An Observing System Simulation Experiment for Hydros Radiometer-Only Soil Moisture Products

Wade T. Crow, *Member, IEEE*, Steven Tsz K. Chan, *Senior Member, IEEE*, Dara Entekhabi, *Senior Member, IEEE*, Paul R. Houser, Ann Y. Hsu, Thomas J. Jackson, *Fellow, IEEE*, Eni G. Njoku, *Fellow, IEEE*, Peggy E. O'Neill, *Senior Member, IEEE*, Jiancheng Shi, *Senior Member, IEEE*, and Xiwu Zhan

***Abstract*—Based on 1-km land surface model geophysical predictions within the United States Southern Great Plains (Red-Arkansas River basin), an observing system simulation experiment (OSSE) is carried out to assess the impact of land surface heterogeneity, instrument error, and parameter uncertainty on soil moisture products derived from the National Aeronautics and Space Administration Hydrosphere State (Hydros) mission. Simulated retrieved soil moisture products are created using three distinct retrieval algorithms based on the characteristics of passive microwave measurements expected from Hydros. The accuracy of retrieval products is evaluated through comparisons with benchmark soil moisture fields obtained from direct aggregation of the original simulated soil moisture fields. The analysis provides a quantitative description of how land surface heterogeneity, instrument error, and inversion parameter uncertainty impacts propagate through the measurement and retrieval process to degrade the accuracy of Hydros soil moisture products. Results demonstrate that the discrete set of error sources captured by the OSSE induce root mean squared errors of between 2.0% and 4.5% volumetric in soil moisture retrievals within the basin. Algorithm robustness is also evaluated for the case of artificially enhanced vegetation water content (W) values within the basin. For large $W(> 3\ \mathrm{kg \cdot m^{-2}})$, a distinct positive bias, attributable to the impact of sub-footprint-scale landcover heterogeneity, is identified in soil moisture retrievals. Prospects for the removal of this bias via a correction strategy for inland water and/or the implementation of an alternative aggregation strategy for surface vegetation and roughness parameters are discussed.**

***Index Terms*—Microwave remote sensing, observing system simulation experiment, soil moisture, spaceborne radiometry.**

Manuscript received July 16, 2004; revised January 19, 2005. This work was supported by the National Aeronautics and Space Administration (NASA) Earth System Science Pathfinder (ESSP) Program Risk-Mitigation Phase funding to the Hydros Science Team. The high-resolution land surface modeling approach and forcing dataset used in this analysis were developed under the direction of E. Wood at Princeton University through support from NASA under Grants NAG8-1517 and NAG5-6494 and from the National Oceanic and Atmospheric Administration under Grant NA96GP0413. The research described in this paper was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under contract to the National Aeronautics and Space Administration.

W. T. Crow, A. Y. Hsu, and T. J. Jackson are with the Hydrology and Remote Sensing Laboratory, U.S. Department of Agriculture, Agricultural Research Service, BARC-West, Beltsville, MD 20705 USA (e-mail: wcrow@hydrolab.arsusda.gov).

S. T. K. Chan and E. G. Njoku are with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA.

D. Entekhabi is with the Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

P. R. Houser, P. E. O'Neill, and X. Zhan are with the Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA.

J. Shi is with the Institute for Computational Earth System Sciences, University of California, Santa Barbara, CA 93106 USA.

Digital Object Identifier 10.1109/TGRS.2005.845645

I. Introduction

An important issue in the development of a dedicated spaceborne soil moisture sensor has been concern over the reliability of soil moisture retrievals in densely vegetated areas and the global extent over which retrievals will be possible. Errors in retrieved soil moisture can originate from a variety of sources within the measurement and retrieval process. In addition to instrument error, two key contributors to retrieval error are the masking of the soil microwave signal by vegetation [1], [2] and the interplay between nonlinear retrieval physics and the relatively poor spatial resolution of passive microwave spaceborne sensors [3]–[6]. Quantification of these errors requires the realistic specification of land surface soil moisture heterogeneity and spatial vegetation patterns. Since detailed soil moisture patterns are currently difficult to obtain from direct observation, an attractive alternative is the application of an observing system simulation experiment (OSSE) in which simulated land surface states are propagated through the sensor measurement and retrieval process to investigate and constrain expected levels of retrieval error [7], [8].

