The Efficiency of Assimilating Satellite Soil Moisture Retrievals in a Land Data Assimilation System Using Different Rainfall Error Models

VIVIANA MAGGIONI

Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut

ROLF H. REICHLE

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

EMMANOUIL N. ANAGNOSTOU

Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut

(Manuscript received 9 July 2012, in final form 12 October 2012)

ABSTRACT

The efficiency of assimilating near-surface soil moisture retrievals from Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) observations in a Land Data Assimilation System (LDAS) is assessed using satellite rainfall forcing and two different satellite rainfall error models: a complex, multidimensional satellite rainfall error model (SREM2D) and the simpler (control) model (CTRL) used in the NASA Goddard Earth Observing System Model, version 5 LDAS. For the study domain of Oklahoma, LDAS soil moisture estimates improve over the satellite retrievals and the open-loop (no assimilation) land surface model estimates, exhibiting higher daily anomaly correlation coefficients (e.g., 0.36 in the open loop, 0.38 in the AMSR-E, and 0.50 in LDAS for surface soil moisture). The LDAS soil moisture estimates also match the performance of a benchmark model simulation forced with high-quality radar precipitation. Compared to using the CTRL rainfall error model in LDAS, using the more complex SREM2D exhibits only slight improvements in soil moisture estimates.

1. Introduction

Soil moisture is a key variable of the hydrological cycle and governs water and energy exchanges at the land–atmosphere interface. Ground-based soil moisture measurements are sparse, but satellite observations of microwave brightness temperature and backscatter can monitor soil moisture at large scales. However, satellite measurements are sensitive only to the top few centimeters of the soil column and are subject to errors due to sensor limitations and the parameterization of the retrieval algorithm. Land surface models, on the other hand, can provide continuous and spatially distributed soil moisture estimates, but model estimates are susceptible to errors in the model forcing, structure, and parameters. The observational and modeling limitations can be partially overcome through a Land Data Assimilation System (LDAS) that combines soil moisture observations with land model estimates to maximize spatial and temporal coverage, consistency, and accuracy (Reichle 2008). This is particularly relevant in the context of current and planned satellite missions: the Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al. 2010), the Soil Moisture Active Passive (SMAP) mission (Entekhabi et al. 2010), and the Global Precipitation Measurement (GPM) mission (Hou et al. 2008).

A key issue of data assimilation is that observational and modeling uncertainties are poorly known, and incorrect assumptions about these errors may compromise LDAS efficiency (Crow and Van Loon 2006). It is thus crucial to investigate the impact of the error characterization on the assimilation of soil moisture observations, in particular because LDASs often use very simplistic error models. As rainfall is the dominant meteorological forcing input to the land surface model for soil moisture

Corresponding author address: Viviana Maggioni, ESSIC/University of Maryland, 5825 University Research Court, Suite 4001, College Park, MD 20740-3823.
E-mail: viviana@umd.edu

DOI: 10.1175/JHM-D-12-0105.1

© 2013 American Meteorological Society

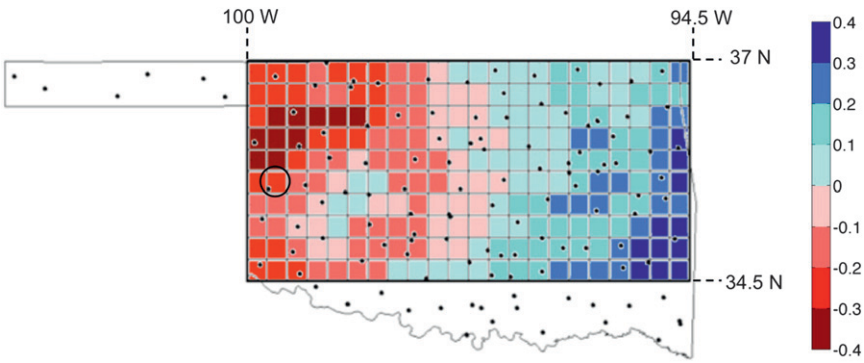

FIG. 1. Map of the rainfall climatology parameter (dimensionless), computed at each grid cell with respect to the 3-yr domain average Stage IV rainfall (Maggioni et al. 2012), overlaid by the 25-km spatial interpolation grid of the study domain and the locations of Mesonet stations. The black circle highlights the location from which soil moisture time series are shown in Fig. 3.

