Evaluating the Utility of Remotely Sensed Soil Moisture Retrievals for Operational Agricultural Drought Monitoring

John D. Bolten, *Member, IEEE*, Wade T. Crow, *Member, IEEE*, Xiwu Zhan, Thomas J. Jackson, *Fellow, IEEE*, and Curt A. Reynolds

***Abstract*—Soil moisture is a fundamental data source used by the United States Department of Agriculture (USDA) International Production Assessment Division (IPAD) to monitor crop growth stage and condition and subsequently, globally forecast agricultural yields. Currently, the USDA IPAD estimates surface and root-zone soil moisture using a two-layer modified Palmer soil moisture model forced by global precipitation and temperature measurements. However, this approach suffers from well-known errors arising from uncertainty in model forcing data and highly simplified model physics. Here, we attempt to correct for these errors by designing and applying an Ensemble Kalman filter (EnKF) data assimilation system to integrate surface soil moisture retrievals from the NASA Advanced Microwave Scanning Radiometer (AMSR-E) into the USDA modified Palmer soil moisture model. An assessment of soil moisture analysis products produced from this assimilation has been completed for a five-year (2002 to 2007) period over the North American continent between 23° N—50° N and 128° W—65° W. In particular, a data denial experimental approach is utilized to isolate the added utility of integrating remotely sensed soil moisture by comparing EnKF soil moisture results obtained using (relatively) low-quality precipitation products obtained from real-time satellite imagery to baseline Palmer model runs forced with higher quality rainfall. An analysis of root-zone anomalies for each model simulation suggests that the assimilation of AMSR-E surface soil moisture retrievals can add significant value to USDA root-zone predictions derived from real-time satellite precipitation products.**

***Index Terms*—Agriculture, data assimilation, remote sensing, soil moisture.**

Manuscript received November 17, 2008; revised April 20, 2009; accepted June 23, 2009. First published December 15, 2009; current version published February 24, 2010. This work was supported by funding from NASA's Applied Sciences Program and grant NNS06AA051 entitled "Integrating NASA's Global Soil Moisture Remote Sensing and Modeling Data into the USDA's Global Crop Production Decision Support System".

J. D. Bolten is with the Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: john.bolten@nasa.gov).

W. T. Crow and T. J. Jackson are with the Hydrology and Remote Sensing Lab, USDA-ARS, Beltsville, MD 20705 USA (e-mail: wade.crow@ars.usda.gov; tom.jackson@ars.usda.gov).

X. Zhan is with the Center for Satellite Applications and Research, NOAA National Environmental Satellite, Data, and Information Service, Camp Springs, MD 20746 USA (e-mail: xiwu.zhan@noaa.gov).

C. A. Reynolds is with the International Production Assessment Division USDA Foreign Agricultural Service, Office of Global Analysis, Washington, DC 20002 USA (e-mail: curt.reynolds@fas.usda.gov).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2009.2037163

I. Introduction

THE International Production Assessment Division (IPAD) is the agricultural forecasting division of the Office of Global Analysis (OGA) within the U.S. Department of Agriculture's (USDA) Foreign Agricultural Service (FAS). IPAD is responsible for providing monthly global crop estimates and projected crop yields to monitor global crop conditions and ensure agricultural economic security. These estimates are used for decision making by the U.S. Government to determine food-insecure geographical regions and their potential for affecting national security. Consequently, IPAD plays an active role in monitoring and enhancing world food security by alerting policymakers of potential food security problems well in advance, working closely with the U.S. Agency for International Development (USAID), and providing direct support to the Famine Early Warning System (FEWS-NET).

IPAD relies on many information sources and utilizes a convergence of evidence methodology for comparing data, minimizing risk of error, and maximizing the reliability of foreign crop production, area, and yield forecasts during the growing season. In an effort to determine anomalous meteorological conditions indicating times of water stress or flooding which impact these crop condition assessments, IPAD analysts compare current global agro-meteorological conditions against a database of archived satellite imagery and crop yields. To this end, estimates from IPAD are derived from a merging of many data sources including satellite remote sensing and ground observations, and more than 20 years of climatology and crop behavior data over key agricultural areas. To most efficiently manage these data sources, IPAD has developed a series of analytical tools, crop models, and hazard calendars within a Crop Condition Data Retrieval and Evaluation (CADRE) Data Base Management System (DBMS).

A crucial requirement of these global crop yield forecasts is the regional characterization of root-zone soil moisture. By capturing the impact of agricultural drought (i.e., the lack of root-zone soil moisture) on crop health and eventual yield, IPAD analysts can better prepare for and prevent possible food shortages and agricultural disasters. However, the accurate estimation of regional soil moisture dynamics based on sparse ground measurements is difficult due to soil moisture heterogeneity caused by the spatial heterogeneity of precipitation events, land cover,

soil properties, and topography. In particular, temporal and spatial resolution of root-zone soil moisture is important for predicting adequate soil profile wetting and drying between precipitation events. Thus, root-zone soil moisture availability is a major factor impacting IPAD yield forecasts.

The CADRE DBMS system estimates root-zone soil moisture (and, therefore, the severity of agricultural drought) using a modified 2-layer Palmer model forced by precipitation and temperature datasets operationally obtained from the World Meteorological Organization (WMO) and U.S. Air Force Weather Agency (AFWA). However, soil moisture estimates derived from this type of global modeling suffer from a range of deficiencies including: poor quality rainfall input, uncertain parameter values and over-simplified vertical and lateral physics [1], [2]. This work aims at reducing the impact of these deficiencies on IPAD root-zone soil moisture estimates by assimilating surface soil moisture observations from the NASA Advanced Microwave Scanning Radiometer (AMSR-E) into CADRE using an Ensemble Kalman Filter (EnKF). In this way, surface soil moisture dynamics observed by AMSR-E can be used to indirectly update the root-zone through the vertical soil moisture coupling of the 2-layer soil moisture model. The hypothesis being that the improved temporal resolution and spatial coverage of AMSR-E retrievals over ground station data and model outputs used by IPAD will provide a better characterization of surface wetness and enable more accurate crop monitoring in key agricultural areas. While assimilation strategies for improved root-zone soil moisture monitoring have recently been presented [2]–[5], there still remains a need for a quantitative evaluation of the utility of these assimilation strategies for soil profile estimation. Within this study, we explicitly test our hypothesis using a data denial experimental design in which AMSR-E soil moisture retrievals are used to correct root-zone model estimates (obtained using a low-quality rainfall product) back to benchmark modeling levels based on the input of high accuracy rainfall measurements. The assimilated surface soil moisture retrievals and IPAD water balance model are described in Sections II and III, respectively. Section IV describes the EnKF-based data assimilation system which is applied within the data denial experimental framework described in Section V. Results are presented in Section VI, and the summary and conclusions are given in Section VII.