This analysis describes an OSSE performed for soil moisture products to be derived from the NASA Hydrosphere State (Hydros) mission. The Hydros mission is an Earth System Science Pathfinder (ESSP) mission selected in 2002 by NASA for further development and currently scheduled for launch in 2010 [9]. Hydros is designed to provide global maps of the earth's soil moisture and freeze/thaw state every 2–3 days for weather and climate prediction, water, energy, and carbon cycle studies, and natural hazards monitoring. Hydros utilizes a unique active and passive L-band (1.2–1.4 GHz) microwave concept to simultaneously measure microwave emission and backscatter from the surface across a wide spatial swath [9], [10]. The Hydros antenna is an approximately 6-m diameter deployable lightweight mesh reflector that provides footprint sizes of approximately 40 km for the radiometer and 30 km for the radar. The radar resolution is enhanced to 1–3 km using synthetic aperture processing. The key derived products are soil moisture at 40-km resolution for hydroclimatology obtained from the radiometer measurements, soil moisture at 10-km resolution for hydrometeorology obtained by combining the radar and radiometer measurements in a joint retrieval algorithm, and freeze/thaw state at 3-km resolution for terrestrial carbon flux dynamics studies obtained from the radar measurements.

For the OSSE described in this paper, we focus on the 40-km soil moisture product, and use a baseline set of passive mi-

crowave retrieval algorithms to illustrate Hydros retrieval capabilities and examine error propagation in the algorithms. The synthetic experiment is driven by realistically heterogeneous land surface geophysical variables generated from a distributed land surface model. These states are used to derive a set of simulated brightness temperatures which are degraded (i.e., spatially aggregated and randomly perturbed with simulated instrument noise) to simulate Hydros measurements and then inverted back into soil moisture products using various retrieval algorithms. Comparison of these retrievals to the original soil moisture field reveals how the performance of soil moisture retrieval algorithms will be impacted by vegetation density, measurement resolution, inversion parameter uncertainty, and land cover heterogeneity. The intent of this analysis is not to capture the full range of error sources impacting the accuracy of spaceborne soil moisture retrievals. Rather, it is to study the error propagation of a discrete number of error sources through the entire Hydros observation and retrieval system and examine prospects for remediation strategies to reduce the impact of these errors on Hydros soil moisture products. OSSE-type experiments provide a controlled-synthetic environment within which such issues can be addressed prior to sensor launch. The comparison of results in this paper obtained from different retrieval algorithms is not intended as a selection test for the algorithms but as a means to investigate the sensitivities of the different algorithm types to the error sources studied in this OSSE. Continuing evaluations of these and other algorithms will be performed in the years ahead, using a variety of simulation and observational datasets, for further algorithm optimization and selection prior to the launch of Hydros. This particular study focuses exclusively on 40-km Hydros soil moisture products derived from radiometer observations and passive-only soil moisture algorithms. Future work will focus on higher resolution Hydros products derived from radar observations.

II. Hydros Instrument Design

The Hydros instrument is designed for a 670-km circular, Sun-synchronous orbit, with equator crossings at 6 A.M. and 6 P.M. local time. The instrument combines radar and radiometer subsystems that share a single feedhorn and parabolic mesh reflector. The radar operates with vertical (VV), horizontal (HH), and horizontal–vertical (HV) transmit–receive polarizations, and uses separate transmit frequencies for the H (1.26 GHz) and V (1.29 GHz) polarizations. The radiometer operates with V, H and U (third Stokes parameter) polarizations at 1.41 GHz. The reflector is offset from nadir and rotates about the nadir axis at 14.6 rpm, providing a conically scanning antenna beam with a surface incidence angle of approximately 40°. The reflector diameter is roughly 6 m, providing a radiometer footprint of approximately 40 km defined by the one-way 3-dB beamwidth. The two-way 3-dB beamwidth defines the real-aperture radar footprint of approximately 30 km. The swath width of 1000 km provides global coverage within 3 days at the equator and 2 days at boreal latitudes ($> 50°$ N or S). To obtain the desired 3-km spatial resolution, the radar employs range and Doppler discrimination. Relevant Hydros instrument characteristics are listed in Table I. Discussion of the instrument measurement errors and sampling characteristics is deferred to Section VI. Details of the Hydros mission and instrument design are provided in [9] and [10].

