides 4x/daily (2 daytime) launches, but as a result of the ESLCS campaign, 11 July produced hourly profiles from radiosonde that could be utilized in the LoCo analysis. Likewise, the Ellis, KS site was a supersite of the PECAN field campaign in Summer 2015 which took additional measurements from ground-based lidar (Weckwerth et al. 2016) to produce temperature and moisture profiles almost continuously during the daytime of 11 July. The radiosondes at CF and DIAL at Ellis were then used to characterize the diurnal structure and evolution of the PBL, derive PBL height estimates, and compare with NU-WRF simulations using LoCo metrics.

f. LoCo Metrics

The integrative nature and application of LoCo metrics to NWP and NU-WRF studies has been described in detail in Santanello et al. (2009, 2011, 2013, 2015). This includes the mixing diagram approach and evaporative fraction versus PBL height metrics that are employed in this study to better understand the impacts of soil moisture ICs on the coupled system, including the PBL response and the relative influence of atmospheric ICs as well. The reader is referred to Santanello et al. (2018) for an overview of LoCo metrics, and resources for the community.

4. Offline Soil Moisture Intercomparison

We first assess the behavior of near-surface soil moisture derived from SMAP, LIS, and in-situ measurements during the offline spinup period, followed by an intercomparison of the suite of soil moisture ICs and the coupled case study impacts.

Timeseries of near-surface soil moisture from the SMAP retrieval, LIS simulations, and in-situ STAMP probes are shown in Fig. 3 for three ARM-SGP extended facilities during Summer 2016. Overall, SMAP shows a comparable dynamic range to in-situ measurements, responding to precipitation events and drydown periods with little evidence of noise or spurious (outlier) values. To this end, the temporal consistency and absolute value of SMAP soil moisture appear realistic, and comparable to the STAMP measurements. Note that none of these sites were used as part of SMAP calibration/validation activities, and this is a true independent test of SMAP performance across sites with varying vegetation and soil characteristics.

The LIS simulations, on the other hand, have a distinct timeseries at E33 (Newkirk, OK) and E38 (Omega, OK) that shows a much narrower dynamic range and values on the wetter end of the soil moisture spectrum (relative to SMAP and STAMP). The spread across the three LIS simulations is rather small overall, but there are brief periods where the quality of atmospheric forcing (GDAS vs. Control; particularly in July and August) and, to a lesser extent, vegetation greenness (VIIRS vs. Control) do impact the soil moisture values. Regardless, the envelope of soil moisture across these simulations is one that is narrow, and it is bounded by a maximum (during precipitation spikes), and a minimum (during dry periods). The timescale of the drying events is controlled by the Noah LSM soil type and hydraulic parameters as specified by the lookup table at each site. At E33 and E38 (and the remaining 17 sites (not shown)), it is apparent that these parameters do not permit the model to dry down at a steep enough rate to reach the drier soil moisture levels observed by SMAP and STAMP.

Site E31 (Anthony, KS; Fig. 3c) is shown as an outlier, where LIS soil moisture tracks very close to that observed in terms of absolute range and drydown behavior. Interestingly, this is a site where the prescribed soil type in LIS is sand, but the observed type is silt loam. The hydraulic parameters corresponding to sand are the most extreme in terms of allowing for rapid drying and overall drier wilting point and minimum soil moisture values. So in a sense, at E31 the model obtains a better result for the wrong reasons by designating the site as sand in order for it to exhibit behavior like that of silt loam.

Figure 4 presents scatterplots of the timeseries data in Fig. 3, and compares the in-situ STAMP data directly with that of LIS and SMAP. The higher range of soil moisture values in LIS is apparent at E33 and E38, as are the comparable SMAP and STAMP values. The higher peaks during precipitation events in SMAP are also evident, and not unexpected as L-band sensing depths are much shallower during wet conditions (Liu et al. 2012; Escorihuela et al. 2010) and retrievals characterize a wetter and more dynamic quantity of soil moisture immediately after rainfall than at other times (Schneeberger et al. 2004; Rondinelli et al. 2015).

The linear slopes that can be seen in the data (e.g. Fig. 4a) actually reflect the inherent drying rates in each product. For example, at E33 LIS has a much lower drying rate (as seen in Figs. 3ab) compared with STAMP, which is reflected in the lesser slope of Fig. 4a (less than 1:1) as compared to that seen in Fig. 4b (nearly 1:1). The greater scatter seen at E38 in Fig. 4c is the result of mismatches in the number of modeled precipitation events vs. those observed locally at this site (as in Fig. 3b). At E31 (Fig. 4e), the tighter relationship and slope approaching 1:1 is the result of the unrealistic sandy soil parameters, as discussed above.

A summary of the performance of LIS and SMAP versus that of STAMP at all 17 ARM-SGP sites is presented in Table 2 in terms of bias, unbiased RMSE, and RMSE statistics over the JJA 2016 period. At 13 of the sites, SMAP bias is within +/- 0.04 m^3m^{-3} and unbiased RMSE is within 0.06 m^3m^{-3}, not far from the mission target of +/- 0.04 m^3m^{-3} at the SMAP calibration/validation sites, which is impressive given these sites are independent (uncalibrated) and varied in terms of soil type and land cover characteristics. The performance of LIS is largely inconsistent, with quite large bias and RMSE values (> 0.10 m^3m^{-3}) at a number of sites possibly due to mismatches in soil type versus those observed and the rigid parameter values that determine the overall soil moisture climatology in the Noah LSM. Other potential influences are the rooting depth, litter, and GVF in the LSM not matching what is observed at these sites. The unbiased RMSE statistics for LIS are much better, further indicating the importance of reducing the systematic error in the LIS results and the potential benefit of LSM calibration and parameter estimation approaches.

As discussed in Section 1, the ultimate impact of soil moisture on coupled NWP is felt through the surface flux connections of latent and sensible heat. Thus, it is not only important to intercompare the different soil moisture products as above, but also to assess what the implications of those soil moisture characteristics (and climatologies) are in terms of the surface energy balance and transfer of heat and moisture to the atmosphere. In Fig. 5, fluxes from the LIS simulations (averaged across the Control, GDAS, and VIIRS runs) at the three sites in Figs. 3–4 are shown versus those observed by the ARM ECOR stations. Averaged daytime diurnal cycles (hourly data) are calculated for the JJA 2016 period, and they show that at E33 and E38 there is a distinct overestimation of latent heat flux by LIS-Noah. Sensible heat fluxes are generally comparable between the model and flux towers.