estimation, a more comprehensive characterization of rainfall uncertainty may improve soil moisture estimates. Since soil moisture temporally integrates antecedent precipitation and is subject to lower and upper limits, the variability of errors in soil moisture is typically smaller than that of errors in precipitation. This error variance relationship is not linear and depends on the error properties of the rainfall fields (Hossain and Anagnostou 2005). Maggioni et al. (2011) showed that the use of a complex error model to characterize the spatial variability of rainfall errors could better capture soil moisture error properties. Furthermore, in a synthetic numerical assimilation experiment, Maggioni et al. (2012) demonstrated that using the more elaborate rainfall error model may slightly improve surface and root zone soil moisture estimates obtained from assimilating soil moisture retrievals. This study expands the synthetic experiment of Maggioni et al. (2012) by evaluating the assimilation of near-surface soil moisture retrievals from the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E). In the following section we discuss the domain and data used in this study. The models and methodology used to generate and evaluate the assimilation estimates are described in section 3. Results and conclusions are discussed in sections 4 and 5.

2. Study region and data

The study domain comprises a 250 km × 550 km modeling grid with 25^2 km^2 resolution covering Oklahoma, United States (Fig. 1). The study period is three years, 1 January 2004 to 31 December 2006. Two rainfall products, a radar and a satellite-based rainfall dataset, are used to force the land surface model along with supplemental surface meteorological forcing data from the Global Land Data Assimilation System (GLDAS) project (Rodell et al. 2004). The radar rainfall dataset is the most accurate available surface rainfall product and is used here for a benchmark model simulation. It is extracted from the offline Stage IV National Weather Service Weather Surveillance Radar-1988 Doppler (hereinafter Stage IV) precipitation estimation algorithm that involves rain gauge adjustment and postprocessing of radar data (Fulton 1998). A rainfall climatological parameter derived from radar rainfall (Maggioni et al. 2012) is shown in Fig. 1. The rainfall spatial pattern exhibits a west to east gradient, which is used in this study to define wet and dry climatological regimes. The satellite rainfall is provided by the NOAA Climate Prediction Center morphing (CMORPH) product (Joyce et al. 2004), which combines passive microwave retrievals from low earth-orbiting satellites and geostationary satellite infrared data. In this study, CMORPH rainfall fields are adjusted to the radar rainfall mean climatology to meet the LDAS assumption of unbiased forcing. Climatological bias adjustment based on long-term ground observations is now used by a number of satellite retrievals (Huffman et al. 2007; Xie and Xiong 2011).

The study domain is covered by the Oklahoma Mesoscale Network (hereinafter Mesonet) (Brock et al. 1995), which is a dense network of meteorological stations (see Fig. 1) that provide soil moisture measurements at 5-cm-, 25-cm-, 60-cm-, and 75-cm depths every 30 min. Mesonet measurements at the 60-cm and 75-cm depths are sparse and of insufficient quality during the study period; therefore, only the 5-cm and 25-cm data are used in this study to provide independent verification of the model and satellite soil moisture estimates. The satellite soil moisture estimates are from the AMSR-E Land Parameter Retrieval Model (LPRM) (Owe et al. 2008), which uses one dual-polarized channel for the retrieval of surface soil moisture and vegetation water content. Here we used

