3.3 SMOS SOIL MOISTURE DATA ASSIMILATION IN THE NASA LAND INFORMATION SYSTEM: IMPACT ON LSM INITIALIZATION AND NWP FORECASTS

Clay B. Blankenship, USRA, Huntsville, Alabama*
Jonathan L. Case, ENSCO, Inc., Huntsville, Alabama
Bradley T. Zavodsky, NASA-MSFC, Huntsville, Alabama

1 INTRODUCTION

Land surface models are important components of numerical weather prediction (NWP) models, partitioning incoming energy into latent and sensitive heat fluxes that affect boundary layer growth and destabilization. During warm-season months, diurnal heating and convective initiation depend strongly on evapotranspiration and available boundary layer moisture, which are substantially affected by soil moisture content. Therefore, to properly simulate warm-season processes in NWP models, an accurate initialization of the land surface state is important for depicting the exchange of heat and moisture between the surface and boundary layer. The ultimate goal of this study is to improve numerical weather forecasts by providing more accurate soil moisture boundary conditions from a land surface model. Toward that end, we assimilate soil moisture retrievals from the Soil Moisture and Ocean Salinity (SMOS) radiometer into the Noah Land Surface Model via an Ensemble Kalman Filter.

2. Models and Data

2.1 LIS

The NASA Land Information System (LIS, Kumar et al. 2006; Peters-Lidard et al. 2007) is a modeling framework for running land surface models and conducting land surface data assimilation (Kumar et al. 2008; Kumar et al. 2009). To facilitate intercomparisons, users may select land surface models, forcing data sources, landcover and soil type data sources, and many other parameters. The Short-term Prediction Research and Transition (SPoRT) Center at NASA/MSFC uses LIS to produce real-time soil moisture products for situational awareness and local NWP over the southeastern CONUS and East Africa (Case and White 2014; Case et al. 2014). These products are shared with Weather Forecast Offices and other users. Model output can be used to monitor and/or predict several phenomena including drought, fire, extreme heat, flooding, convective initiation, and water-borne diseases.

2.2 SMOS and SMAP

The SMOS satellite, launched in 2009 by the European Space Agency, carries a 1.4-GHz synthetic aperture radiometer which can be used to retrieve soil moisture with a resolution of 35-50 km and a volumetric accuracy of 4%. Due to the lower-frequency observation channel, SMOS retrievals are more sensitive to soil moisture and can be made in more highly vegetated areas than previous instruments such as AMSR-E.

NASA launched the Soil Moisture Active/Passive (SMAP) mission in January 2015 (Entekhabi et al. 2010). The SMAP satellite carries a 1.41 GHz radiometer which has similar resolution and accuracy to SMOS (although it uses a conventional dish rather than a synthetic aperture) and a synthetic aperture radar operating at 1.2 GHz. The radar can be used to retrieve surface soil moisture with a horizontal resolution of ~3 km. A combined active/passive product with intermediate resolution at 9 km will also be available (Das et al. 2011). NASA SPoRT is part of the SMAP Early Adopters team (Brown et al. 2013), and plans to assimilate the blended SMAP product when these data become available operationally.

3. Soil Moisture Data Assimilation

We have adapted LIS to assimilate Level 2 Soil Moisture User Data Products from the ESA into the Noah version 3.2 land surface model (Ek et al. 2003) using an Ensemble Kalman Filter (Blankenship et al. 2014). Kalman filtering is a data assimilation method that

---

**Corresponding author address:* Clay B. Blankenship, NASA-MSFC SPoRT Center, 320 Sparkman Dr., Huntsville, AL 35805; e-mail: clay.blankenship@nasa.gov.

a)