When looking at all six ECOR sites (Fig. 5d), the overestimation of latent heat flux is more apparent, with differences approaching 200 Wm^{-2}. However, there is not the typical Bowen ratio compensation of lower sensible heat flux, which indicates (and is confirmed in Fig. 5e) that there is a significant overestimation of available energy in LIS at these sites. The tendency for LIS to have ample soil moisture then leads to the partitioning of excess available energy into latent, rather than sensible, heat flux. A detailed radiation analysis at these sites indicates that the extra available energy in LIS is a result of slight phase differences in downward shortwave radiation from LIS (NLDAS-2 forcing) vs. observed, in combination with a lower albedo in LIS vs. observed over this region. Overall, these higher evaporation rates into the atmosphere should have coupled L-A implications, which may lead to reduced PBL growth, more humidity and lower temperatures near the surface and in the PBL, and could impact moist processes and feedbacks that support clouds and precipitation.

5. Coupled Case Study Results

a. Intercomparison of Soil Moisture ICs

Near-surface soil moisture from the suite of IC approaches discussed in Section 2 are shown in Fig. 6, valid on 12 UTC on 9 June, 11 July, and 28 August 2015. Although the coupled case study focuses on 11 July, the June and August dates are shown to compare soil moisture conditions earlier and later in the summer season, both before and after the typical seasonal drydown in the SGP region. The spatial heterogeneity of SMAP across the domain is comparable to that of the other products in terms of overall variability and range of soil moisture (from dry to wet). The LIS simulations have the highest spatial resolution (1km) and therefore depict more local-scale features, many of which reflect the soil type dataset. The 'off-the-shelf' products show only coarser features, limited by the model resolution in each (ranging from 12.5–33 km). SMAP also shows regions of higher soil moisture (compared to other products) just after precipitation events (e.g. 11 July in Eastern Kansas), which is consistent with the behavior seen in Fig. 3 regarding the precipitation peaks in SMAP being larger than observed or modeled using thicker soil layers.

Generally, over dry regions SMAP tends to be drier than the other products, while GFS and NARR tend to be wetter. SMAP is known to dry down faster than the Noah LSM (Shellito et al. 2016), and likely has a true retrieval depth that is shallower than the published '5 cm', as near-surface soil layers (top 2–3 cm) dry down much faster than the 5–10 cm layer. The true SMAP retrieval depth is further complicated by vegetation effects and the soil moisture itself, and likely varies both in time and space as a result. The coarse model products (using Noah LSM and the 0-10cm layer) are wetter. This is likely due to a combination of a deeper top soil layer (0–10-cm depth), inaccuracy of the soil hydraulic parameters, and the coarse horizontal model resolution, which all contribute to restricting the model's response to higher resolution land surface data. LIS, on the other hand, shows regions of very dry soil (consistent with the SMAP patterns) as a result of retaining the 1 km soils information as well as local vegetation and precipitation patterns that allow for more extensive dry downs (particularly in late August).

It should be noted that it is not possible to objectively evaluate which is the most accurate soil moisture IC. Each IC approach provides a representation of soil moisture that is reliant on (physical or retrieval) model assumptions, and in-situ data is too sparse to convincingly validate each across a large 1 km resolution domain. However, understanding the differences and what causes them (e.g. resolution, SMAP retrieval, layer depth, input parameters, etc.) is key to understanding the potential coupled impacts of each IC. To better parse out these differences, Fig. 7 presents PDFs of the soil moisture ICs in Fig. 6 on each of the three dates. SMAP is skewed towards drier values, as discussed above, and the coarser products are much wetter (GFS, NLDAS) with the LIS runs in between. On 11 July, there is a bimodal distribution in SMAP and the LIS runs, as a result of each capturing the spatial heterogeneity generated by recent localized precipitation over part of the domain creating distinct wet and dry regimes. Notably, the coarser products (GFS) do not capture this bimodal distribution nearly as well. Another striking difference can be seen on 28 August, where GFS is much wetter and more narrowly distributed as compared to LIS, SMAP, and NLDAS-2.

Overall, there are three important takeaways from Figs. 6–7, in that 1) the climatologies of soil moisture differ significantly based on the source of the IC (i.e. SMAP, high-resolution LSM spinup, or off-the-shelf products), 2) ICs based solely on SMAP tend to be drier overall, but capture the spatial variability of the region, and 3) stark differences in spatial distributions suggest that the choice of IC is likely to have significant downstream coupled impacts across the domain.

b. Coupled Case Study – 11 July 2015

I. LOCO ANALYSIS

Coupled NU-WRF simulations, initialized by the suite of soil moisture conditions in Fig. 6b, were performed for 24h beginning at 12 UTC on 11 July, 2015. As described in Section 3, SMAP soil moisture values were directly inserted as the top layer soil moisture in the NLDAS-2 and NARR profiles and are referred to as SMAP+NARR and SMAP+NLDAS. The impact of each IC is reflected in the process chain (Eq. 1) variables that connect soil moisture to evaporation (sensible and latent heat flux), PBL evolution (mixed-layer temperature, humidity), PBL height, ambient weather (2-meter temperature and humidity), and clouds and precipitation. Seven hours into the simulation (19 UTC), there are large differences across the runs in each of these variables, reflecting the relative impact of wetter soil moisture conditions on increased evaporation, decreased PBL growth, lower temperature, higher humidity, and modification of precipitation intensity and location. An example is shown in Fig. 8 in terms of soil moisture IC differences between NARR and LIS-GDAS at 12 UTC, and the downstream impacts on latent and sensible heat flux (~200-300 Wm-2), PBL height (~800-1000 meters), and 2-meter temperature and humidity (~2-3K; ~3-4 g kg-1, respectively) in midafternoon (19 UTC). The impact of the soil moisture IC differences and drier (wetter) regions can clearly be seen carried through towards higher (lower) sensible (latent) heat fluxes, larger (reduced) PBL growth, warmer (cooler) temperatures, and lower (higher) humidity at 2-meters. The relatively drier LIS-GDAS conditions overall support higher sensible heat flux, PBL heights, and temperatures later in the day throughout much of the domain, illustrating the coupled impacts of soil moisture ICs.

While Fig. 8 presents a standard single variable assessment, the integrated metrics of LoCo can be used for a more qualitative and comprehensive assessment of the fully coupled impacts of soil moisture ICs. Figure 9 presents the mixing diagram analyses at the ARM CF and Ellis, KS sites for each simulation, along with the derived Bowen and entrainment ratios (as in Santanello et al. 2009). As shown in Fig. 6b and based on SMAP, the CF site is located in a wet region having just received precipitation and Ellis is in the western much drier part of the domain, so there is a natural contrast in conditions at these two sites. At the CF site, the mixing diagram signatures are vertically oriented with little change in humidity throughout the day, as a result of only moderate PBL growth and entrainment and little spread across simulations (Fig. 10a, 11a), as might be expected for a wet site. There is only small divergence in the co-evolution of 2-meter temperature and humidity across the runs, with the exception of the GFS simulation, which employs GFS atmospheric IC/BCs, whereas the remainder of the simulations use NARR. Thus, soil moisture does not seem to impact the results nearly as much as the choice of GFS or NARR atmospheric data at this location.