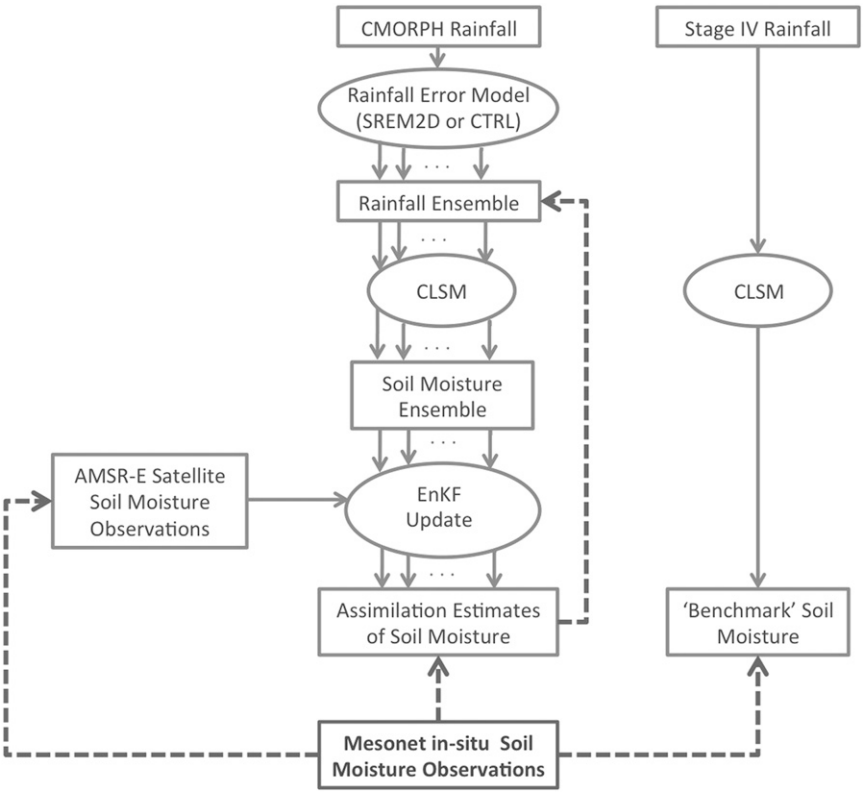

FIG. 2. Schematic of the experimental setup.

retrievals based on X-band (10.7 GHz) brightness temperatures from ascending and descending overpasses.

3. Methodology

The methodology, shown in Fig. 2, follows Maggioni et al. (2012) with the difference that, in this study, we used actual satellite soil moisture estimates. Namely, the NASA Goddard Earth Observing System Model, version 5 (GEOS-5) LDAS is used to correct soil moisture generated by the Catchment Land Surface Model (CLSM) (Koster et al. 2000) toward AMSR-E soil moisture retrievals based on the ensemble Kalman filter (EnKF) approach (Reichle et al. 2007). The AMSR-E retrievals were scaled prior to data assimilation so that their climatology matched that of CLSM surface soil moisture. The EnKF dynamically updates model error covariance information by producing an ensemble of model predictions. Each member of the ensemble experiences perturbations in the observed forcing fields (representing forcing errors) and randomly generated noise added to the prognostic variables (representing model errors).

To study the rainfall error complexity effect on data assimilation, precipitation forcing error is introduced using two different rainfall error models: the standard model currently used as part of the GEOS-5 LDAS (control model, henceforth CTRL) and the more comprehensive two-dimensional satellite rainfall error model (SREM2D) of Hossain and Anagnostou (2006). Both error models apply stochastic space–time formulations to describe the satellite-retrieval error structure. While CTRL assumes a perfect delineation of rainy and nonrainy areas by simply scaling the precipitation forcing with an ensemble of spatially distributed, mean-unity, lognormal, multiplicative perturbations, SREM2D explicitly models the joint probability distribution of successful delineation of rainy and nonrainy areas. The result is that SREM2D, unlike CTRL, would introduce rain in areas where the satellite does not detect rain (missed rain) or assign zero rain where the satellite retrieves rain (false alarms), while nonzero rain rates are perturbed as in CTRL. SREM2D and CTRL parameters were calibrated to generate replicates of CMORPH precipitation that stochastically reproduce the overall variability of the satellite rainfall errors across different spatial scales (Maggioni et al. 2011). Although in this study calibration was based on high-quality reference rainfall from the Stage IV product, application of the techniques at the global scale may introduce ambiguities in the calibration of the rainfall error model parameters,

which would cause uncertainty in the estimation of modeling error variance. In the absence of quality ground validation rainfall data, error model parameters could be estimated using the Tropical Rainfall Measuring Mission (TRMM) precipitation radar rainfall fields (Dinku and Anagnostou 2005).

CTRL and SREM2D are used independently with CMORPH rainfall estimates to generate precipitation ensembles. Four experiments are performed: each of the two satellite rainfall error model ensembles is used in LDAS without assimilation [open-loop (OL) runs: OL-CTRL and OL-SREM2D] and with the assimilation of AMSR-E soil moisture retrievals [assimilation runs (DA): DA-CTRL and DA-SREM2D] using a one-dimensional EnKF (Reichle et al. 2007). The retrieval (observation) error standard deviation is 0.08 $m^3\ m^{-3}$ (Liu et al. 2011). We use 24 ensemble members for each simulation. In addition, a fifth experiment, labeled "Stage IV," is performed by forcing CLSM with Stage IV radar rainfall to obtain (single member, unperturbed) benchmark model soil moisture estimates (without assimilating AMSR-E retrievals).