TABLE I
Hydros Instrument Characteristics

System:			
Antenna aperture diameter	6 m		
Antenna rotation rate	14.6 rpm		
Beam incidence angle	40°		
Orbit	Sun-synchronous		
Altitude	670 km		
Nodal crossing times (local)	6 am/6 pm		
Radiometer:			
Center frequency	1.41 GHz		
Footprint (root ellipsoidal area)	38 km		
Channels	H, V, U		
Bandwidth, Integration time	25 MHz, 84 msec		
Precision (noise equivalent ΔT combining fore/aft meas)	0.4 K		
Relative calibration error	0.5 K		
Total relative error	0.64 K		
Radar:			
Transmit Frequencies	1.26 GHz (H), 1.29 GHz (V)		
Channels	HH, VV, HV		
PRF, Pulse Length	3.5 kHz, 15 μsec		
Maximum aperture length	32 ms		
Transmit bandwidth	1 MHz		
Peak transmit power	500 W		
Noise equivalent σ°	-39 dB		
Product resolution	3 km	10 km	30 km
Total # looks (worst case, outer swath)	24	267	2400
Relative calibration error	0.4 dB	0.35 dB	0.3 dB
Total relative error (1σ each channel, HH and VV)	1.0 dB	0.45 dB	0.3 dB

III. OSSE Design

The OSSE is designed to simulate Hydros sensor and orbital characteristics. The simulation involves four elements: 1) a land surface model (LSM); 2) a forward microwave emission model (MEM); 3) an orbit and sensor model (OSM); and 4) an inverse retrieval model (RM). A flowchart of the forward and inverse modeling is shown in Fig. 1. The LSM is used to generate a simulated 2-D spatial scene as a 1-km grid of geophysical parameters covering a sufficiently large region of the U.S. to incorporate a diversity of land cover types and representative dynamic ranges of soil moisture, temperature, and other surface characteristics.

Brightness temperatures (T_B) are computed at the Hydros frequencies, polarizations and look angle via the MEM. Inputs to the MEM are provided by 1-km gridded geophysical variables—predicted by the LSM—and observed land surface characteristics. The OSM applies representative orbit sampling and antenna spatial resolution characteristics to the simulated measurements, and adds expected instrument noise and relative calibration error. It outputs simulated Hydros radiometer measurements at 40 km which are run through a series of RMs to obtain 40-km soil moisture estimates. These retrieved soil moistures are compared with the original soil moisture "truth" fields (after degrading these to 40 km) to obtain the error characteristics of the radiometer-based soil moisture retrievals. Consequently, the OSSE can conceptually be divided up into two parts. The