At the Ellis site, there is much larger diurnal variability in temperature and humidity, and larger entrainment fluxes of dry and warm air into the PBL, as would be expected at a dry site. Bowen ratios also vary from 0.73 to 2.55 depending on the choice of soil moisture IC. Once again, GFS is the outlier and evolves differently over time. Overall, the soil moisture ICs are directly reflected in the surface Bowen ratio and EFs (Fig. 9b) with the SMAP-based runs (driest soil moisture) having the lowest EF (high sensible heat flux) and the wetter off-the-shelf products (GFS, NLDAS-2) producing the highest values (high latent heat flux). These differences in surface energy balance are amplified at Ellis (vs. that seen at CF) by the much larger PBL growth and spread across simulations (Fig. 10b, 11b). At this site, the wetter ICs of GFS and NLDAS tend to limit PBL growth while the SMAP and LIS-GDAS runs easily reach over 3 km. It should also be noted that all runs tend to overestimate PBL height at both sites throughout the daytime period, so those with wetter ICs overall that limit PBL growth are closer to observed.

Overall, these two contrasting sites demonstrate that soil moisture ICs can impact PBL and L-A coupling, but the magnitude depends on the relative range and spread of soil moisture (and atmosphere vs. soil limited regime) across the ICs, and whether the PBL is sensitive to the surface flux partitioning. A look at the four Noah LSM soil layers for each of the simulations provides further insight as to the potential role of soil moisture ICs. At the CF site (Fig. 12a), the second layer of soil moisture is generally similar across all IC products and greater than 0.25 m3 m-3.

This second layer (10-40 cm) represents more of the root zone, and thus controls the majority of evapotranspiration. These soil moisture values are all in the atmosphere-limited regime, and there is little difference across simulations with each producing high evaporative fraction that limits PBL growth.

On the other hand, at the Ellis site (Fig. 12b) conditions are much drier, particularly in the root zone, for the LIS and NARR ICs, but wetter in the second layer in NLDAS and GFS. This creates a disparity in EF (Fig. 10b) across ICs. Furthermore, the SMAP direct insertion into NLDAS-2 versus NARR produces different surface flux and resultant PBL growth as a result of the wet (NLDAS-2) vs. dry (NARR) root zone of each. At this site, it is not the dry SMAP near-surface soil moisture that controls the surface energy balance, rather the deeper soil layers which in this case are derived from NLDAS-2 and NARR. This is an important result in that it highlights the potential limited role of SMAP on L-A coupling if combined with other products that are not consistent (and why a direct insertion approach is not recommended), even when SMAP is much drier than the layers below.

II. LAND VS. ATMOSPHERIC IC IMPACTS

Having parsed out the soil moisture IC impacts, it is worthwhile isolating and examining the impact of the atmospheric ICs as well, given the outlier behavior of the GFS simulation seen in Fig. 9. Fig. 13 shows mixing diagrams and evaporative fraction versus PBL height analyses, for only the GFS and NARR simulations, at the CF and Ellis sites in addition to two other sites (36.0 N, 100.0 W; 39.0 N, 97.0 W) across the domain that represent different surface and atmospheric conditions. Mixing diagrams can be used to show the diurnal behavior of humidity (x-axis) and temperature (y-axis) simultaneously with the surface latent and sensible heat flux vectors and atmospheric response (PBL entrainment and advection) vectors. In Fig. 13, the mixing diagram plots show an initial dry bias in GFS 2-meter humidity at 12 UTC (7am local time) relative to that of NARR and that observed at CF and Ellis. The ensuing daytime evolution of temperature and humidity then differs considerably across the sites. At the CF site, GFS and NARR remain parallel to each other, with the initial dry GFS bias persisting throughout the day (with comparable temperature evolution in each). Fig. 13b shows that the EF (i.e. land ICs) and PBL height (i.e. initial atmospheric profiles) are similar in GFS and NARR, and hence there was no coupled mechanism to impact the GFS humidity bias during the day. In addition, low-level winds were weak and variable, thus limiting any potential impact of horizontal advection.

In contrast, at Ellis the initial dry bias in GFS is overcome by NARR by the end of the day, with NARR drying out significantly. This can be explained by the drier soil moisture IC at this site in NARR, which promotes lower EF, combined with significant PBL growth and dry air entrainment, and generates a PBL with lower humidity. NARR humidity ends up further from the observations at this site, likely due to a dry IC of soil moisture. As NARR atmospheric IC/BCs were used to drive all the soil moisture IC permutations except for GFS, this is an important result particularly as the LIS and SMAP ICs tend to be even drier than the default NARR ICs (and thus even further from observed 2-m humidity).

At the third site (Fig. 13ef), the GFS initial dry bias is apparent, but then eroded over time with GFS approaching similar humidity values of NARR (which does not dry out during the day). Each has similar soil moisture ICs at this site, so the land influence is eliminated. However, a closer look at the vertical profiles of temperature and humidity (not shown) indicate that in GFS the PBL grows more slowly and into a more humid layer than NARR, which tends to increase and cap the overall humidity in the PBL including at 2 meters. At the final site (Fig. 13gh), the GFS dry bias is also reduced over time as a combination of afternoon moistening in GFS and drying in

NARR, despite having similar soil moisture ICs and evaporative fraction. This can be attributed to a combination of deeper PBL growth and dry air entrainment in NARR, along with GFS eventually reaching a phase of PBL growth into a more humid layer in the afternoon (not shown).

Overall, these results demonstrate that land, atmospheric and PBL ICs and processes can have varying relative impacts on L-A coupling and ambient weather prediction. They also suggest that it is unlikely that changes in a single component of initialization (e.g. soil moisture) will have uniform or spatially/temporally consistent impacts on NWP across the domain of interest, and will still be modulated by the land or atmospheric conditions (and inherent biases) being introduced elsewhere in the coupled system.

III. AMBIENT WEATHER STATISTICS

Following this approach, typical NWP benchmarking statistics can now be examined under the context of LoCo and land versus atmospheric ICs. The LoCo analysis in Figs. 9–12 focused on two well-instrumented sites with contrasting soil moisture conditions. However, as shown in Fig. 8, there are widespread and larger impacts seen across the full domain particularly with respect to 2-meter temperature (~2–6 K) and humidity (~2–6 g kg^{-1}). It is these ambient weather impacts that are particularly important to NWP operational centers in terms of forecast performance and improvement, as well as public perception.