II. AMSR-E Surface Soil Moisture Retrievals

This work is based on near-daily surface soil moisture estimates derived from the satellite-based Advanced Microwave Scanning Radiometer (AMSR-E). AMSR-E was launched in 2002 on board the NASA EOS Aqua satellite to provide global coverage of passive microwave measurements of terrestrial, oceanic, and atmospheric variables for the investigation of global water and energy cycles [6]. Aqua follows a sun-synchronous orbit with equatorial crossing at approximately 1330 Local Standard Time (LST) and completes full global coverage every 2–3 days at the equator and more frequently at higher altitudes. AMSR-E measures brightness temperatures at six frequencies, 6.92, 10.65, 18.7, 23.8, 36.5, and 89.0 GHz, with vertical and horizontal polarizations at each frequency. With a fixed incidence angle of 54.8° and an altitude of 705 km, AMSR-E provides a conically scanning footprint pattern with a swath width of 1445 km. The mean footprint diameter ranges from 56 km at 6.92 GHz to 5 km at 89 GHz.

In our approach, we retrieve surface soil moisture estimates from AMSR-E brightness temperatures based upon an algorithm developed by Jackson [7]. Jackson's retrieval method utilizes a physically based forward model of microwave emission from the soil-vegetation-atmosphere medium. The algorithm uses horizontally polarized AMSR-E brightness temperatures at 10.7 GHz re-scaled to a 1/4° grid. The effective emissivity from each pixel is calculated by independently modeling the microwave emission from the bare soil layer and the emission and attenuation from the vegetation layer. For proper implementation, the model requires ancillary input data including soil texture and porosity, land cover, the Normalized Difference Vegetation Index (NDVI). Surface temperature is estimated from the AMSR-E vertically polarized 37 GHz brightness temperature according to the equation presented in [8]. Effects of atmospheric scattering at these wavelengths (10.7 GHz) are considered to be minimal and neglected. Within this framework, observed brightness temperatures from AMSR-E are divided by measured soil temperature to estimate emissivity over each pixel. This measured emissivity is then isolated and used to solve for volumetric soil moisture by computing the dielectric constant from the dielectric mixing model [9], soil reflectivity from the Fresnel equations [10], and corrected for vegetation effects as in [11]. In this way, a near-daily soil moisture product is obtained by combining retrievals obtained from both ascending (1:30 pm) and descending (1:30 am) AMSR-E overpasses. The version of this algorithm applied to AMSR-E brightness temperatures can be found in [12]. Recent evaluation results suggest that this single-polarization approach is more effective at retrieving soil moisture over the continental United States than competing approaches based on multipolarization brightness temperature [13]. Traditionally, the sampling depth of soil moisture estimates at 10.7 GHz is assumed to be approximately 1 cm depending on soil type, moisture content, etc. For this study, the soil moisture estimates from AMSR-E are assumed to represent a soil depth comparable to the surface layer used by the IPAD modified Palmer model. Within the context described here, the risk of error introduced from this mismatch of sampling and modeling depth is considered minimal.

III. IPAD Water Balance Model

The IPAD DBMS utilizes a wealth of data sources including over 3000 ground observations from the WMO and climatological estimates provided by the AFWA. The primary inputs for the IPAD water balance model are daily maximum and minimum temperature and precipitation accumulation. The AFWA precipitation and temperature products are provided by the AFWA Agricultural Meteorology modeling system (AGRMET). Daily estimates are calculated from three-hourly analyses of merged gauge reports from AFWA's global surface observation database and remotely sensed climatological data. Consequently, the AFWA rain product is of high quality in specific regions of the world (e.g., North America) in which high-quality daily rain gauge observations are operationally available. However, in other more data poor regions of the

world, it relies heavily on uncertain satellite-based rainfall retrievals, and is prone to high levels of error in short-term rainfall accumulation estimates.

The two-layer soil moisture model used by IPAD was first described by Palmer [14]. The two-layer soil moisture model is a bookkeeping method that accounts for the water gained or lost in the soil profile by recording the amount of water withdrawn by evapotranspiration and replenished by precipitation. The available soil moisture capacity in for the soil column is calculated from the available water content (AWC) of both layers (i.e., AWC_{sz} and AWC_{rz} for the surface and root-zone model layers respectively) derived from the FAO Digital Soil Map of the World [15], soil texture and the total depth of the soil column. The Palmer model surface layer is assumed to contain 2.54 cm of water at soil saturation (i.e., $AWC_{sz} = 2.54$ cm), and the lower layer depends on the depth of the effective root-zone calculated from AWC_{rz}. Daily estimates of minimum and maximum temperature and precipitation are applied to the soil moisture model in order to calculate the daily amount of soil moisture withdrawn by evapotranspiration and replenished by precipitation for two layers of soil. Although the Palmer soil moisture model is simplistic relative to more recent advances in hydrologic modeling, IPAD continues to use this model in order to take advantage of its computational efficiency and historical database of global soil moisture climatology within their DBMS. The purpose of this analysis is to quantify the added value of assimilating surface soil moisture relative to an established operational DBMS. Therefore, in order to accurately represent the current operational IPAD baseline system, we also utilize the Palmer model with our data assimilation system. For each day, soil moisture in the Palmer model surface layer θ_{sz} is calculated as

$$\theta_{sz} = \theta'_{sz} + P - PE - F \tag{1}$$

where θ'_{sz} is soil moisture content in the surface layer from the previous day, P is precipitation, PE is potential evaporation, and F is a diffusion term discussed below. No runoff is calculated for this version of the model; excess water is lost from the system. The root-zone soil moisture θ_{rz} is adjusted as

$$\theta_{rz} = \theta'_{rz} + (P - PE + \Delta\theta_{sz}) + F \tag{2}$$

where θ'_{rz} is the previous day root-zone soil moisture and $\Delta\theta_{sz}$ the net daily change in surface soil moisture. Thus, loss from the root-zone depends on initial moisture content as well as on the PE and P. A no flow boundary condition is assumed for the bottom of the root-zone layer. In this context, all soil moisture model calculations are in units of depth. For compatibility with the volumetric soil moisture retrievals from AMSR-E, we transformed soil moisture units from depth to volumetric content during implementation of our data assimilation system (Section IV) using soil porosity data from the Soil Survey Geographic (SSURGO) database. Consequently, data assimilation results shown later are in volumetric units.

IPAD has altered Palmer's original two-layer soil moisture model shown above via a number of specific modifications. First, moisture is depleted from the lower layer before the surface layer is completely dry, thus allowing for a more gradual and realistic depletion of the surface layer. The modified extraction function allows moisture to be depleted from the surface layer at the potential evapotranspiration rate up to 75% of the surface capacity (or 75% of 2.54 cm of water). When the surface layer is below 75% capacity, moisture is extracted from the surface at a reduced rate with the lower layer making up the remaining requirement. In addition, IPAD has changed the model calculation of potential evapotranspiration to be based on the modified FAO Penman-Monteith equation described in [16]. Due to the lack of a global data set of wind speed and relative humidity within the IPAD DBMS, the potential evapotranspiration calculations assume a constant wind speed of 2 m/s and estimate vapor pressure deficit from minimum and maximum temperature. Atmospheric pressure and extraterrestrial radiation are calculated from station latitude and elevation. In its current operational implementation, IPAD applies the Palmer model at daily time steps within a stereographic projection with approximately 47 km horizontal grid spacing at 60° latitude. Here, all data have been regridded to a 1/4° resolution mesh for compatibility with the AMSR-E soil moisture observations as discussed below.