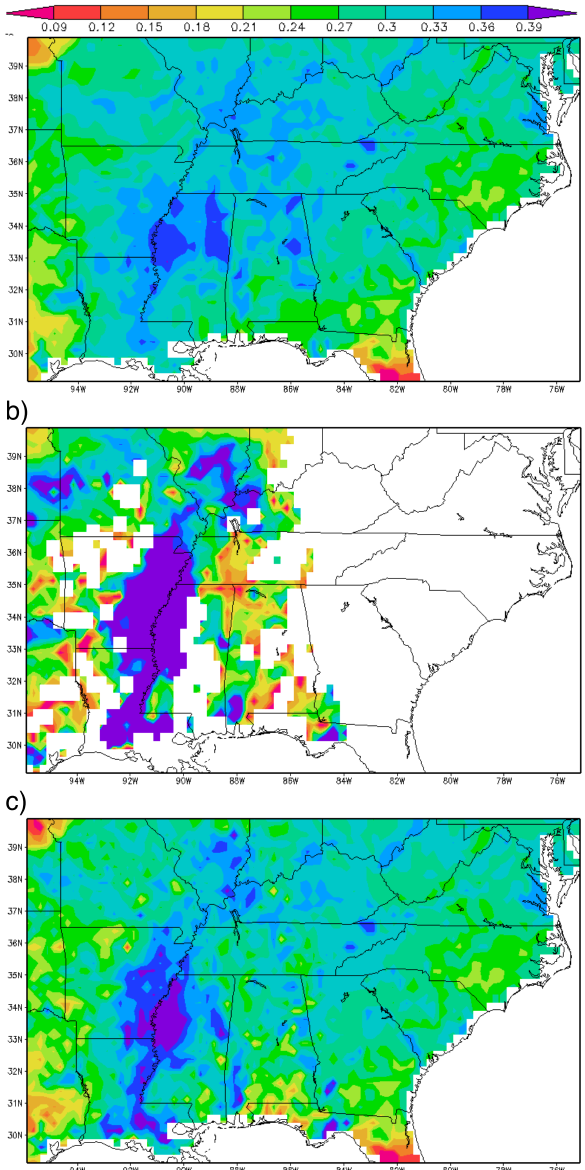


Figure 1. Soil moisture data assimilation case from 1 April 2013: a) Model background 0-10 cm soil moisture (cm^3/cm^3); b) SMOS soil moisture retrieval; and c) 0-10 cm soil moisture analysis.

combines a forecast (background) with observations (soil moisture retrievals, in this case) to generate an improved estimate of a model variable. A Kalman Filter calculates an optimal weighting between the background and the observation. The EnKF uses the spread of the ensemble to represent the forecast error covariance. Within LIS, data are screened for radio frequency interference, frozen soil, snowcover, falling precipitation, heavy vegetation, and other data quality flags.

An example of the assimilation process is shown in Figure 1. Panel a depicts the 0-10 cm soil moisture from the model at 0000 UTC 1 April 2013, prior to any assimilation. Panel b shows the observations (retrieved SMOS soil moisture) valid at the same time. In this particular case, there is a very prominent soil moisture feature in the retrievals in the lower Mississippi Valley, likely due to rice irrigation. The feature coincides closely with previously identified irrigated and rice-growing areas (Figure 2). Since irrigation is not present in the precipitation forcing used to drive LIS-Noah, the modeled field lacks this feature. Panel c depicts the analysis, which combines the model background and the observations, and which now shows the moist soils within the area under irrigation.

a)