Figure 14 shows the 2-m RMSE and bias statistics timeseries for temperature and humidity from each of the coupled NU-WRF simulations, for the 24h period beginning at 12 UTC on 1 July 2015. These statistics were calculated hourly based on the NCEP ADP Global Upper Air Surface Weather Observations (https://rda.ucar.edu/datasets/ds337.0/) dataset that includes 153 sites sampled across the SGP domain. This is a typical NWP center approach, focused on sensible weather impacts that are readily observable, when assessing the impacts of new datasets, parameters, physics, ICs, and data assimilation. A bird's eye assessment of Fig. 14 in this context may be that soil moisture ICs do not have large or systematic impacts on temperature and humidity forecasts, and in effect it would be difficult to conclude which is the 'best' IC. For the daytime period (7am–7pm), it could be argued that NLDAS-2 and GFS have the lowest RMSE values and biases, and that SMAP and the LIS runs have the largest. This would be counterintuitive to the idea that NLDAS-2 and GFS are coarse, default and off-the-shelf products whereas LIS and SMAP are higher-resolution and observationally-driven. This may lead to conclusions that improved land ICs do not improve NWP.

However, based on the knowledge gleaned in the prior sections using integrated LoCo metrics, we can better understand these results in the context of the role of land versus atmospheric ICs, in particular that of soil moisture and SMAP. The lowest daytime temperature errors (Fig. 14ab) are seen in GFS and NLDAS-2, and the highest are in the SMAP+NARR simulation. As GFS and NLDAS are the wettest ICs in terms of soil moisture, these act to reduce the overall warm bias across the domain, while the SMAP and NARR runs are the driest, which tends to amplify the warm bias over the domain. In terms of temperature bias overall, there is a slight warm bias at initialization that is then amplified throughout the day, which is likely a result of a net radiation (driven by downward shortwave and underestimation of localized cloud cover) overestimation at the surface driven by the NU-WRF (GSFC) radiation and microphysics schemes in combination with lower than observed surface albedo in LIS (as in the offline case). However, while the temperature biases (Fig. 14b) appear to remain relatively constant over the daytime in each of the runs, the actual locations of these biases shift significantly over time from the northern to southern part of the domain (Fig. 15bcde).

The humidity statistics (Fig. 14d) also indicate that NLDAS-2 tends to have the lowest bias, and that all the NARR-driven runs tend to dry out rapidly during the daytime despite higher quality atmospheric ICs, with the driest (SMAP+NARR) runs performing worst. The GFS run and the IC dry bias can be clearly seen here, even across all 153 sites in this analysis, and overall remains constant during the daytime. Once again, the wetter soil moisture ICs tend to perform best, as they are countering an inherent warm, dry bias in the coupled system. The envelope of RMSE and biases in these plots is generally narrow, but these are averages across many points and a large domain.

The inflection point seen in the temperature bias at 7–9pm is a notable feature as well in these results, and is due to late afternoon precipitation in the northern part of the domain which is overestimated compared to observations, and tends to cool down the region overall (compensating for the warm biases in the south). What follows in the SMAP direct insertion runs is a linear decrease in temperature bias, and a significant cooling that takes place during the entire nighttime period (unrelated to precipitation) particularly in the western part of the domain. This is, in fact, due to the direct insertion of much drier SMAP values on top of NARR and NLDAS-2 profiles. As discussed earlier, this approach is not recommended as it disrupts the soil moisture and temperature equilibrium from the top layer vs. three deeper layers in the Noah LSM. This direct insertion did not show negative impacts during the daytime, and as mentioned it was often the root zone soil moisture of NARR and NLDAS that dominated the surface energy balance. At nighttime, however, the very low SMAP soil moisture and 5-cm upper soil layer led to changes in the thermal properties of the near-surface soil that promote rapid cooling at night. The thermal impacts of the daytime were overcome by the dominance of evaporation, but it is evident that a more robust approach to merging SMAP with existing soil profiles from other products should be performed if using as ICs for NWP.

Fig. 15 shows an example of how these 2-meter statistics vary in space and time, and in response to the initial SM differences in GFS and NARR. These results indicate that there is much more divergence across runs regionally and at specific sites than is evident in the lumped timeseries statistics, often ranging in magnitude to near 6 K and 6 g kg^{-1} in temperature and humidity, respectively. The initial dry bias in GFS is evident across much of the central and southern SGP (Fig. 15g), while NARR shows a slight wet bias. Because NARR soil moisture is drier than GFS (Fig. 15a) especially over the central part of the domain, NARR ends up with a strong dry humidity bias by the end of the day, whereas GFS improves its initial dry atmospheric bias where the soil tends to be wetter. The location of the warm bias and shift from north to south over the course of the day mentioned above is also apparent in Figs. 15bd. These plots show the components that must be simultaneously considered when interpreting NWP statistics and assessing new parameterization or initialization approaches, including the background atmospheric IC biases, the change in land surface ICs, and the evolution of each as dictated by the LSM and PBL schemes throughout the day.

Overall, this analysis demonstrates that the aggregated statistics commonly employed by NWP centers are often not systematic in space or time, and can miss the important nuances and drivers behind them thus confounding the conclusions made. In essence, each point in Fig. 15 has its own 'coupling story' that is dependent on many factors, and thus a change to the land IC (or LSM physics) is unlikely to produce uniform impacts or improvements. At the same time, any perceived improvement could be compensating for errors elsewhere in the system. Although it takes a bit more work using integrated analyses to understand these impacts, it becomes necessary for a true assessment of the impact of soil moisture (or any other IC, physics package, or parameter dataset) in coupled prediction.

6. Discussion and Conclusions

This study provides a review of current soil moisture initialization approaches used in NWP, and in particular those employed by the regional weather and climate (i.e. WRF) research and operational communities. Land ICs are often overlooked by atmospheric scientists, and as a result there have been a wide range of approaches employed using vastly different datasets in terms of quality and resolution. Soil moisture tends to get most of the focus (vs. soil temperature) due to its strong control on surface energy balance and surface fluxes which are the only true LSM variables that the atmospheric model is sensitive (and coupled) to. Here, we isolate the impacts of these varied soil moisture initialization approaches on coupled forecasts using a very pragmatic, yet integrative (in the L-A sense) approach using NASA's LIS, SMAP, and NU-WRF assets.

Results and their implications for NWP modeling communities are as follows:

1) Offline analysis of satellite, in-situ, and LSM products confirms that SMAP soil moisture performs quite well in terms in spatial and temporal consistency (i.e. low noise), capturing heterogeneity, precipitation and drydown events, and overall looks like a 'real' observable soil moisture field.

2) There remains an observability issue due to differing LSM and observed (satellite and in-situ) soil moisture climatologies that are largely due to differences between LSM physics and the actual soil hydraulic properties and vegetation characteristics which affect the satellite and in-situ measurements.

3) There is a wide variation in the spatial distribution of soil moisture across commonly used NWP initialization approaches, including those from satellite-infused, high-resolution LSM spinup, and 'off-the-shelf' atmospheric model based products.

4) The sensitivity of coupled impacts is not limited to the near-surface soil layer as the root zone may still play a dominant role in governing surface fluxes and L-A coupling, thus limiting the potential impact of near-surface layer observations in isolation.

5) Coupled impacts of land ICs are clearly visible downstream in the NWP forecasts (including surface fluxes, PBL evolution and entrainment, and ambient weather), and can be better understood and quantified using integrated LoCo metrics.