For each of the five experiments, the simulated surface (0–2 cm) and root zone (0–100 cm) soil moisture is compared against Mesonet ground measurements at 5-cm and 25-cm depth, respectively. The daily anomaly correlation coefficient (ACC) is used as a quantitative performance metric, where anomalies are defined as differences between the daily values and the monthly climatological average values of the 3-yr time series. For each experiment (OL or DA), the ACC was computed separately for each grid cell from the (ensemble mean) simulated soil moisture anomalies and the corresponding Mesonet data. Only time steps and model grid cells for which satellite observations were assimilated are included in the computation. For each ACC estimate a 95% confidence interval is calculated using a Fisher transformation (Hawkins 1989). Because soil moisture anomaly time series are autocorrelated, the number of degrees of freedom is smaller than the number of observations included in the ACC calculations. We account for this through the use of an effective sample size defined as $N_{\mathrm{eff}} = N(1 - r_x r_y)/(1 + r_x r_y)$, where N is the number of samples and r_x and r_y are the lag-one autocorrelations of the two time series (Dawdy and Matalas 1964).

4. Results and discussion

Figure 3 shows soil moisture anomaly time series for 1 March 2005 to 31 October 2005 for a single Mesonet station located at 99.73°W, 35.55°N (see Fig. 1) and the corresponding model grid cell. Modeled soil moisture anomalies are largely consistent with anomalies of ground observations and satellite retrievals. As expected, DA (using either CTRL or SREM2D) brings surface soil moisture model predictions closer to the satellite surface soil moisture observations, which in most cases leads to better agreement with ground data. For example, during June 2005, DA estimates of surface soil moisture anomalies are closer to the Mesonet anomalies than the OL predictions. For root zone soil moisture, model soil moisture anomaly time series are consistent with the Mesonet time series, with DA exhibiting better agreement than the OL time series (e.g., September 2005). Notably, DA often performs better than the benchmark Stage IV simulation at both soil moisture depths. No significant improvement is detected when SREM2D is adopted compared to the CTRL rainfall error model.

Next, we quantitatively verify for the entire domain and experiment period what is qualitatively presented in Fig. 3 for a single station and a single summer. Figure 4 shows the domain-average ACCs calculated against Mesonet data and computed from all model grid cells where both satellite and ground observations were available, along with 95% confidence intervals. Anomaly correlation coefficient values are shown to increase significantly when assimilating satellite near-surface soil moisture data. Specifically, for surface soil moisture the mean ACC increases from 0.36 in the OL simulations to 0.50 in the DA experiments for both rainfall error models (39% relative improvement due to DA). In the case of root zone soil moisture, the improvement obtained from DA is slightly lower. Specifically, ACC is 0.35 (0.34) in the OL-CTRL (OL-SREM2D) experiment, and it increases to a value of 0.44 for both DA experiments, for a relative ACC improvement of 26% (29%) for CTRL (SREM2D). The increases are significant at the 5% confidence level (nonoverlapping 95% confidence intervals). These results are consistent with the findings of Maggioni et al. (2012), who showed significantly improved estimates of soil moisture due to the assimilation of synthetic satellite soil moisture observations but only slight improvements due to the use within LDAS of the more complex rainfall error model.

The mean ACC of the AMSR-E retrievals is 0.38, which is slightly better than the OL performance and significantly lower than the mean ACC in DA for both rainfall error models (Fig. 4). This indicates the ability of DA to filter two poorer estimates into a superior soil moisture estimate. The performance of the DA estimates is comparable to what is obtained by forcing the model with the most accurate available rainfall product (i.e., Stage IV). For surface soil moisture, DA exhibits better ACC values (0.50) than those achieved using the benchmark simulation (0.45). For root zone soil moisture, the mean ACC for the Stage IV simulation is 0.44, which is equal to that of DA-CTRL and DA-SREM2D.