In data assimilation systems, the constraint of root-zone soil moisture values using surface observations is based on the presence of cross-correlation between errors in surface and root-zone soil moisture predictions made by the model. Such cross-correlation typically requires the presence of diffusion by which anomalies in one layer are vertically propagated into neighboring layers. Within the 2-layer Palmer model used by IPAD, vertical coupling between layers is relatively simplistic and allows recharge of the top layer (and diffusion to the root-zone) based on an assigned value of fractional water volume. Below the fractional water threshold of 75%, root-zone moisture recharge is halted. This artificial truncation of soil moisture diffusion results in a complete loss of coupling between the two layers and reduces the value of surface soil moisture retrievals for constraining deeper root-zone moisture. However, this problem can be eased with the addition of a simple linear diffusion term F

$$F = \alpha(\theta_{sz}/AWC_{sz} - \theta_{rz}/AWC_{rz}) \tag{3}$$

where α is a constant diffusion coefficient. Surface and root-zone soil moisture contents are then adjusted by F following (1) and (2).

This added diffusion term results in a more gradual soil moisture gradient between the surface and root-zone layers, and ensures sufficient vertical communication between the two layers. Here the α coefficient in (3) has been assigned a constant value of 17.9 mm based on a sensitivity analysis as described in Section VI. The impact of this modification on data assimilation results will be discussed later.

IV. Ensemble Kalman Filter

The increased availability of satellite remote sensing products has led to improved meteorological, oceanographic, and land surface predictions through the merging of satellite observations with numerical models [17], [18]. Sequential data assimilation techniques use auto-recursive analyses to optimally

merge model estimates with state observations. If properly done, such merging should yield improved state predictions relative to the accuracy of either the model or observations. Filtering-based applications of data assimilation techniques rely on the availability of an observation y at time k that can be related to the state vector x via a known observation operator H

$$y_k = H_k(x_k) + v_k \tag{4}$$

where v_k represents a random perturbation of the observations. Here, such perturbations are assumed to be Gaussian with a known covariance of R.

In this study, a 1-D Ensemble Kalman filter (EnKF) is applied. The EnKF is a nonlinear extension of the standard Kalman filter first demonstrated by [19] and has been successfully applied to land surface forecasting problems [20]. It is based on adding random errors to the model's internal states or forcing to produce an ensemble of model states x_k^i and predicted observations $H_k(x_k^i)$ where i refers to a particular realization within the ensemble. This Monte Carlo ensemble is then used to sample the error covariance of the remotely sensed observations (CM_k) and the cross-correlation between these observations and each forecasted state variable (CYM_k). The updating step of the EnKF utilizes these error covariance estimates to optimally update forecasts in response to observations based on the calculation of the Kalman gain, defined as

$$K_k = \frac{CYM_k}{(CM_k + R)} \tag{5}$$

and the application of the Kalman filter updating equation

$$x_k^{i,+} = x_k^{i,-} + K_k \left[y_k^i - H_k \left(x_k^{i,-} \right) + v_k^i \right] \tag{6}$$

to each realization within the ensemble. Note that "$-$" and "$+$" notation is used to signify state estimates made before and after updating in response to observations at time k. As in (4), v_k represents a random Gaussian perturbation of the observation with covariance R. After updating via (6), each ensemble member is forecasted in time until the next available observation using the dynamic model. The EnKF state estimate at any time is obtained by sampling the mean of the ensemble.

Our particular implementation of the EnKF integrates soil moisture observations from AMSR-E with the modified Palmer two-layer soil moisture model described in Section III by applying a 1-D EnKF at daily time-steps when AMSR-E observations are available. However, before the AMSR-E soil moisture retrievals can be assimilated, the modeled and observed (AMSR-E) data must be scaled to a common climatology to reduce potential biases and differences in dynamic range that commonly exist between modeled and observed surface soil moisture products. By removing time-invariant biases from the observation data, the two datasets can be optimally merged to allow more efficient assimilation [1]. The removal of multiplicative and additive errors in this way also provides an objective basis for the comparison of soil moisture anomalies and a basis for properly validating the system.

To establish a representative climatology for both AMSR-E and IPAD modified Palmer model surface soil moisture datasets, we constructed a retrospective analysis of archived data from June 2002 to June 2007. The AMSR-E product was rescaled using standard normal deviates based on a mean and standard deviation climatology of AMSR-E soil moisture retrievals and the surface layer of the IPAD soil moisture model within a 31-day sampling window centered on a particular day of the year. In this way, the AMSR-E retrievals are transformed such that their climatology is comparable to the climatology for top layer soil moisture estimates produced by the IPAD-modified Palmer model. The climatologically rescaled AMSR-E data are then introduced as observations to the EnKF using sequential observations of AMSR-E and climatological data. Since AMSR-E observations are preprocessed into surface soil moisture estimates (assumed to be consistent with the top layer of the Palmer model), our observation operator in (4, 6) is simply $H = (1, 0)$.

As noted above, the implementation of the EnKF requires that a Monte Carlo ensemble of Palmer model predictions be generated. Because modeling parameters are static within the Palmer water balance, we assume a temporally constant modeling error for the surface layer. Each member of the model ensemble is generated by applying independent Gaussian noise directly to the Palmer model surface and root-zone soil moisture states. The statistical properties of this noise determine the size of CM and CYM in (5). Here, surface layer perturbations were assumed to be mean-zero with a standard deviation of 0.03 $\mathrm{cm}^3/\mathrm{cm}^3$. Our choice of model error is based on *a priori* model runs over the United States and is considered realistic. The same stochastic perturbation is also applied to the root-zone after it has been scaled by the AWC_{rz} to ensure a physically consistent volumetric perturbation for the entire soil column (i.e., shifts in the root-zone layer are dampened relative to the surface layer according to the soil column water capacity). In this way, the perturbations diminish with increases in soil column depth, porosity, and water holding capacity.