6) By simultaneously assessing land vs. atmospheric ICs in a LoCo framework, the question of whether improvements in traditional NWP statistics are achieved for the right reasons can better be addressed, and in turn shed light on the true potential impact of improved soil moisture ICs.

It should be noted that additional case study simulations were performed in June 2015, (Fig. 4a), and the results were largely consistent with those from 11 July. Specifically, a warm atmospheric IC bias dominated the region, and as a result the wettest soil moisture ICs (once again the coarse GFS and NLDAS-2 products) produced the best 2-meter statistics. As for the July case, in isolation this would suggest that the coarse soil moisture products are better than the high-resolution or observed products, when actually the coarse products are only best for this particular modeling system and atmospheric forcing where they are correcting inherent biases. Studies that show uniform impacts (e.g. drying) after satellite assimilation across a wide domain are likely to see some improvements in sub-regions simply as a matter of luck, correcting for inherent model biases (and vice-versa for degradation).

This underscores the importance of understanding inherent coupled model behavior before introducing new datasets or ICs, so that their impacts can be more accurately assessed. As satellite data continues to improve in quality and resolution, there can be greater incorporation of more accurate observations into coupled models. Understanding their impacts requires quantification of process-chain impacts in order to avoid compensating errors. It will be difficult for highly-tuned systems to incorporate new datasets and see direct improvements as a result, but the approaches here will help aid in identifying what remaining model biases and deficiencies exist in order to further fully integrated model development and improvement.

While the statistical significance of the limited number of deterministic simulations performed here is clearly lacking, the methodology remains valid as an approach that can be adopted by operational and cycled modeling centers. It is clear from this work that biases and forecast errors can be best understood and improved via integrated (LoCo-type) assessment of relative impacts of land surface and atmospheric (specifically in PBL vertical profiles) ICs that are used to drive the coupled simulations. One example where this can be applied is in the work of Fang et al. (2018), where novel approaches to assimilating satellite-based land surface temperature were performed, and the regional and temporal impacts on 2-meter statistics could be better understood with more integrated process-level understanding. This approach also reduces the potential for mischaracterizing forecast impacts and improvements that may result from compensating errors or misattribution.

In terms of the offline soil moisture analysis, it is clear that the governing soil and vegetation physics and parameters in the Noah LSM do not allow for soil drying behavior that is observed from satellite or in-situ. While likely due primarily to soil texture and rigid lookup tables of soil hydraulic properties (setting the maximum and minimum range of soil moisture), there are also potential impacts on soil moisture dynamics from improper rooting depth specification, lack of leaf litter, and inconsistencies in GVF in the LSM versus what is observed. Incidentally, the wetter LIS runs (compared with SMAP) were actually advantageous in the coupled runs due to the warm, dry bias of NARR and GFS. Even when modifying the upper 10cm of soil layering to create a 2, 3, or 5cm top layer, the soil drying dynamics were only marginally impacted indicating there are structural limitations in the LSM that prohibit it from having the observability necessary for unbiased data assimilation or direct comparison of soil moisture with satellite or in-situ observations. Clearly, the structural deficiencies in the LSM and systematic errors need to be addressed via calibration and parameter estimation approaches in order to better match the soil moisture dynamics with those observed. However, avoiding so-called 'effective' parameters that absorb additional unrelated model errors and estimating physically meaningful soil characteristics remains a challenge.

These results highlight the critical nature of soil type information, parameter lookup tables, and the difficulty in modeling soil moisture dynamics at the local scale using only coarse soils information. It also highlights the relative inflexibility of LSM parameters and soil physics, whereby soil moisture results can only be improved to a limited degree when introducing improved, high-resolution inputs such as atmospheric forcing and vegetation characteristics. To address this, there are community efforts underway in GEWEX focused on reexamination of pedotransfer functions and soils in LSMs, and also to improve the collaboration between the soils and LSM communities themselves.

Another interesting result (not shown) is that the soil moisture IC differences at 12 UTC tend to diminish over time throughout the domain (e.g. by 19 UTC). When examining timeseries at specific sites, it is apparent that when comparing two simulations with different ICs, the wetter of the simulations tends to dry down over time and at a more rapid rate than the drier simulations. This can be traced once again to the Noah LSM soil physics and hydraulic parameters that determine the levels of atmosphere and soil limited evaporation. In the case of NLDAS-2 vs. SMAP, for example, SMAP is already very dry and soil-limited such that it doesn't change much or dry out further while at the same time the wetter NLDAS-2 is in a very active evaporative stage and dries out rapidly, thus converging towards the SMAP values. As a result, it is common for IC differences to be dampened over time due to evaporative physics, as opposed to an initial perturbation that is amplified. Exceptions to this occur when wetter soil moisture promotes precipitation, and vice-versa, over a more extended period of time.

A related variant of soil moisture ICs can be generated by performing data assimilation during an offline spinup (e.g. Santanello et al. 2016). Based on the largely incremental soil moisture DA impacts in studies to date combined with the results here in terms of the narrow envelope of LIS simulations with different parameters and forcing, it is likely that SMAP assimilation will not lead to vastly different results or ICs. The CDF matching approach to bias correction makes large impacts even less likely, as discussed in Kumar et al. 2015 and Navari et al. 2018. The SMAP direct insertion approach taken here, while not advisable (but still used/published in the community), was chosen as a brute force approach to see what the maximum impact of satellite soil moisture might be on the IC, while acknowledging that any proper EnKF assimilation is likely to impact the ICs to a much lesser degree and be just another permutation of a LIS run. Only via model calibration (discussed above, specifically targeting hydraulic properties) that addresses systematic errors would we expect more distinct ICs and impacts on the LSM climatology and drydown behavior.

Ongoing and future work on this topic includes performing formal EnKF data assimilation with SMAP and LIS, as well as LIS calibration using in-situ networks in an effort to improve LSM observability, and reduce the negative impacts of typical satellite bias correction approaches. In addition, the capabilities of SMAP to detect agricultural and irrigation practices (largely missing or mischaracterized in LSMs) are being evaluated in an effort to improve model-data-fusion efforts and aid in offline and coupled model development. It is clear that the community is now demonstrating that the land states and strength of land-atmosphere coupling can play a significant role in the accuracy of ambient weather forecasts. Improving the initial conditions of soil moisture, temperature, and vegetation using NASA satellite observations and assimilation systems therefore becomes even more critical, and the combination of NASA's SMAP, LIS and NU-WRF resources will continue to be used to develop and test these approaches and coupled impacts. As a result, the continuity of missions (beyond SMAP) to provide accurate, global data records of near-surface soil moisture remains important to consider going forward at NASA and other space agencies.

Acknowledgements. This work was supported by the NASA Science Utilization of SMAP (SUSMAP) program and Jared Entin. The authors would like to thank David Cook for extensive consultation on the ARM-SGP flux and soil moisture measurements, and Rajat Bindlish for guidance on interpretation of the SMAP retrievals.