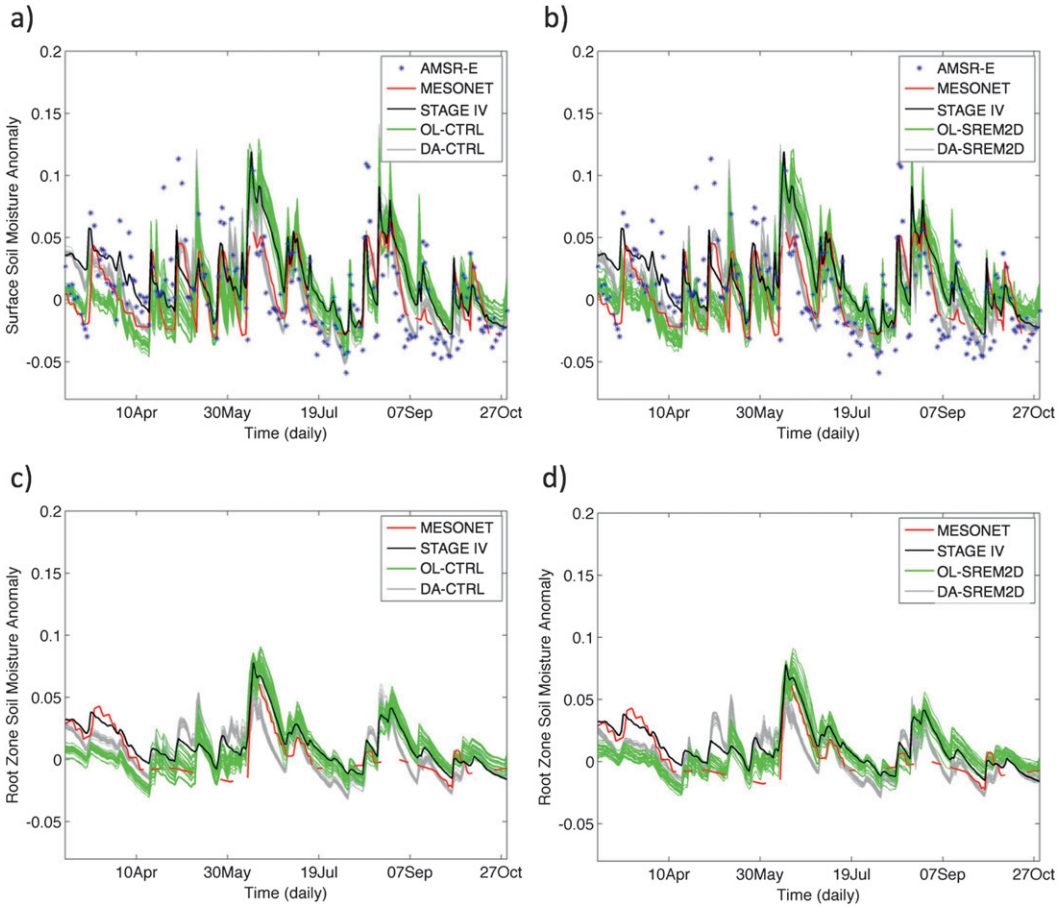

FIG. 3. (a),(b) Surface and (c),(d) root zone soil moisture anomalies ($m^3\ m^{-3}$) calculated for the Mesonet station located at 99.73°W, 35.55°N; the Stage IV benchmark simulation; and OL and DA results when the (a),(c) CTRL and (b),(d) SREM2D rainfall error models are used in LDAS.

Analogously to the study by Maggioni et al. (2012), we separated the domain into climatologically drier and wetter areas (negative and positive values shown in Fig. 1) for additional analysis. The corresponding performance metrics, also shown in Fig. 4, demonstrate that all simulations (benchmark, OL, and DA) perform better for the wetter conditions. In most cases, the difference between ACCs in the drier and wetter conditions is statistically significant at the 5% confidence level, as the 95% confidence intervals are distinct for almost all simulations.

For both surface and root zone soil moisture, the relative improvement in ACCs due to DA with respect to the OL simulation is larger in drier conditions. In particular, for surface soil moisture the relative improvement is equal to 44% (47%) under drier conditions and 38% (40%) under wetter conditions when CTRL (SREM2D) is considered. Analogously, for root zone soil moisture the ACC relative improvement equals 37% (40%) in drier regimes and 28% (32%) in wetter regimes for the CTRL (SREM2D) case. It is worth noting that in all cases SREM2D provides a slightly higher relative improvement than CTRL, consistent with the findings of the Maggioni et al. (2012) synthetic experiment. As satellite-based rainfall accumulations are prone to false positives, especially in arid and semiarid areas (due to below-cloud evaporation in dry atmospheres and potential effects of soil wetness on the detection of low rain rates by microwave techniques), it is encouraging that SREM2D adds slightly more marginal value in the drier climatological regime than in the wetter regime. Future studies should investigate the significance of these improvements.