An additional challenge in applying the EnKF to the assimilation of surface soil moisture retrievals into a land surface model is the accurate estimation of the observation error R in (5). The relative magnitude of R (versus competing errors assumed to be present in the model, i.e., CM) determines the size of K in (5), and, thus, the amount of weight applied to observations upon the implementation of (6). The accuracy of remotely sensed passive microwave observations vary greatly over different land cover types due to signal attenuation by vegetation and increased scattering over rough terrain [10], [21]. At the wavelengths used by AMSR-E, the accuracy of observed soil moisture is significantly degraded over areas of vegetation water content greater than approximately 5 $\mathrm{kg/m}^2$ [22]. Therefore, regarding AMSR-E soil moisture accuracy, R should correlate significantly with vegetation density. We exploit this relation by adjusting the magnitude of R in relation to vegetation type and canopy opacity τ based on vegetation water content w_c as in

$$R = R_o \exp(\tau) \tag{7}$$

and

$$\tau = \frac{b * w_c}{\cos\theta} \tag{8}$$

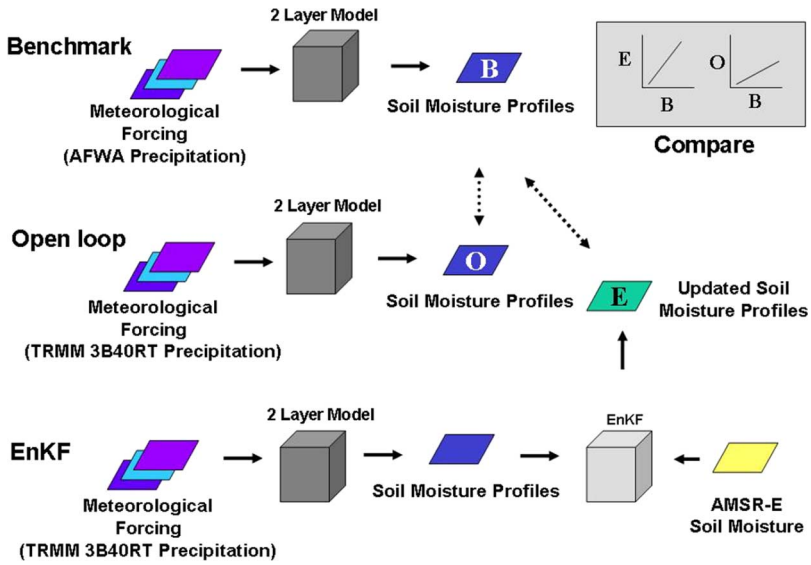

Fig. 1. Schematic of the data denial experiment utilized in this analysis. The experiment is based on the use of AMSR-E soil moisture retrievals and an EnKF to correct the open loop experiment (forced by the satellite-based TRMM 3B40RT rainfall product) back to benchmark results (obtained using the gauge-correct AFWA rainfall product).

where θ is the AMSR-E incidence angle and b is a coefficient that depends on vegetation type and is assigned based on published values for similar sensor parameters and land cover conditions [11]. In this case, b is set to 0.3 for wooded grasslands and shrubs, grasslands, and croplands, and 0.28 for closed bushlands, open shrublands, and bare soil. Since the accuracy of AMSR-E soil moisture retrievals quickly diminishes in areas of vegetation water content above 5.0 kg/m^2, R is set to 1.00 cm^3/cm^3 (effectively driving K to zero) for these pixels (e.g., pixels classified as forested). To discern land cover classes, we apply the 8-km MODIS land cover classification data set produced by the University of Maryland (http://glcf.umiacs.umd.edu/). The EnKF was employed with an ensemble of 30 members perturbed with Gaussian noise having an initial standard deviation of $R_o = 0.15$ cm^3/cm^3. Equation (7) was derived by performing multiple runs of the filter over many different land cover types (i.e., various b and w_c combinations) for the entire dataset and incrementally adjusting R until the normalized filter innovations best matched those expected from a properly parameterized filter (i.e., serially uncorrelated and having a temporal second moment of one) [23]. This dynamic approach ensures that the filter is placing an appropriate relative weight on the model predictions and AMSR-E observations over varied landscapes. As noted above, forested areas of the domain are masked from the analysis due to the inability of AMSR-E to estimate soil moisture in such areas.

V. Experiment Description

For evaluation of the data assimilation system described above, we focus on the creation of a five-year (06/19/2002–06/19/2007), 1/4° latitude/longitude root-zone soil moisture analysis product over North America. A data denial framework is employed to evaluate the analysis product by comparing three separate root-zone soil moisture products. The three products are created by 1) forcing the two-layer Palmer model with a high-quality precipitation product (the benchmark run), 2) forcing the model with a lower-quality satellite-based precipitation data set (the open loop run), and 3) employing an EnKF to assimilate AMSR-E soil moisture retrievals into the open loop run (the EnKF run). The AFWA gauge-corrected precipitation data described in Section III was used for the benchmark loop. This product is a merged analysis of blended surface observations and remotely sensed estimates. Consequently, it is of relatively high-quality in areas of world like our North American study domain which possess good rain gauge coverage. For precipitation forcing in the open loop and EnKF loop, we applied the uncorrected real-time precipitation 3B40RT product provided by the Tropical Rainfall Measuring Mission (TRMM) [24]. 3B40RT is a real-time, satellite-only precipitation product that accurately reflects the challenges of obtaining operational rainfall information within data-poor land areas lacking adequate ground observations. Unlike the AFWA product, it is not corrected using ground rain gauge data over North America. This lack of ground gage correction typically introduces root-mean-squared errors (RMSE) of 5 to 10 mm/day into TRMM 3B40RT daily rainfall accumulation amounts over the central United States [25].

Following the procedure outlined in Section IV, the AMSR-E observations have been scaled to the climatology of the TRMM-forced Palmer model. Using this approach, the application of the EnKF to assimilate remotely sensed soil moisture retrievals in the EnKF run can be evaluated based on how efficiently it transforms root-zone soil moisture results from the open run (generated with the least accurate rainfall product) to match benchmark root-zone soil moisture (generated using the most accurate rainfall product). A flowchart of this data denial process is demonstrated in Fig. 1. Note that similar approaches have been successfully applied in previous attempts to evaluate the added benefit of assimilating remotely sensed observations into a land surface model (see, e.g., [3] and [13]).

Modeled and observed soil moistures are often not representative of each other due to inconsistent modeling and sensing depths and time-invariant multiplicative and additive errors [26]. In this case, we are evaluating the performance of the EnKF using two distinct rainfall products, each having a unique climatology and bias. As a result, proper validation of the soil moisture analysis products is challenging. Therefore, to maintain continuity between the rainfall products and ensure a true evaluation of the EnKF performance, we removed the long-term bias from the TRMM data (i.e., matched long-term mean of the AFWA precipitation) and scaled the AMSR-E observations to the bias corrected TRMM forced (i.e., "open loop") model run estimates prior the its inclusion into either the open loop or EnKF runs.

Using the approach in Fig. 1, we quantitatively evaluate the filter performance by calculating the improvement in RMSE, determined relative to the benchmark run, found between the open loop and EnKF runs. The RMSE difference noted upon assimilation of AMSR-E soil moisture is referred to as the "delta RMSE". By taking these differences, we can evaluate the degree of improvement provided by the EnKF application from the sign and size of delta RMSE values.