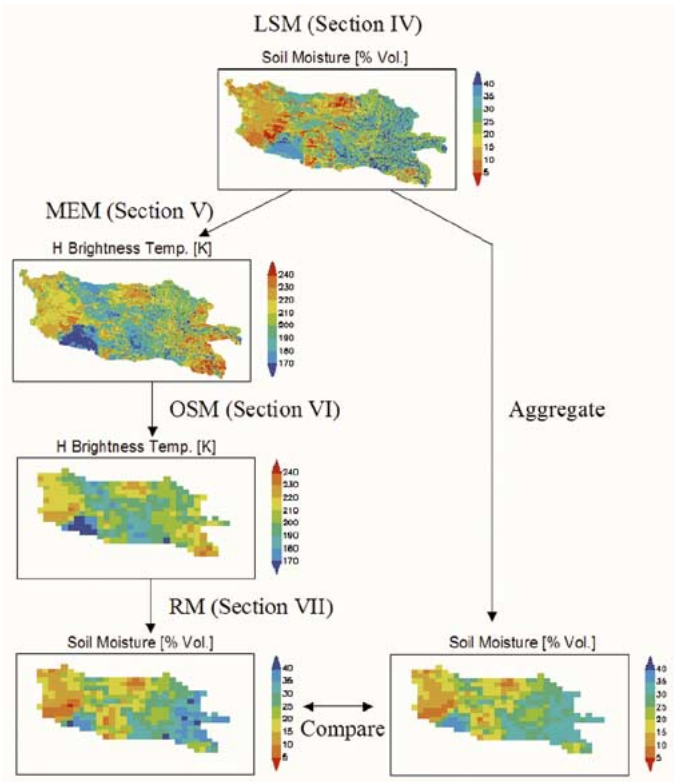


Fig. 1. Illustration of OSSE procedure using simulated soil moisture and brightness temperature imagery. Key components include the land surface modeling (LSM), microwave emission modeling (MEM), orbital and sensor modeling (OSM), and soil moisture retrieval modeling (RM).

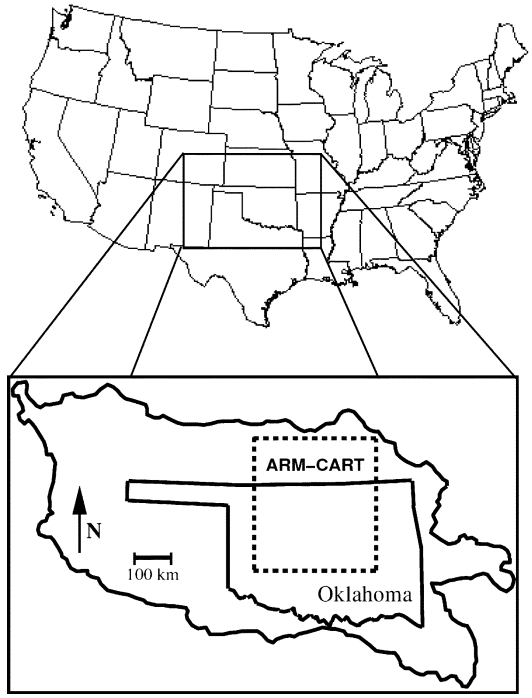


Fig. 2. Location of OSSE domain within the continental United States.

"forward" portion of the simulation—described in Sections IV and V—which consists of the generation, aggregation, and perturbation of synthetic Hydros observations, and the "retrieval" portion—described in Section VI—which entails the inversion of synthetic microwave observations into Hydros soil moisture products.

IV. Land Surface Modeling

The starting point for the simulations is the generation of high-resolution (1-km) geophysical fields using a land surface model (LSM) applied within the United States Southern Great Plains. The LSM used in the computations is the TOPLATS hydrological model [11]. The spatial domain is the 575 000-km^2 Red-Arkansas River Basin in the south-central United States (see Fig. 2). The approach to the land surface modeling and database generation is based on [7] and [12]. The simulations over the spatial domain are forced by the following static datasets: 1) a U.S. Geologic Survey land cover database[1] at 1 km derived from Advanced Very High-Resolution Radiometer (AVHRR) overpasses between April 1992 and March 1993 using the classification methodology presented in [13] and [14]; 2) a soil texture database at 1 km derived from merged State Soil Geographic (Penn State University) products;[2] 3) a USGS digital elevation model at 1 km;[3] and 4) a normalized difference vegetation index (NDVI) database[4] at 1 km derived from AVHRR data and based on composite imagery from June 1995. The land cover and soil texture classifications are listed in Tables II and III.