References

Ament, F., and C. Simmer, 2006: Improved representation of land-surface heterogeneity in a non-hydrostatic numerical weather prediction model. *Bouldary-Layer Meteorol.,* **121,** 153-174.

Angevine, W. M., E. Bazile, D. Legain, and D. Pino, 2014: Land surface spinup for episodic modeling. *Atmos. Chem. Phys.,* **14,** 8165-8172.

Carrera, M. L., B. Bilodeau, S. Belair, M. Abrahamowicz, A. Russell, and X. Wang, 2018: Assimilation of passive L-band microwave brightness temperatures in the Canadian Land Data Assimilation System: Impacts on short-range Numerical Weather Prediction. *J. Hydrometeorol.,* under review

Case, J. L., W. L. Crosson, S. V. Kumar, W. M. Lapenta, C. D. Peters-Lidard, 2008: Impacts of High-Resolution Land Surface Initialization on Regional Sensible Weather Forecasts from the WRF Model. *J. Hydrometeor*, **9**, 1249–1266.

Case, J. L., S. V. Kumar, J. Srikishen, G. J. Jedlovec, 2011: Improving Numerical Weather Predictions of Summertime Precipitation over the Southeastern United States through a High-Resolution Initialization of the Surface State. *Wea. Forecasting*, **26**, 785–807.

Chan, S. K., R. Bindlish, P. O'Neill, T. Jackson, E. Njoku, S. Dunbar, J. Chaubell, J. Piepmeier, S. Yueh, D. Entekhabi, A. Colliander, F. Chen, M.H. Cosh, T. Caldwell, J. Walker, A. Berg, H. McNairn, M. Thibeault, J. Martínez-Fernández, F. Uldall, M. Seyfried, D. Bosch, P. Starks, C. Holifield Collins, J. Prueger, R. van der Velde, J. Asanuma, M. Palecki, E.E. Small, M. Zreda, J. Calvet, W.T. Crow, Y. Kerr, 2018: Development and assessment of the SMAP enhanced passive soil moisture product. *Rem. Sens. Environ.,* **204,** 931-941.

Chen, F., K. W. Manning, M. A. LeMone, S. B. Trier, J. G. Alfieri, R. Roberts, M. Tewari, D. Niyogi, T. W. Horst, S. P. Oncley, J. B. Basara, and P. D. Blanken, 2007: Description and Evaluation of the Characteristics of the NCAR High-Resolution Land Data Assimilation System. *J. Appl. Meteor. Climatol.,* **46,** 694–713.
doi: http://dx.doi.org/10.1175/JAM2463.1

Collow, T. W., A. Robock, and W. Wu, 2014: Influences of soil moisture and vegetation on convective precipitation forecasts over the United States Great Plains. *J. Geophys. Res. Atmos.,* **119,** 9338-9358.

Cook, D. R., 2018. Soil Temperature and Moisture Profile (STAMP) System Instrument Handbook. Ed. by Robert Stafford, DOE ARM Climate Research Facility. DOE/SC-ARM-TR-186.

Daniels, E. E., R. W. A. Hutjes, G. Lenderink, R. J. Ronda, and A. A. M. Holtslag, 2014: Land Surface Feedbacks on Spring Precipitation in the Netherlands. *J. Hydrometeorol.,* **16,** 232-243.

Dillon, M. E., E. A. Collini, and L. J. Ferreira, 2016: Sensitivity of WRF short-term forecasts to different soil moisture initializations from the GLDAS database over South America in March 2009. *Atmospheric Research,* **167,** 196-207.

Dirmeyer, P.A., L. Chen, J. Wu, C. Shin, B. Huang, B.A. Cash, M.G. Bosilovich, S. Mahanama, R.D. Koster, J.A. Santanello, M.B. Ek, G. Balsamo, E. Dutra, and D.M. Lawrence, 2018: Verification of Land–Atmosphere Coupling in Forecast Models, Reanalyses, and Land Surface Models Using Flux Site Observations. *J. Hydrometeor.,* **19,** 375–392, https://doi.org/10.1175/JHM-D-17-0152.1

Dirmeyer, P. A., and Co-authors, 2016: Confronting weather and climate models with observational data from soil moisture networks over the United States. *J. Hydrometeor.,* **17,** 1049–1067.

———, and S. Halder, 2016: Sensitivity of surface fluxes and atmospheric boundary layer properties to initial soil moisture variations in CFSv2. Wea. Fcst., **31,** 1973-1983, doi: 10.1175/WAF-D-16-0049.1.

Dy, C. Y., and J. C. H. Fung, 2016: Updated global soil map for the Weather Research and Forecasting model and soil moisture initialization for the Noah land surface model. *J. Geophys. Res. Atmos.,* **121,** 8777-8800.

Environmental Modeling Center, 2003: The GFS atmospheric model. NCEP Office Note 442, 14 pp. [Available online at http://www. emc.ncep.noaa.gov/officenotes/newernotes/on442.pdf.]

Escorihuela, M. J., A. Chanzy, J. P. Wigneron, and Y. J. Kerr, 2010: Effective soil moisture sampling depth of L-band radiometry: A case study. *Remote Sens. Env.,* 7 pp.

Fang, L., X. Zhan, C. R. Hain, J. Yin, J. Liu, and M. A. Schull, 2018: An Assessment of the Impact of Land Thermal Infrared Observation on Regional Weather Forecasts Using Two Different Data Assimilation Approaches. *Remote Sens.,* **10,** 625.

Ferguson CR, JA Santanello, and P Gentine. 2016. Enhanced Soundings for Local Coupling Studies Field Campaign Report. Ed. by Robert Stafford, DOE ARM Climate Research Facility. DOE/SC-ARM-16-023.

Findell, K. L. and E. A. B. Eltahir, 2003a. Atmospheric controls on soil moisture-boundary layer interactions: Part I: Framework development. *J. Hydrometeor.,* **4,** 552-69.

Ford, T. W., A. D. Rapp, and S. M. Quiring, 2015: Does Afternoon Precipitation Occur Preferentially over Dry or Wet Soils in Oklahoma? *J Hydrometeorol*, **16,** 874-888.

Geerts, B., D. Parsons, C.L. Ziegler, T.M. Weckwerth, M.I. Biggerstaff, R.D. Clark, M.C. Coniglio, B.B. Demoz, R.A. Ferrare, W.A. Gallus, K. Haghi, J.M. Hanesiak, P.M. Klein, K.R. Knupp, K. Kosiba, G.M. McFarquhar, J.A. Moore, A.R. Nehrir, M.D. Parker, J.O. Pinto, R.M. Rauber, R.S. Schumacher, D.D. Turner, Q. Wang, X. Wang, Z. Wang, and J. Wurman, 2017: The 2015 Plains Elevated Convection at Night Field Project. *Bull. Amer. Meteor. Soc.,* **98,** 767–786, https://doi.org/10.1175/BAMS-D-15-00257.1

Gomez, I., V. Caselles, M. J. Estrela, J. M. Sanchez, E. Rubio, 2018: Simulation of surface energy fluxes and meteorological variables using the Regional Atmospheric Modeling System (RAMS): Evaluating the impact of land-atmosphere coupling on short-term forecasts. *Ag. Forest Meteorol.,* **249,** 319-334.