In addition, after each run was completed, anomalies were calculated from climatological expectations for each model run. The calculation of climatological expectations on a given day of the year are based on the available multiyear heritage of each soil moisture product (2002 to 2007) and the 31-day mean sampling approach described in Section IV. From these anomaly values, another diagnostic was calculated from the difference in Pearson's correlation coefficient r between the runs (i.e., "delta r"). Because seasonal variations are removed during this procedure, improvements in such an "anomaly" delta r metric can be attributed solely to enhanced skill with regard to anomaly detection and is independent of a product's particular climatology and/or bias. To accurately forecast deviations in yield versus historical expectations, IPAD analysts are primarily interested in identifying anomalies in root-zone soil water anomaly conditions (i.e., relative wet or relative dry versus an expected seasonal soil moisture climatology). Therefore, it is expected that a comparison of delta r anomalies, reflecting areas/conditions of interest, is an appropriate test of the potential added value of the integrated product for IPAD agricultural drought monitoring applications.

VI. Results

We aim to show through our data denial strategy that complementary information from AMSR-E observations can improve the spatial and temporal characterization of root-zone soil moisture when applied to the IPAD modified Palmer model. The target areas for IPAD are nonforested land—particularly regions with potential for agricultural development. As an initial exercise, the impact of including the vertical diffusion term in (2) was investigated by applying the data denial strategy with multiple choices for α ranging from 0 to 25 mm. In general, results associated with the inclusion of the diffusion term (i.e., $\alpha > 0$) lead to improved data denial delta r and delta RMSE results relative to the baseline case of no diffusion ($\alpha = 0$). This indicates that inclusion of the term results in a more efficient correction of root-zone soil moisture errors associated with poor rainfall forcing. Since it appears to provide a more realistic representation of vertical water flow, the diffusive term in (2) was included in all subsequent denial exercises. A specific value of $\alpha = 17.9$ mm was selected based on an analysis of the sensitivity between delta r and α. However, modest levels of sensitivity between delta r and α suggest that the following results are relatively insensitive to the detailed specification of α.

Assessment of the system for particular land cover types is possible by evaluating time series generated during the entire five-year simulation period over selected areas. Fig. 2 shows a time series of the benchmark (blue lines), open loop (green lines), EnKF cases (red lines), and AFWA precipitation product for a 1/4° pixel box in the South-Central United States (39° N and 90° W). Land cover in this pixel is predominantly cropland, with maximum vegetation water content approaching 4 kg/m^2. Note how contrasts in the accuracy of rainfall products forcing between the benchmark and open runs leads to substantial differences in both surface and root-zone soil moisture predictions. These differences are partially compensated for via the EnKF-based assimilation of surface soil moisture retrievals. Despite the fact that AMSR-E cannot directly observe the root-zone, application of the EnKF (red lines) ensures that the benefits of the surface observations extend downward into the root-zone, and the EnKF root-zone soil moisture predictions are better able to capture benchmark variations relative to the open loop case. These improvements are realized in both the surface and root-zone soil moisture layers throughout much of the time-series. For example, during the latter half of 2003, multiple precipitation events are missed by the open loop. It is clear during this time period that assimilation of the AMSR-E observations leads to more realistic soil moisture predictions (i.e., higher correlation with benchmark run) in both layers. For this particular location, the application of the EnKF leads to a delta r of 0.09 for root-zone anomalies and a delta RMSE of -0.02 cm^3/cm^3 for root-zone raw values. The agricultural landscape at this site is a good representation of an IPAD target area and effectively demonstrates the added value of assimilating AMSR-E observations into the TRMM-forced, IPAD-modified Palmer model.

In order to examine the geographic extent of areas in which the assimilation of AMSR-E soil moisture retrievals adds value to IPAD root-zone soil moisture model estimates, the analysis in Fig. 2 was extended over the North American continent between 23° N—50° N and 128° W—65° W. For each 1/4° grid in this domain, delta RMSE and delta r values were calculated using daily soil moisture estimates obtained between 06/19/2002–06/19/2007. To reduce the impact of snow on our analysis, we examined only the growing season (May–October) for the region north of 42° N and east of 112° W. The analysis for the remainder of the North American region was applied to the entire year. Fig. 3 plots the root-zone soil moisture delta RMSE. Red (negative) pixels indicate areas of reduced delta RMSE and improvement upon the nonupdated, TRMM-forced open loop. In comparison, blue (positive) pixels represent an increase in error and degradation of the soil moisture estimates. Because K in (5) is effectively set to zero for very densely

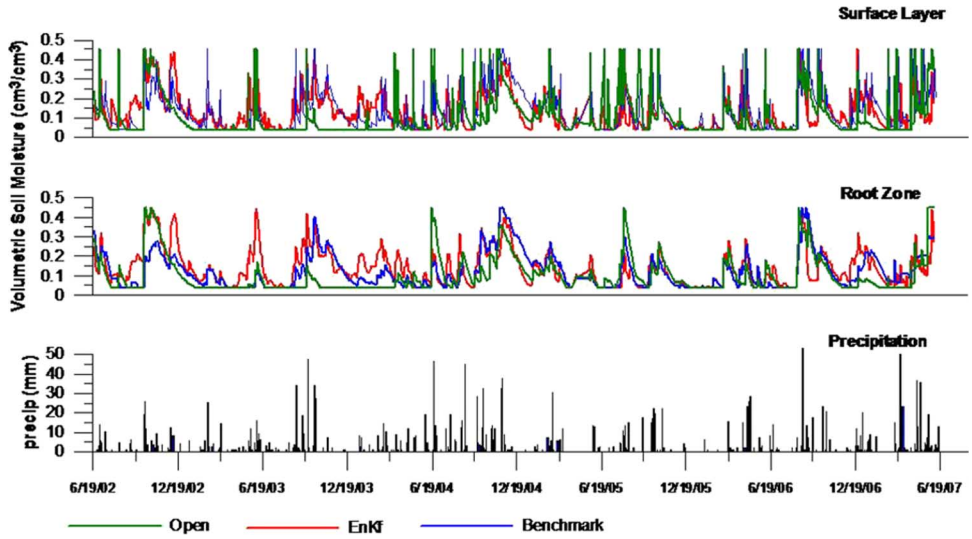

Fig. 2. Time-series of benchmark soil moisture (blue), EnKF soil moisture (red), open loop run soil moisture (green), and AFWA precipitation (black) over cropland area for the 1/4° grid cell location at 39° N and 90° W.