TOPLATS is run with three water and energy balance layers. The top water balance layer extends from the surface to a depth of 5 cm, and the second layer from 5 cm to the top of the water table (depth to water table is also a dynamic state variable). The top layer soil moisture value is taken to be an integrated value over the top 5 cm. The energy balance discretization has a node at the surface, a node at 5 cm and third node at 50 cm. The 50-cm node is prescribed to vary seasonally. Input forcings are derived from 4-km Weather Surveillance Radar precipitation imagery, 1-km Geostationary Operational Environmental Satellite (GOES) solar radiation imagery, and spatially interpolated National Climate Data Center surface airways meteorology data. Model outputs are generated at hourly time steps and saved in the output database every 12 h at 6 A.M. and 6 P.M. local times to simulate the Hydros overpass times. The outputs generated by the model include: 1) vertically integrated 0- to 5-cm soil moisture expressed as volumetric soil moisture; 2) surface (soil or vegetation) skin temperature; and 3) soil temperature at 5 cm below the surface. Model validation against ground measurements from the Southern Great Plains Atmospheric Radiation Measurement Cloud and Radiation Test bed (ARM-CART) site is described in [12]. The time period for the LSM simulations is from April 1, 1994 to July 31, 1994. Basin-averaged rainfall and simulated soil moisture conditions for this time period are plotted in Fig. 3. However, the actual OSSE is conducted only for the subset of this period between May 26, 1994 to June 28, 1994 that demonstrates a pronounced transition between wet and dry surface conditions.

[1]See http://www.nationalatlas.gov/landcvm.html.

[2]See http://www.essc.psu.edu/.

[3]See http://edcdaac.usgs.gov/gtopo30/gtopo30.asp.

[4]See http://edc.usgs.gov/products/landcover/ndvi.html.

TABLE II
LAND COVER CLASSIFICATIONS AND REPRESENTATIVE ROUGHNESS AND VEGETATION PARAMETERS (DESCRIBED IN TEXT)

Class	Category Name	s (cm)	h	ω	b	b_v	b_h	f_T
1	Crop/mixed farming	1.5	0.15	0.05	0.13	0.143	0.117	0.2
2	Short grass	1.0	0.10	0.05	0.10	0.11	0.09	0.0
3	Evergreen needleleaf tree	1.0	0.10	0.12	0.10	0.12	0.08	0.8
4	Deciduous needleleaf tree	1.0	0.10	0.12	0.10	0.12	0.08	0.8
5	Deciduous broadleaf tree	1.0	0.10	0.12	0.12	0.144	0.096	0.8
6	Evergreen broadleaf tree	1.0	0.10	0.12	0.12	0.144	0.096	0.8
7	Tall grass	1.0	0.10	0.05	0.10	0.11	0.09	0.0
8	Desert	1.0	0.10	0.00	0.00	0.00	0.00	0.0
9	Tundra	1.0	0.10	0.05	0.10	0.11	0.09	0.0
10	Irrigated crop	1.5	0.15	0.05	0.11	0.121	0.099	0.0
11	Semidesert	1.0	0.10	0.05	0.10	0.11	0.09	0.0
12	Bog or marsh	1.0	0.10	0.05	0.10	0.11	0.09	0.0
13	Inland water	0.1	0.01	0.00	0.00	0.00	0.00	0.0
14	Evergreen shrub	1.0	0.10	0.12	0.11	0.121	0.099	0.0
15	Deciduous shrub	1.0	0.10	0.12	0.11	0.121	0.099	0.0
16	Mixed woodland	1.0	0.10	0.12	0.11	0.132	0.088	0.8
17	Short grass/crop	1.2	0.12	0.05	0.13	0.143	0.117	0.1
18	Tall grass/crop	1.2	0.12	0.05	0.13	0.143	0.117	0.1
19	Crop/mixed woodland	1.2	0.12	0.08	0.12	0.138	0.102	0.5
20	Crop/evergreen needleleaf tree	1.2	0.12	0.08	0.12	0.138	0.102	0.5
21	Crop/deciduous broadleaf tree	1.2	0.12	0.08	0.12	0.138	0.102	0.5
22	Irrigated crop/deciduous broadleaf tree	1.2	0.12	0.08	0.12	0.138	0.102	0.5
23	Short grass/mixed woodland	1.0	0.10	0.08	0.11	0.127	0.094	0.4
24	Evergreen needleleaf/short grass	1.0	0.10	0.08	0.11	0.127	0.094	0.4
25	Evergreen needleleaf/ evergreen broadleaf	1.0	0.10	0.12	0.11	0.132	0.088	0.8