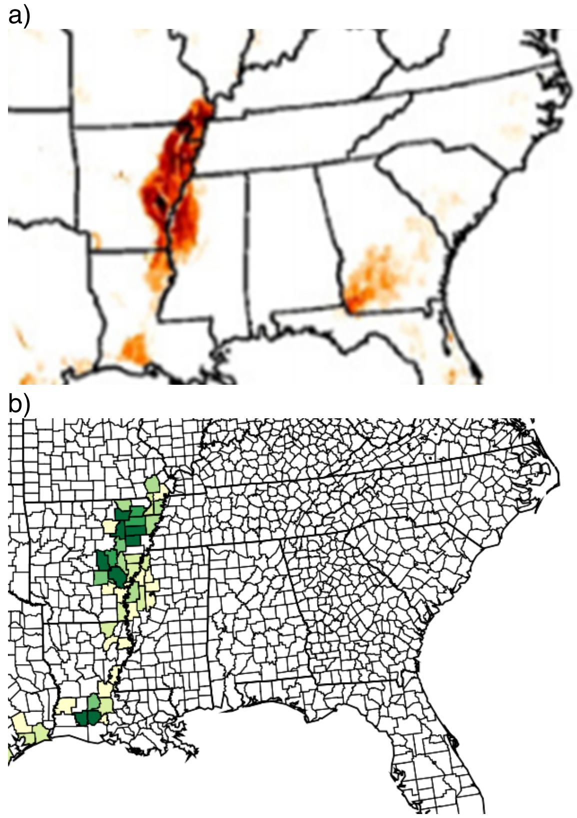
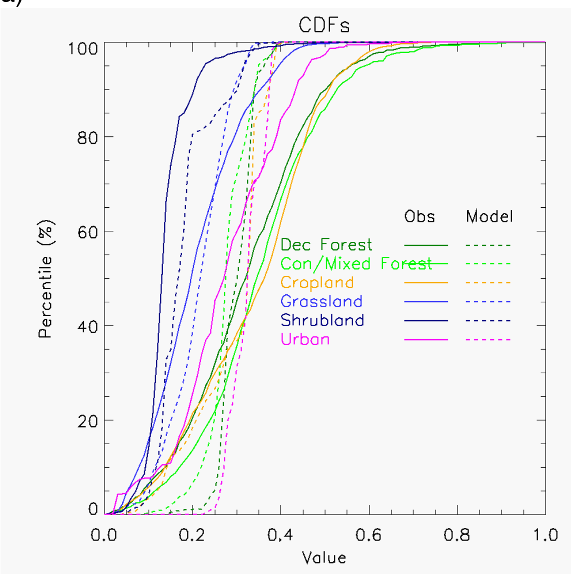


Figure 2. a) University of Frankfurt/FAO map of irrigated areas (from Ozdogan and Gutman 2001), and b) 2013 USDA map of rice production by county (National Agricultural Statistics Service 2015).

3.1 Bias Correction

We apply a bias correction to the SMOS retrievals using a Cumulative Distribution Function (CDF)-Matching technique (Reichle and Koster 2004), with a separate correction curve calculated for several broad land cover categories (Blankenship and Crosson 2011).

a)



b)

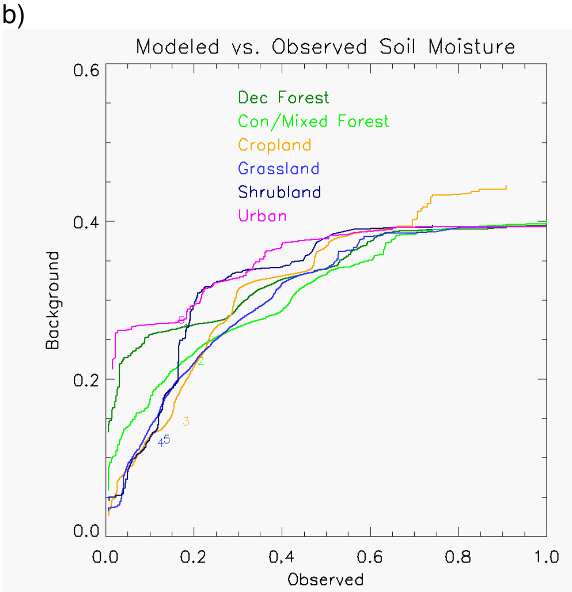


Figure 3. a) Cumulative Distribution Functions (CDFs) of SMOS retrieved surface soil moisture (solid curves) and modeled 0-10 cm soil moisture (dashed curves) for 6 landcover types, for a southeastern US domain. b) Correction curves derived from the CDFs.

Figure 3a shows the CDFs for soil moisture for both model (control run with no data assimilation; dashed curves) and observations (uncorrected satellite retrievals, solid curves) for six land cover types, generated from a model run and observations from March 1 to October 1 of 2013 over the experiment domain. The resulting correction curves are shown in Figure 3b. These curves are used to adjust an observation to an equivalent model value. The intent is to preserve the observations' information content regarding spatial patterns and relative values while not changing the model's soil moisture climatology. The observations are drier on average so the mean correction is a moistening of the signal. The observations also have a larger dynamic range, with values up to 1.0 volumetric soil moisture being allowed. This was an intentional choice by ESA to preserve the signal in flooded areas (Arnaud Mialon, personal communication, 10 April 2014).