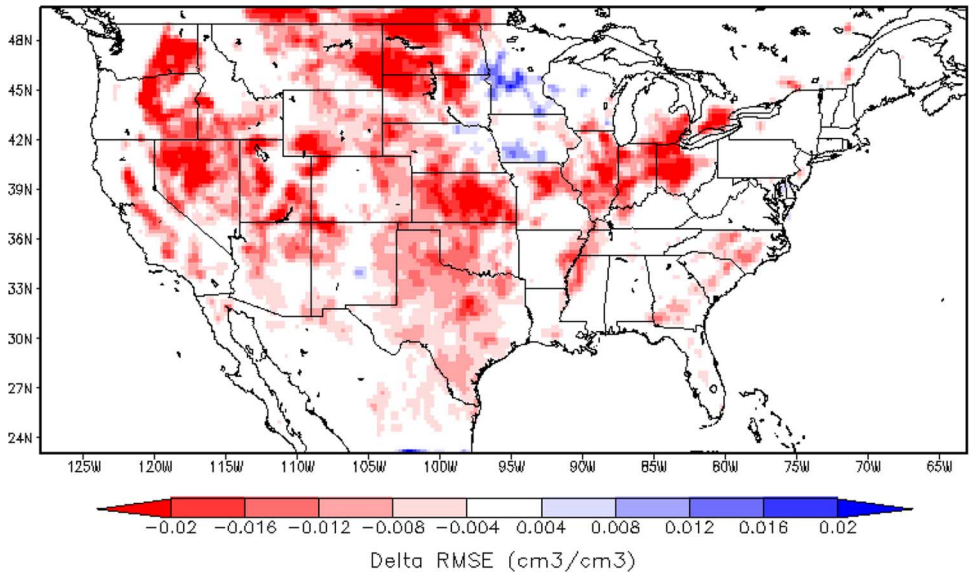

Fig. 3. Root-zone delta RMSE soil moisture results over North America for the five-year data denial experiment. Negative (positive) values shaded in red (blue) indicate areas of improvement (degradation) relative to the open loop upon application of the EnKF.

vegetated areas (i.e., $w_c > 5.0$ kg/m^2), a large portion of the domain has no delta RMSE, particularly in the Northeast and Northwest United States. This figure illustrates the improved correlation of the EnKF root-zone soil moisture results with benchmark values for most of the un-masked regions of the Midwest and Pacific regions, particularly in northern Texas, Oklahoma, Kansas, and within the Ohio River Basin. Negative delta RMSE values are found for much of the lightly vegetated or bare soil regions of the domain, with a majority of the improvements in RMSE being greater than 0.016 cm^3/cm^3. Degradation of root-zone RMSE indicated by positive (blue) shading in the figure is limited mostly to Iowa, Minnesota, and Wisconsin. These areas of degradation are relatively isolated and of small intensity (majority lower than 0.06 cm^3/cm^3) yet indicate error introduced into the moisture estimates from the assimilated AMSR-E observations. Some possible explanations for this introduced error are discussed below.

In a similar manner to Fig. 3, Fig. 4 illustrates root-zone anomaly delta r for the same time period. Positive (red) gains in correlation coefficients indicate that the assimilation of AMSR-E surface soil moisture retrievals is enhancing IPAD's ability to characterize soil moisture anomalies (by creating a higher degree of consistency with the benchmark results). Conversely, negative (blue) values indicate areas of degradation with respect to the benchmark run upon implementation of the EnKF. Positive soil moisture impacts are observed along a wide swath of the central United States. Negative differences are generally restricted to mountainous (e.g., Western Colorado, Western Wyoming and Idaho), closed canopy shrubland areas (e.g., Eastern Montana, Central New Mexico) and/or heavily vegetated regions (e.g., the Upper Midwest) known to be

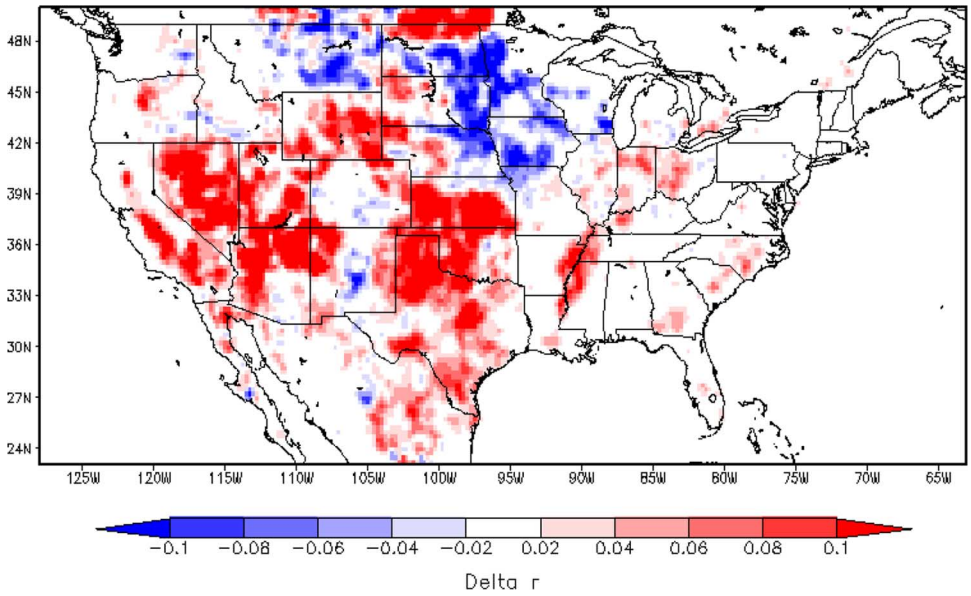

Fig. 4. Root-zone delta r soil moisture anomaly results over North America for the five-year data denial experiment. Positive (negative) values shaded in red (blue) indicate areas of improvement (degradation) relative to the open loop upon application of the EnKF.

challenging for soil moisture remote sensing at the 10.7-GHz frequency used here. However, the spatial pattern of improved root-zone soil moisture correlation is similar to the delta RMSE results, with well-defined increases of delta r (predominantly greater than 0.08) along the Southern United States. The calculation of significance levels for these correlation improvements is complicated by the presence of substantial temporal and spatial auto-correlation within root-zone soil moisture fields [27]. However, even very conservative hypothesis testing indicates that the large-scale pattern of improvement in the Southern United States is significant at a 95% confidence level. In addition, it has been shown that even modest improvements in correlation-based skill are often of great value in operational settings [2].

Poor performance in desert areas of Southern Arizona and Nevada reflect known difficulties with microwave-based surface soil moisture retrievals in highly arid regions. In contrast, the spatial pattern of negative delta r values near Minnesota and Iowa in Fig. 4 is more difficult to interpret. The removal of the cold season from the analysis should mask the potential degrading impact of ground snow cover on AMSR-E retrievals. Several other potential factors may contribute to poor results in this area. First, this area has an extensive network of wetlands, rivers, and streams. In addition, it is relatively densely vegetated; the primary crop grown in these regions is corn which can exceed 5 kg/m^2 during the growing season. Both of these factors can have a negative impact on AMSR-E soil moisture retrievals. Also, it is evident from the AWC dataset that this region also has a higher maximum water capacity (>25 cm) than most other areas of North American (not shown). The combination of high AWC and low annual evapotranspiration in these areas may lead to slowly varying root-zone soil moisture dynamics and, therefore, little basis on which to evaluate the improved detection of temporal root-zone soil moisture variations.