Gomez, I., R. J. Ronda, V. Caselles, M. J. Estrela, 2016: Implementation of non-local boundary layer schemes in the Regional Atmospheric Modeling System and its impact on simulated mesoscale circulations. *Atmospheric Research,* **180,** 24-41.

Gomez, I., V. Caselles, and M. J. Estrela, 2015: Impacts of soil moisture content on simulated mesoscale circulations during the summer over eastern Spain. *Atmospheric Research,* **164-165,** 9-26.

Hirsch, A.L., J. Kala, A.J. Pitman, C. Carouge, J.P. Evans, V. Haverd, and D. Mocko, 2014: Impact of Land Surface Initialization Approach on Subseasonal Forecast Skill: A Regional Analysis in the Southern Hemisphere. *J. Hydrometeor.,* **15**, 300–319, https://doi.org/10.1175/JHM-D-13-05.1

Jacobs, S. J., A. J. E. Gallant, and N. J. Tapper, 2017: The Sensitivity of Urban Meteorology to Soil Moisture Boundary Conditions: A Case Study in Melbourne, Australia. *J. Applied Met. Clim.,* **56,** 2155-2172.

Johnson, Z.F. and N.M. Hitchens, 2018: Effects of Soil Moisture on the Longitudinal Dryline Position in the Southern Great Plains. J. Hydrometeor., **19,** 273–287, https://doi.org/10.1175/JHM-D-17-0091.1

Kala, J. , Evans, J. P. and Pitman, A. J., 2015: Influence of antecedent soil moisture conditions on the synoptic meteorology of the Black Saturday bushfire event in southeast Australia. *Q.J.R. Meteorol. Soc.,* **141,** 3118-3129. doi:10.1002/qj.2596

Kalverla, P. C., G.-J., Duine, G.-J. Steeneveld, T. Hedde, 2016: Evaluation of the Weather \ Research and Forecasting Model in the Durance Valley Complex Terrain during the KASCADE Field Campaign. *J. Appl. Meteor. Clim.,* **55,** 861-882.

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, https://doi.org/10.1175/2011JCLI3978.1

Koster, R. D., and Co-authors, 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.

———, and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **306**, 1138–1140.

Kumar, S. V., P. A. Dirmeyer, C. D. Peters-Lidard, R. Bindlish, and J. Bolten. 2017: Information theoretic evaluation of satellite soil moisture retrievals. Remote Sensing of Environment, 204, 392-400 [10.1016/j.rse.2017.10.016]

———, Peters-Lidard, C. D., Santanello, J. A., Reichle, R. H., Draper, C. S., Koster, R. D., Nearing, G., and Jasinski, M. F., 2015: Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. Hydrol. Earth Syst. Sci., 19, 4463-4478, https://doi.org/10.5194/hess-19-4463-2015.

———, R.H. Reichle, C.D. Peters-Lidard, R.D. Koster, X. Zhan, W.T. Crow, J.B. Eylander, and P. R. Houser, 2008: A Land Surface Data Assimilation Framework using the Land Information System: Description and Applications. *Advances in Water Resources*, **31,** 1419-1432, DOI:10.1016/j.advwatres.2008.01.013

Lawston, P. L., J. A. Santanello, S. V. Kumar, 2017: Irrigation signals detected from SMAP soil moisture retrievals. *Geophys. Res. Letters,* **44,** 11,693-12,036.

Lawston, P.M., J.A. Santanello, B.F. Zaitchik, and M. Rodell, 2015: Impact of Irrigation Methods on Land Surface Model Spinup and Initialization of WRF Forecasts. *J. Hydrometeor.,* **16,** 1135–1154, https://doi.org/10.1175/JHM-D-14-0203.1

Lin, T.-S., and F.-Y. Cheng, 2016: Impact of Soil Moisture Initialization and Soil Texture on Simulated Land–Atmosphere Interaction in Taiwan. *J. Hydrometeorol.,* **17,** 1337-1355.

Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M. , Wagner, W., McCabe, M.F., Evans, J.P., van Dijk, A.I.J.M., 2012: Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sensing of Environment,* **123,** 280-297, doi: 10.1016/j.rse.2012.03.014.

Massey, J. D., W. J. Steenburgh, J. C. Knievel, and W. Y. Y. Cheng, 2016: Regional Soil Moisture Biases and Their Influence on WRF Model Temperature Forecasts over the Intermountain West. *Wea. Forecasting,* **31,** 197-216.

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

Mueller, B., and S. I. Seneviratne, 2012: Hot days induced by precipitation deficits at the global scale. *Proc. Nat. Acad. Sci.*, **109**, 12398–12403.

Peters-Lidard, C.D., E. M. Kemp, T. Matsui, J. A. Santanello, Jr., S. V., Kumar, J. P. Jacob, T. Clune, W.-K. Tao, M. Chin, A. Hou, J. L. Case,, D. Kim, K.-M. Kim, W. Lau, Y. Liu, J.-J. Shi, D. Starr, Q. Tan,, Z. Tao, B. F. Zaitchik, B. Zavodsky, S. Q. Zhang, M. Zupanski, 2015:

Integrated modeling of aerosol, cloud, precipitation and land processes at satellite-resolved scales. *Environmental Modelling & Software*, **67**, 149–159.

Rajesh, P. V., S. Pattnaik, U. C. Mohanty, D. Rai, H. Baisya, P. C. Pandey, 2017: Land urface sensitivity of monsoon depressions formed over Bay of Bengal using improved high-resolution land state. *Dyn. Of Atmosphere and Oceans,* **80,** 155-172.

Reichle, R. H., and R. D. Koster, 2004: Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.,* **31,** L19501, doi: 10.1029/2004GL020938.

Rodell, M., P.R. Houser, A.A. Berg, and J.S. Famiglietti, 2005: Evaluation of 10 Methods for Initializing a Land Surface Model. *J. Hydrometeor.,* **6,** 146–155, https://doi.org/10.1175/JHM414.1

Rodell, M., P.R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J.K. Entin, J.P. Walker, D. Lohmann, and D. Toll, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.,* **85(3),** 381-394, 2004.

Rondinelli, W. J., B. K. Hornbuckle, J. C. Patton, M. H. Cosh, V. A. Walker, and B. D. Carr, 2015: Different Rates of Soil Drying after Rainfall Are Observed by the SMOS Satellite and the South Fork in situ Soil Moisture Network. *J. Hydrometeorol.,* **16,** 889-903.