3.2 Model Configuration

Three separate ensemble runs of the Noah 3.2 LSM were performed within LIS for the period of 1 Feb. 2011 to 1 Jun. 2011, on a 3-km domain in the southeastern United States. These simulations were all initialized with a one-year spinup from 1 Jan. 2010, followed by a one-month ensemble spinup (applying perturbations) during January 2011. The ensembles each consisted of 32 members generated by perturbing 4 state variables (4 layers of soil moisture), and 1 observation variable (SMOS soil moisture). The three runs were 1) OPL: an open loop control run, with no assimilation, 2) DA: an uncorrected SMOS data assimilation run, and 3) DA+BC: a SMOS assimilation run with bias correction. This setup allows us to separately assess the results of the assimilation of the uncorrected observations and of the bias correction. The SMOS satellite is a polar orbiter, so there are at most 2 passes per day available for assimilation at mid-latitudes (sometimes fewer due to gaps between swaths). For this domain, the assimilation took place at the 0000 and 1200 UTC timesteps.

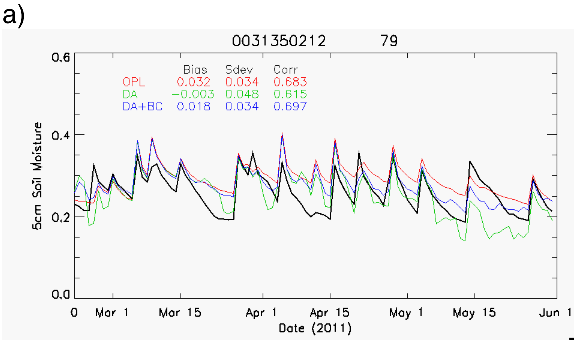
For all cases, the Noah model was run with a 30-minute timestep and configured with the IGBP MODIS landcover database (Hansen et al. 2000) and the STATSGO soil database

(Miller and White 1998). It was forced with temperature, humidity, winds, and incident radiation at the surface from the Global Data Assimilation System (GDAS; Wu et al. 2002; NCEP EMC 2004) and precipitation from the North American Land Data Assimilation Phase 2 (NLDAS2; Xia et al. 2012). Green vegetation fraction was derived from daily MODIS data (Case et al. 2014).

3.3 Results

Modeled soil moisture was validated against in situ observations from the North American Soil Moisture Database obtained from Texas A&M University (Ford and Quiring 2012). 112 stations within the domain were chosen, after screening out stations with limited variability and excessive periods of saturation in the in situ data. Time series from two of these stations are shown in Figure 4, along with 0-10 cm layer soil moisture at the same location from the three model runs. Results from the New Hope station in north Alabama (panel a) reveal that the DA run better captures the full dynamic range of drying compared to the OPL run, with

a)



b)

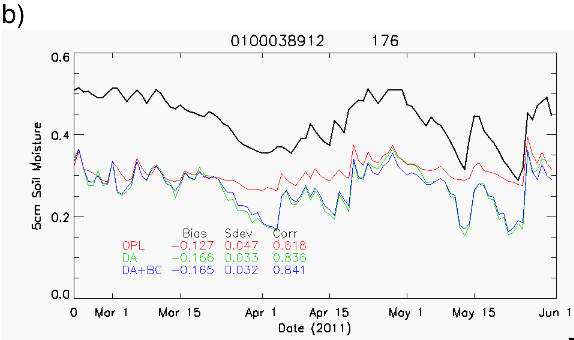


Figure 4. Time series of in 5 cm in situ soil moisture (black) and modeled 0-10 cm soil moisture at same point from 1 Feb 2011 to 1 Jun 2011. a) New Hope, AL. b) Champaign, IL.

Table 1. Statistics from 112 stations in the southeastern US, verified against the North American Soil Moisture Database for the period 1 Feb 2011 to 1 Jun 2011, for three model runs: open loop, SMOS data assimilation (DA), and SMOS data assimilation with bias correction (DA+BC).

Experiment	Open Loop	DA	DA+BC
Bias	-2.2%	-3.3%	-2.8%
Err. Std. Dev.	9.0%	9.2%	9.0%
Correlation	0.59	0.62	0.63

the DA+BC run in between. The second case from Champaign, IL (panel b), has a large initial bias, perhaps due to inconsistent soil properties between the site and the model grid cell. In this case, it appears that the two data assimilation runs better capture the dry periods in the second half of the run. For this station, there is a large improvement in correlation coefficient (Pearson r) from 0.62 for OPL to 0.84 for both data assimilation runs. Aggregate statistics from all 112 stations (Table 1) indicate a small increase in bias for the DA only run, which is lessened in the DA+BC run, and a small improvement in average correlation from 0.59 for OPL to 0.62 for DA and 0.63 for DA+BC.