Fig. 5 sub-divides the root-zone delta r correlation results in Fig. 4 into five main land cover classes included in the study to demonstrate the effectiveness of the system over selected land cover types. Land cover classes shown are: 1-wooded grasslands and shrubs, 2-closed bushlands, 3-open shrublands, 4-grasslands, 5-croplands, and 6-bare soil. Each box is a culmination of all pixels of similar land cover. It is evident from the figure that there is a net improvement (i.e., positive mean delta r value as shown by the horizontal line) for all land cover classes—with an overall mean increase in delta r of 0.04. Pixels dominated by grasslands (i.e., land cover 4) show the most improvement with a mean delta r value of 0.05. The analysis also indicates that the filter performs reasonably well in croplands, giving a mean increase in delta r of 0.04. Improvements in such agricultural areas are, of course, the highest priority for the USDA IPAD DBMS. However, there are areas of reduced performance. The delta r analysis varied most over pixels dominated by land cover type 2 (closed bushlands). Some of this variance can simply be explained by the diversity of land surface conditions encapsulated within the rather broad "closed bushlands" classification (e.g., evergreen, deciduous, and herbaceous vegetation found within areas of varying topographic relief). This leads to a large number of pixels of mixed performance being lumped into land cover type 2 relative to the other land cover types (e.g., 3449 pixels are classified as land cover type 2 versus 615 for land cover type 5). In addition, AMSR-E performance over areas of wooded grassland (i.e., land cover type 1) and closed shrublands (i.e., land cover type 2) is not expected to be optimal due to vegetation density limitations (see Section II). Still, the assimilation of AMSR-E observations into the IPAD modified Palmer model improves root-zone anomaly correlations for many pixels within these domains and demonstrates the utility of this methodology.

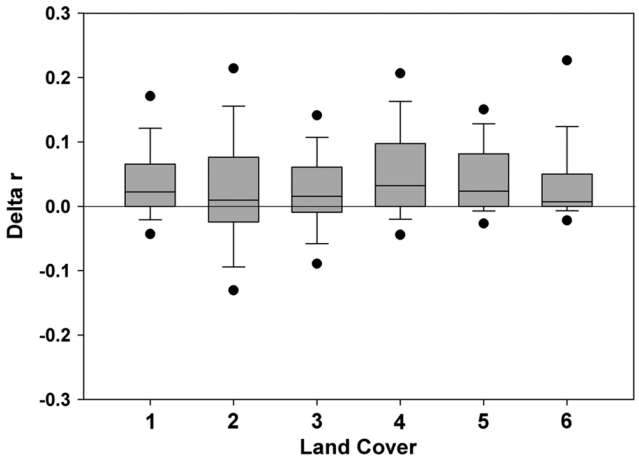

Fig. 5. Box plot of root-zone delta r soil moisture anomaly results over North America for the five-year data denial experiment. Positive values indicate areas of improvement over the open loop realized upon AMSR-E soil moisture data assimilation. Land cover values are: 1-wooded grasslands and shrubs, 2-closed bushlands, 3-open shrublands, 4-grasslands, 5-croplands, and 6-bare soil.

VII. Summary and Conclusions

This analysis attempts to define the contribution of integrating AMSR-E soil moisture retrievals into the drought detecting capability of the USDA IPAD soil moisture model. Given that the TRMM 3B40RT rainfall product (used to force the open loop case within the data denial experiment) accurately reflects the quality of real-time rainfall accumulation data available in data-poor areas, results in Figs. 2–5 provide a credible estimate of the added utility provided by AMSR-E surface soil moisture retrievals for drought applications (like the USDA IPAD DSS) requiring near real-time root-zone soil moisture estimates within (potentially) data-poor land regions. Net improvement is noted in our ability to track root-zone soil moisture temporal dynamics (Figs. 2–4) and is observed for all nonforested land cover types within the North American study domain—most notably cropland areas of prime importance for the IPAD agricultural drought DSS (Fig. 5). The data denial experiment conducted here is limited in that it focuses solely on the ability of AMSR-E soil moisture retrievals to reduce modeling errors associated with poor-quality rainfall errors. However, for global modeling applications based on real-time rainfall observations obtained from satellite sensors, such errors are expected to be large, and may dominate the total modeling error budget [28], [29]. Follow-up work with an alternative data denial design is required to examine the potential skill associated with correcting other error sources (e.g., the poor internal estimation of evapotranspiration by the model). In addition, our data denial experimental design can be extended by employing soil moisture estimates from each model run within a crop forecasting model. In this manner, the value of integrating AMSR-E soil moisture retrievals into the IPAD modified Palmer model can be further evaluated by identifying times of water stress on crop forecasts initialized by each model run.

This methodology holds promise for applying remotely sensed soil moisture observations for more accurate characterization of root-zone conditions at the regional scale, with possible application in crop yield forecasting and the monitoring of anomalous agro-meteorological events. Studies demonstrating the added benefit of using remotely sensed soil moisture observations as shown here are essential given the expected launch of several soil moisture-focused missions in the near future. For example, the European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) mission in 2009 and the National Aeronautics and Space Administration (NASA) Soil Moisture Active/Passive mission scheduled for launch before 2014 will both provide improved global soil moisture observations that can be used to further enhance the global characterization of agricultural drought conditions.

References

[1] R. H. Reichle and R. D. Koster, "Bias reduction in short records of satellite soil moisture," *Geophys. Res. Lett.*, vol. 31, p. L19501, 2004, DOI: 10.1029/2004GL020938.

[2] R. H. Reichle and R. D. Koster, "Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model," *Geophys. Res. Lett.*, vol. 32, p. L02404, 2005, DOI: 10.1029/2004GL021700.

[3] V. Lakshmi, "A simple surface temperature assimilation scheme for use in land surface models," *Wat. Resour. Res.*, vol. 36, no. 12, pp. 3687–3700, 2000.

[4] N. Montaldo, J. D. Albertson, M. Marcini, and G. Kiely, "Robust prediction of root zone soil moisture from assimilation of surface soil moisture," *Wat. Resour. Res.*, vol. 37, pp. 2889–2900, 2001.

[5] W. T. Crow, W. P. Kustas, and J. H. Prueger, "Monitoring root-zone soil moisture through the assimilation of a thermal remote sensing-based soil moisture proxy into a water balance model," *Remote Sens. Environ.*, vol. 112, pp. 1268–1281, 2008.

[6] E. G. Njoku, T. J. Jackson, V. Lakshmi, T. K. Chan, and S. V. Nghiem, "Soil moisture retrieval from AMSR-E," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, pp. 215–229, 2003.