Santanello, J.A., P.A. Dirmeyer, C.R. Ferguson, K.L. Findell, A.B. Tawfik, A. Berg, M. Ek, P. Gentine, B.P. Guillod, C. van Heerwaarden, J. Roundy, and V. Wulfmeyer, 2018: Land–Atmosphere Interactions: The LoCo Perspective. *Bull. Amer. Meteor. Soc.,* **99**, 1253–1272, https://doi.org/10.1175/BAMS-D-17-0001.1

———, S. V. Kumar, C. D. Peters-Lidard, and P. M. Lawston, 2016: Impact of Soil Moisture Assimilation on Land Surface Model Spinup and Coupled Land-Atmosphere Prediction. *J. Hydrometeorol.,* **17,** 517-540.

———, J. Roundy, and P. A. Dirmeyer, 2015: Quantifying the Land–Atmosphere Coupling Behavior in Modern Reanalysis Products over the U.S. Southern Great Plains. *J. Climate*, **28(14),** 5813-5829.

———, S. V. Kumar, C. D. Peters-Lidard, K. Harrison, and S. Zhou, 2013a: Impact of Land Model Calibration on Coupled Land-Atmosphere Prediction. *J. Hydromet*., **14,** 1373-1400.

———, C. D. Peters-Lidard, A. Kennedy, and S. V. 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,** 3–24.

———, M. Friedl, and M. Ek. 2007: Convective Planetary Boundary Layer Interactions with the Land Surface at Diurnal Time Scales: Diagnostics and Feedbacks. *J Hydrometeorol*. **8,** 1082-1097.

———, M. Friedl, and W. Kustas. 2005: Empirical Investigation of Convective Planetary Boundary Layer Evolution and its Relationship with the Land Surface. *J Appl Meteorol*., **44,** 917-932.

Schneeberger, K., M. Schwank, C. Stamm, P. de Rosnay, C. Matzler, and H. Fluhler, 2004: Topsoil structure influencing soil water retrieval by microwave radiometery. *Vadose Zone,* **3,** 1169-1179.

Seneviratne, S. I., and Coauthors, 2013: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. *Geophys Res Lett*, **40,** 5212-5217.

Shellito, P. J., and E. E. Small, 2017: Controls on surface soil drying rates observed by SMAP and simulated by the Noah land surface model. *Hydrol., Earth Syst. Sci.,* **22,** 1649-1663.

———, and co-authors, 2016: SMAP soil moisture drying more rapid than observed in situ following rainfall events. *Geophys. Res. Letters,* **43**.

Taylor, Christopher M., de Jeu, Richard A. M., Guichard, Francoise, Harris, Phil P. and Dorigo, Wouter A., 2012: Afternoon rain more likely over drier soils. *Nature,* **489,** 423-426. 10.1038/nature11377.

Weckwerth, T.M., K. Weber, D.D. Turner, and S.M. Spuler, 2016: Validation of a new water vapor micropulse differential absorption lidar (DIAL). *J. Atmos. Oceanic Technol*., **33,** 2353-2372, DOI:10.1175/JTECH-D-16-0119.1.

Welty, J. and X. Zeng, 2018: Does soil moisture affect warm season precipitation over the Southern Great Plains? Geophysical Research Letters. https://doi.org/10.1029/2018GL078598

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.

Xiang, T., E. R. Vivoni, and D. J. Gochis, 2017: Influence of initial soil moisture and vegetation conditions on monsoon precipitation events in northwest Mexico. *Atmosfera,* **31,** 25-45.

Zaitchik, B.F., J.A. Santanello, S.V. Kumar, and C.D. Peters-Lidard, 2013: Representation of Soil Moisture Feedbacks during Drought in NASA Unified WRF (NU-WRF). *J. Hydrometeor.,* **14,** 360–367, https://doi.org/10.1175/JHM-D-12-069.1

LIST OF TABLES

LIST OF FIGURES

Figure 4: Scatterplots of LIS vs. STAMP and SMAP vs. STAMP near-surface (0-5cm) soil moisture from Fig. 3 during JJA 2016 at the ARM-SGP a-b) E33, c-d) E38, and e-f) E31 sites.

Figure 5: Scatterplots of hourly mean diurnal cycle sensible and latent heat fluxes over JJA 2016 from the LIS (averaged across all three LIS simulations) vs. ECOR measurements at the ARM-SGP a) E33, b) E38, c) E31 sites, d) all 6 sites, and e) the available energy (sensible + latent heat flux) at all 6 sites.

Figure 6: Near-surface soil moisture from SMAP, the three LIS permutations, and 'off the shelf' products from NLDAS-2, GFS, and NARR valid at 12Z on a) 9 June, b) 11 July, and c) 28 August 2015. Stars indicate the locations of the ARM CF (Lamont, OK) and Ellis, KS sites.

Figure 7: Spatial PDFs of near-surface soil moisture at 12Z on a) 9 June, b) 11 July, and c) 28 August 2015 from the suite of products and across the full domains shown in Fig. 6.

Figure 8: Difference in a) soil moisture ICs (LIS-GDAS – NARR) at 12 UTC, and corresponding differences at 19 UTC in NU-WRF simulated b) sensible heat flux, c) latent heat flux, d) PBL height, e) 2-meter temperature, and f) 2-meter humidity.

Figure 9: Mixing diagrams and associated Bowen and entrainment ratios (as in Santanello et al. 2009) at a) the ARM-SGP CF and b) Ellis, KS sites derived from each of the NU-WRF simulations valid from 12Z on 11 July to 00Z on 12 July 2015.

Figure 10: Mean daytime evaporative fraction vs. PBL height derived from the simulations in Fig. 9 at the a) ARM-SGP CF and b) Ellis, KS sites.

Figure 11: Diurnal cycle of PBL height at for each of the simulations in Fig. 9 and observed at the a) ARM-SGP CF and b) Ellis, KS sites.

Figure 12: Initial near-surface soil moisture values at each of the 4 Noah LSM soil layers for each of the simulations in Fig. 9 at the a) ARM-SGP CF and b) Ellis, KS sites.

Figure 13: Mixing diagrams and evaporative fraction (-) vs. PBL height (m) analyses for the NU-WRF simulations with GFS and NARR initialization on 11 July 2015 at the ab) ARM-SGP CF, cd) Ellis, KS, ef) (36.0, -100), and gh) (39.0, -97.0) sites.

Figure 14: Time series of 2-meter temperature (T2) and specific humidity (Q2) RMSE and bias statistics from the suite of NU-WRF simulations versus observations at the 153 pairs of MET sites across the SGP domain.

Figure 15: **a)** Difference in GFS – NARR soil moisture ICs, along with bias statistics for temperature and humidity for the NARR and GFS simulations valid at the initial (12 UTC) and evening (00 UTC) times on 11 July 2015.