4. NWP INITIALIZATION EXPERIMENT

4.1 Model Configuration

The ultimate goal of this research is to improve numerical weather forecasts by initializing an NWP model with more accurate surface soil moisture fields. We used the output of the three LIS runs—open loop (OPL), data assimilation (DA), and data assimilation with bias correction (DA+BC) LIS—to initialize forecasts from the Weather Research and Forecasting (WRF) NWP model, beginning on 1 Jun 2011, for the same 3-km southeastern US domain. Initial conditions for the atmosphere were taken from the North American Model (NAM) operational run. This area was chosen, in part, because of the large soil moisture anomalies due to rice irrigation discussed in section 3. It is hypothesized that the relatively large changes in soil moisture over the lower Mississippi Valley between the two runs will have significant impacts on evapotranspiration, diurnal heating, and convective initiation.

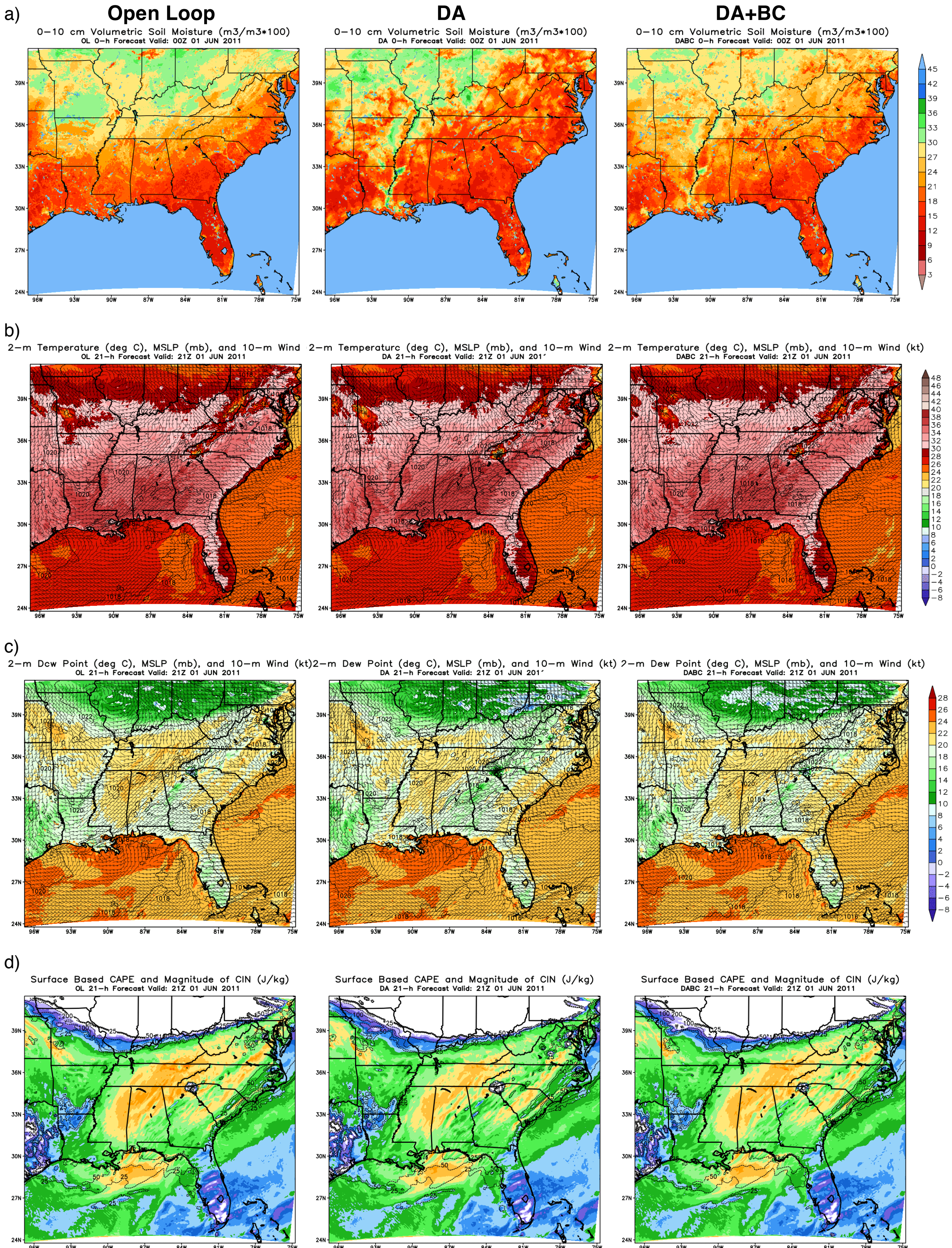


Figure 5. a)Initial 0-10 cm soil moisture, and 48-hr forecasts of b) 2-m temperature, c) 2-m dewpoint, and d) surface-based CAPE for OPL (left), DA (center), and DA+BC (right) runs.

Figure 5a shows the initial 0-10 cm volumetric soil moisture for all three model runs. The DA run is noticeably drier, with the exception of an area of increased moisture along the lower Mississippi River. By design, the DA+BC run closely matches the OPL run in a mean sense, but it retains the enhanced moisture near the Mississippi River, and overall has a weaker north-south moisture gradient compared to the OPL run.

4.2 Results

Results from a 48-h forecast initialized on 1 Jun 2011 are shown in Figure 5b-d for three different fields. The DA run, with a drier initial surface, has lower 2-m dewpoints and increased 2-m temperatures, as expected. These changes are mitigated in the DA+BC run, but the geographic patterns are still changed. The final field shown is the Convective Available Potential Energy (CAPE). There are competing effects from the dewpoint and temperature changes, but the moisture effect tends to dominate, with the drier DA run having reduced CAPE overall. However, the DA run does exhibit small cells of enhanced CAPE. While we cannot yet say if these patterns are more accurate, it does illustrate that the initial surface moisture conditions can impact the evolution of the model, including the development of convection (not shown). We plan to evaluate both the sensitivity and the accuracy of the weather forecasts, validating forecasts against station and upper-air observations over a longer time period, and by investigation of selected case studies.