Overall, DA improves the performance of the model in both wetter and drier regions, making it comparable to the benchmark simulation. When analyzing the relative improvement due to DA with respect to the benchmark Stage IV-forced simulation, for surface soil moisture the relative improvement is larger in the drier regime, which is encouraging as model estimates are generally poorer under this condition: 14% (16%) under drier

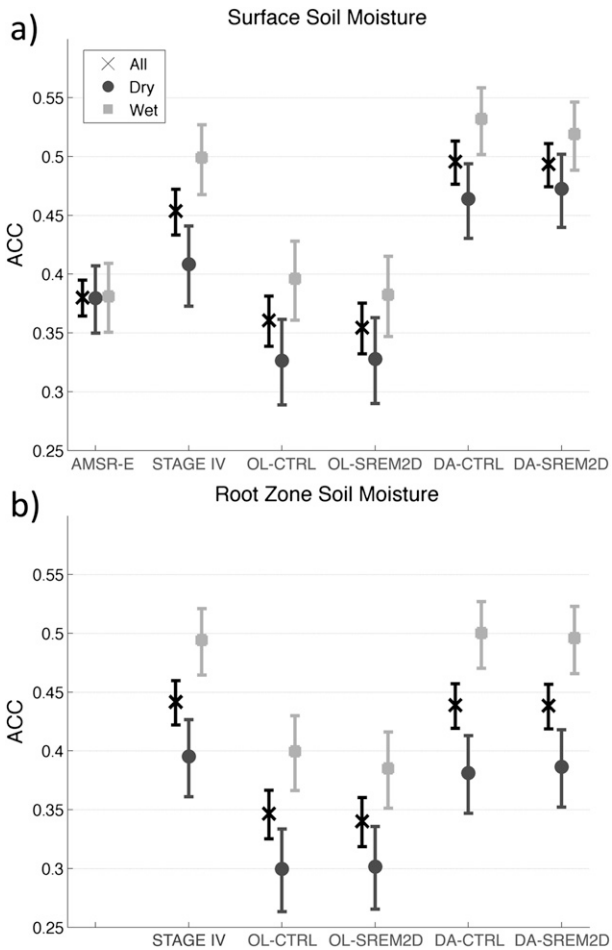

FIG. 4. Domain-average of daily ACCs (dimensionless) calculated against Mesonet station observations for (a) surface and (b) root zone soil moisture with 95% confidence intervals for all grid cells and binned by dry and wet conditions.

conditions and 9% (7%) under wetter conditions for DA-CTRL (DA-SREM2D). On the other hand, for root zone soil moisture we note slight improvement in the wetter regimes (3% for both rainfall error model), whereas in drier conditions DA exhibits worse performances than the benchmark run. SREM2D provided slightly better results than CTRL in this regime.

5. Summary and conclusions

This study investigated the efficiency of assimilating satellite near-surface soil moisture retrievals into CLSM with the GEOS-5 LDAS. Specifically, five experiments were considered: a benchmark simulation forced with Stage IV radar rainfall and four experiments obtained by perturbing satellite rainfall fields (i.e., CMORPH) with two rainfall error models of different complexity: with and without the assimilation of AMSR-E soil moisture retrievals. Surface and root zone soil moisture outputs from each experiment were compared against Mesonet ground measurements.

Results show that the assimilation of satellite soil moisture retrievals provides a significant improvement of surface and root zone soil moisture estimates, indicating the ability of the model update to propagate to deeper soil levels. The improvement due to assimilation is apparent also in comparison to the AMSR-E retrievals themselves; that is, starting from two poorer estimates of soil moisture, DA provides a superior estimate. We also note that soil moisture estimates from DA exhibited correlations higher than, or at least as high as, those estimated from the model forced with the most accurate rainfall (i.e., Stage IV). The use of a more complex rainfall error model leads to only marginally better soil moisture analyses, which suggests that the simpler rainfall error model may be adequate in soil moisture data assimilation. Nevertheless, the use of a more sophisticated error model, such as SREM2D, is suggested in future land data assimilation studies in order to provide a more realistic representation of the sources and nature of errors in precipitation retrievals. Furthermore, comparison of the two error modeling techniques across various hydroclimatic conditions is needed to provide a comprehensive understanding about the use of SREM2D in land data assimilation.