[7] T. J. Jackson, "III. Measuring surface soil moisture using passive microwave remote sensing," *Hydro. Proc.*, vol. 7, pp. 139–152, 1993.

[8] R. A. M. De Jeu, "Retrieval of Land Surface Parameters Using Passive Microwave Remote Sensing," Ph.D. dissertation, Faculty of Earth Science, Vrije Univ., Amsterdam, The Netherlands, 2003.

[9] J. R. Wang and T. J. Schmugge, "An empirical model for the complex dielectric permittivity of soils as a function of water content," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-18, pp. 288–295, 1980.

[10] F. T. Ulaby and A. K. Fung, *Microwave Remote Sensing: Active and Passive-Volume I Scattering and Emission Theory, Advanced Systems and Applications*. Dedham, MA: Artech House, 1986.

[11] T. J. Jackson and T. J. Schmugge, "Vegetation effects on the microwave emission of soils," *Remote Sens. Environ.*, vol. 363, pp. 203–212, 1991.

[12] T. J. Jackson, R. Hurkmans, A. Hsu, and M. H. Cosh, "Soil moisture algorithm validation using data from the Advanced Microwave Scanning Radiometer (AMSR-E) in Mongolia," *Ital. J. Remote. Sens.*, vol. 30/31, pp. 39–52, 2004.

[13] W. T. Crow and X. Zhan, "Continental-scale evaluation of remotely-sensed soil moisture products," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 451–455, Mar. 2007.

[14] W. C. Palmer, "Meteorological drought," U.S. Weather Bureau Research Paper 45, 1965.

[15] Digital Soil Map of the World and Derived Soil Properties FAO, United Nations [Online]. Available: http://www.fao.org/ag/agl/lwdms.stm#cd1

[16] R. G. Allen, L. S. Pereira, D. Raes, and M. Smith, "Crop evapotranspiration guidelines for computing crop water requirements," FAO Irrigation and Drainage Paper 56, pp. 27–65, 1998.

[17] W. T. Crow and E. Wood, "The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97," *Adv. Wat. Res.*, vol. 26, pp. 137–149, 2003.

[18] D. McLaughlin, A. O'Neill, J. Derber, and M. Kamachi, "Opportunities for enhanced collaboration within the data assimilation community," *Quart. J. Roy. Meterol. Soc.*, vol. 131, pp. 3683–3693, 2005.

[19] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *J. Geophys. Res.*, vol. 99, pp. 10143–10162, 1994.

[20] R. H. Reichle, D. B. McLaughlin, and D. Entekhabi, "Hydrologic data assimilation with the ensemble Kalman filter," *Mon. Weat. Rev.*, vol. 130, pp. 103–114, 2002.
[21] J. D. Bolten, V. Lakshmi, and E. G. Njoku, "Soil moisture retrieval using the passive/active L- and S-band radar/radiometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, pp. 2792–2801, 2003.
[22] E. Njoku and S. K. Chan, "Vegetation and surface roughness effects on AMSR-E land observations," *Remote. Sens. Environ.*, vol. 100, pp. 190–199, 2006.
[23] R. K. Mehra, "On-line identification of linear dynamic systems with applications to Kalman filtering," *IEEE Trans. Autom. Control.*, vol. AC-16, pp. 12–21, 1971.
[24] G. J. Huffman, R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff, "The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale," *J. Hydrometeorol.*, vol. 8, pp. 28–55, 2007.
[25] W. T. Crow and J. D. Bolten, "Estimating precipitation errors using spaceborne surface soil moisture retrievals," *Geophys. Res. Lett.*, vol. 34, p. L08403, 2007, DOI: 10.1029/2007GL029450.
[26] W. T. Crow, R. D. Koster, R. H. Reichle, and H. O. Sharif, "Relevance of time-varying and time-invariant retrieval error sources on the utility of spaceborne soil moisture products," *Geophys. Res. Lett.*, vol. 32, p. L24405, 2005, DOI: 10.1029/2005GL024889.
[27] H. Van Storch and F. W. Zwiers, *Statistical Climatology*. Cambridge, U.K.: Cambridge Univ. Press, 2002, 496 pp..
[28] J. Gottschalck, J. Meng, M. Rodell, and P. Houser, "Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states," *J. Hydrometeorol.*, vol. 6, pp. 573–598, 2005.
[29] T. Oki, T. Nishimura, and P. Dirmeyer, "Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP)," *J. Meteor. Soc. Jpn.*, vol. 78, pp. 235–255, 1999.

John D. Bolten (M'05) received the M.S. and Ph.D. degrees in geological science with an emphasis in remote sensing from the University of South Carolina, Columbia in 2001 and 2005, respectively.

He is currently a Physical Scientist at the Hydrological Sciences Branch, NASA Goddard Space Flight Center, Beltsville, MD. His primary interests are the utilization of microwave remote sensing in land surface modeling and data assimilation systems.

Wade T. Crow (M'03) received the Ph.D. degree from Princeton University, Princeton, NJ, in 2001.

He is currently a Research Physical Scientist with the Hydrology and Remote Sensing Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, MD. His research involves the development of land data assimilation techniques to enhance the utility of remote sensing observations for hydrologic and agricultural applications.

Xiwu Zhan received the Ph.D. degree in soil, crops, and atmospheric sciences from Cornell University, Ithaca, NY, in 1995.

He is currently with the Center for Satellite Applications and Research of NOAA-NESDIS. His research interests include land surface soil moisture and vegetation remote sensing, modeling, and data assimilation for regional and global water, energy, and carbon cycle studies.

Thomas J. Jackson (F'02) received the Ph.D. degree in civil engineering from the University of Maryland, College Park, in 1976.

He joined ARS in 1977. He is a Research Hydrologist with the U.S. Department of Agriculture, Agricultural Research Service, Hydrology, and Remote Sensing Lab. His research involves the application and development of remote sensing technology in hydrology and agriculture, primarily microwave measurement of soil moisture. He is or has been a member of the science and validation teams of numerous satellite missions including Aqua, ADEOS-II, ALOS, SMOS, and Aquarius. He is a member of the SMAP Science Definition Team and leads the Calibration and Validation component.

Dr. Jackson is a Fellow of the IEEE, the American Meteorological Society, and the American Geophysical Union. In 2003, he received the William T. Pecora Award (NASA and Dept. of Interior) for outstanding contributions toward understanding the Earth by means of remote sensing.

Curt A. Reynolds received the B.S. degree in civil and environmental engineering from the University of Wisconsin-Madison in 1983 and from the University of Arizona, and the M.S. degree in civil engineering and engineering mechanics in 1993 and the Ph.D. degree in agricultural and biosystems engineering in 1998.

He is currently a Remote Sensing Crop Analyst and GIS Crop Modeler for the U.S Department of Agriculture's Foreign Agricultural Service (FAS). His primary interests are in spatial hydrology and spatial crop yield modeling.