5. SUMMARY AND FUTURE WORK

The NASA SPoRT Center has implemented the assimilation of soil moisture retrievals from the SMOS satellite within the Land Information System. A bias correction using a CDF-matching technique, stratified by vegetation class, has also been developed. Preliminary validation against the North American Soil Moisture Database indicates an improved correlation of model time series of soil moisture for both the uncorrected and corrected data assimilation runs. LIS output from data assimilation runs was used to initialize 48-hr forecasts in WRF. The changes in initial surface conditions led to model changes in 2-m temperature and dewpoint as well as CAPE. We plan to investigate the WRF forecasts to assess both the sensitivity to using SMOS observations in this manner, and the accuracy of the different experiments (no assimilation and assimilation with and without bias correction). We also plan to assimilate SMAP data when they become available, and to assess the impact of SMAP on NWP forecasts. After validation, the data assimilation will be implemented in SPoRT's near-real-time LIS modeling efforts.

6. ACKNOWLEDGEMENTS

This work is supported by Tsengdar Lee of the NASA Science Mission Directorate through SPoRT. We would also like to thank Susan Moran, Vanessa Escobar, and Molly Brown for including us as SMAP Early Adopters.

7. REFERENCES

Blankenship, C. B. and W. Crosson, 2011: Bias correction based on regime-dependent Cumulative Distribution Functions for soil moisture data assimilation in a Land Surface Model. *25th Conf. on Hydrology,* Seattle, WA, Amer. Meteor. Soc., P48. [Available online at https://ams.confex.com/ams/91Annual/webprogram/Paper185532.html]

Blankenship, C. B., B. T. Zavodsky, and J. L. Case, 2014: Assimilation of SMOS soil moisture retrievals in the Land Information System. Preprints, *28th Conf. on Hydrology,* Atlanta, GA, Amer. Meteor. Soc., P53. [Available online at https://ams.confex.com/ams/94Annual/webprogram/Manuscript/Paper233646/AMS-2014-SMOS-ext-abs.pdf]

Brown, Molly E., Vanessa Escobar, Susan Moran, Dara Entekhabi, Peggy E. O'Neill, Eni G. Njoku, Brad Doorn, Jared K. Entin, 2013: NASA's Soil Moisture Active Passive (SMAP) Mission and Opportunities for Applications Users. *Bull. Amer. Meteor. Soc.*, **94**, 1125–1128.