The results are encouraging toward the use of satellite retrievals (of both soil moisture and precipitation) in LDAS, especially because of the increasing availability of satellite soil moisture and rainfall observations from the Soil Moisture and Ocean Salinity (SMOS), Soil Moisture Active and Passive (SMAP), and Global Precipitation Measurement (GPM) missions. In summary, LDAS driven by satellite observations can greatly benefit soil moisture estimates in areas where ground-based precipitation observations are sparse and thereby support, for example, studies of land–atmosphere interactions or seasonal prediction.

Acknowledgments. V. Maggioni was supported by a NASA Earth System Science Graduate Fellowship. R. Reichle was supported by the NASA research program "The Science of Terra and Aqua" and the SMAP Science Definition Team. E. Anagnostou was supported by NASA Precipitation Science Team Grant NNX07AE31G. Computing was supported by the NASA High End Computing Program.

REFERENCES

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma mesonet: A technical overview. *J. Atmos. Oceanic Technol.,* **12,** 5–19.

Crow, W. T., and E. Van Loon, 2006: Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture. *J. Hydrometeor.,* **7,** 421–432.

Dawdy, D., and N. Matalas, 1964: Statistical and probability analysis of hydrologic data. Part III: Analysis of variance, covariance and time series. *Handbook of Applied Hydrology: A Compendium of Water-Resources Technology,* McGraw-Hill, 8.68–8.91.

Dinku, T., and E. N. Anagnostou, 2005: Regional differences in overland rainfall estimation from PR-calibrated TMI algorithm. *J. Appl. Meteor.,* **44,** 189–205.

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

Fulton, R. A., 1998: WSR-88D polar-to-HRAP mapping. National Weather Service Hydrologic Research Laboratory Tech. Memo., 33 pp.

Hawkins, D. L., 1989: Using U statistics to derive the asymptotic distribution of Fisher's *Z* statistic. *Amer. Stat.,* **43,** 235–237, doi:10.2307/2685369.

Hossain, F., and E. N. Anagnostou, 2005: Numerical investigation of the impact of uncertainties in satellite rainfall estimation and land surface model parameters on simulation of soil moisture. *Adv. Water Resour.,* **28,** 1336–1350, doi:10.1016/j.advwatres.2005.03.013.

——, and ——, 2006: A two-dimensional satellite rainfall error model. *IEEE Trans. Geosci. Remote Sens.,* **44,** 1511–1522.

Hou, A. Y., C. Kummerow, G. Skofronick-Jackson, and J. M. Shepherd, 2008: Global precipitation measurement. *Precipitation: Advances in Measurement, Estimation and Prediction,* Springer-Verlag, 131–169.

Huffman, G. J., and Coauthors, 2007: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.,* **8,** 38–55.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.,* **5,** 487–503.

Kerr, Y. H., and Coauthors, 2010: The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proc. IEEE,* **98,** 666–687, doi:10.1109/JPROC.2010.2043032.

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

Liu, Q., and Coauthors, 2011: The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. *J. Hydrometeor.,* **12,** 750–765.

Maggioni, V., R. H. Reichle, and E. N. Anagnostou, 2011: The effect of satellite rainfall error modeling on soil moisture prediction uncertainty. *J. Hydrometeor.,* **12,** 413–428.

——, ——, and ——, 2012: The impact of rainfall error characterization on the estimation of soil moisture fields in a land data assimilation system. *J. Hydrometeor.,* **13,** 1107–1118.

Owe, M., R. de Jeu, and T. Holmes, 2008: Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res.,* **113,** F01002, doi:10.1029/2007JF000769.

Reichle, R. H., 2008: Data assimilation methods in the earth sciences. *Adv. Water Resour.,* **31,** 1411–1418, doi:10.1016/j.advwatres.2008.01.001.

——, R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe, 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSRE) and the Scanning Multichannel Microwave Radiometer (SMMR). *J. Geophys. Res.,* **112,** D09108, doi:10.1029/2006JD008033.

Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.,* **85,** 381–394.

Xie, P., and A.-Y. Xiong, 2011: A conceptual model for constructing high-resolution gauge-satellite merged precipitation analyses. *J. Geophys. Res.,* **116,** D21106, doi:10.1029/2011JD016118.