Case, J. L., F. J. LaFontaine, J. R. Bell, G. J. Jedlovec, S. V. Kumar, and C. D. Peters-Lidard, 2014: A real-time MODIS vegetation product for land surface and numerical weather prediction models. *IEEE Trans. Geosci. Remote Sens.,* **52(3),** 1772-1786.

Case, J. L., and K. D. White, 2014: Expansion of the real-time SPoRT-Land Information System for NOAA/National Weather Service situational awareness and local modeling applications. Preprints, *26th Conf. on Weather Analysis and Forecasting / 22nd Conf. on Numerical Weather Prediction,* Atlanta, GA, Amer. Meteor. Soc., P162. [Available online at https://ams.confex.com/ams/94Annual/webprogram/Paper235310.html]

Case, J. L., J. Mungai, V. Sakwa, E. Kabuchanga, B. T. Zavodsky, and A. S. Limaye, 2014: Toward improved land surface initialization in support of regional WRF forecasts at the Kenya Meteorological Service. Preprints, *26th Conf. on Weather Analysis and Forecasting / 22nd Conf. on Numerical Weather Prediction,* Atlanta, GA, Amer. Meteor. Soc., J9.4. [Available online at https://ams.confex.com/ams/94Annual/webprogram/Paper235263.html]

Das, N. N., D. Entekhabi, and E. G. Njoku, 2011: An algorithm for merging SMAP radiometer and radar data for high-resolution soil-moisture retrieval. *IEEE Trans. Geosci. Remote Sens.,* **49 (5),** 1504-1512.

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 (D22),** 8851.

Entekhabi, D. and Coauthors, 2010: "The Soil Moisture Active Passive (SMAP) mission," *Proc. IEEE*, **98 (5),** 704-716.

European Space Agency, 2002: *Mission Objectives and Scientific Requirements of the Soil Moisture and Ocean Salinity (SMOS) Mission,* Version 5. http://esamultimedia.esa.int/docs/SMOS_MRD_V5.pdf

Ford, T. W., and S. M. Quiring (2012) The North American Soil Moisture Database. American Geophysical Union Fall Meeting, San Francisco, California, December 3 – 7.

Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E.F. Wood and J. Sheffield, 2006: Land Information System - An interoperable framework for high resolution land surface modeling. *Environmental Modelling and Software*, **21,** 1402-1415.

______, R. H. Reichle, R. D. Koster, W. T. Crow, and C. D. Peters-Lidard, 2009: Role of subsurface physics in the assimilation of surface soil moisture observations. *J. Hydrometeor.,* **10,** 1534-1547.

______, ______, 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. *Adv. Water Res.,* **31,** 1419-1432.

Miller, D. A. and R. A. White, 1998: A Conterminous United States multi-layer soil characteristics data set for regional climate and hydrology modeling. *Earth Interactions*, 2. [Available on-line at http://EarthInteractions.org].

National Agricultural Statistics Service, cited 2015: Rice County Maps, Production by County [Available online at http://www.nass.usda.gov/Charts_and_Maps/Crops_County/pdf/AR-PR13-RGBChor.pdf]

NCEP EMC, 2004: SSI Analysis System 2004. NOAA / NCEP / Environmental Modeling Center Office Note 443, 11 pp., April, 2004. [Available online at http://www.emc.ncep.noaa.gov/officenotes/newernotes/on443.pdf]

Ozdogan, M. and G. Gutman, 2008: A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US, *Remote Sensing of Environment*, **112**: 3520-3537.

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

Reichle, R. H., and R. Koster, 2004: Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.,* L19501

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.,* **130,** 2905-2916.

Xia, Y., and Coauthors, “Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products,” J. Geophys. Res., vol. 117, no. D3, pp. 27, February, 2012.