TABLE III
SOIL TEXTURE CLASSES

Class	Description	Sand %	Clay %
1	sand	95	3
2	loamy sand	92	6
3	sandy loam	51	14
4	silt loam	31	14
5	silt	10	10
6	loam	42	8.5
7	sandy clay loam	60	28
8	silty clay loam	10	34
9	clay loam	35	34
10	sandy clay	50	43
11	silty clay	5	47
12	clay	20	63

V. MICROWAVE EMISSION MODELING

Based on the LSM output fields, a microwave emission model (MEM) is used to simulate radiometer observations of the geophysical scene. The parameterizations of the MEM represent tradeoffs between the need to adequately represent the key effects of surface characteristics on microwave signatures at the spatial scale of interest (40 km), and the need for a sufficiently simple representation for application to satellite retrieval algorithms. The MEM incorporates the effects of dynamic features, e.g., surface soil moisture and soil temperature, and static features (on the simulation time scale) such as soil texture, soil surface roughness, land-cover and vegetation type, and vegetation water content. Effects of soil freezing, thawing, and snow cover are not considered in the simulations, and effects of atmospheric variability are assumed negligible at L-band for nonraining conditions. Computations are performed at the 40° look angle and at 1.41-GHz V and H polarizations. Heterogeneity is taken into account by aggregating from the 1-km scale up to the 40-km product scale for Hydros radiometer-based soil moisture retrievals.

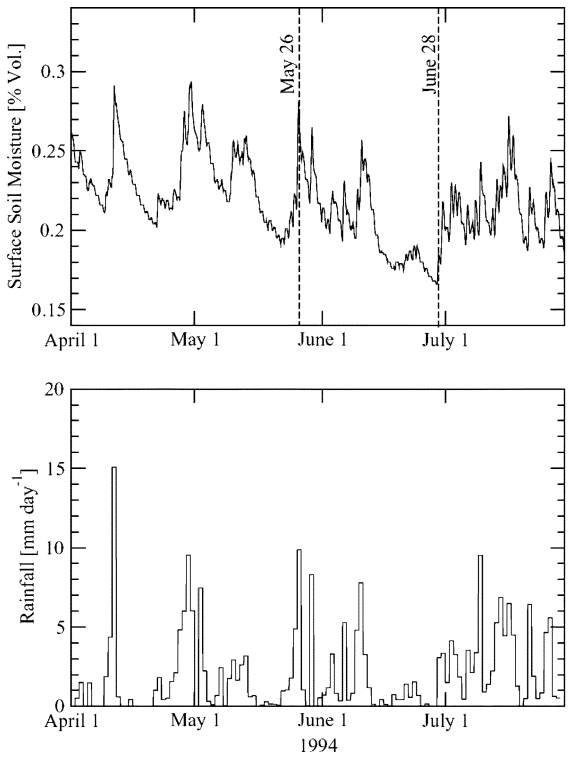


Fig. 3. Time series of areally averaged precipitation and LSM (TOPLATS) soil moisture simulations within the OSSE domain. Dashed vertical lines represent the starting and stopping times of the OSSE simulation.

The microwave emission model is based on a layered single-scattering soil–vegetation model commonly used for passive microwave sensing at L-band [1]. The modeled brightness temperature T_{Bp} includes components from the soil and vegetation and is expressed as

$$T_{Bp} = T_s e_p \exp\left(-\frac{\tau_p}{\cos\theta}\right) + T_c(1-\omega_p) \times \left[1-\exp\left(-\frac{\tau_p}{\cos\theta}\right)\right]\left[1+r_p\exp\left(-\frac{\tau_p}{\cos\theta}\right)\right]. \quad (1)$$

The subscript p refers to polarization (v or h), T_s is the soil effective temperature, T_c is the vegetation temperature, τ_p is the nadir vegetation opacity, ω_p is the vegetation single scattering albedo, and r_p is the soil reflectivity. The reflectivity is related to the emissivity by $e_p = (1-r_p)$, and ω_p, r_p, and e_p are values at the Hydros look angle of $\theta = 40°$. Equation (1) assumes that vegetation multiple scattering and reflection at the vegetation–air interface are negligible. Specular reflection is assumed at the soil surface, but with reflectivity modified for surface roughness. The parameters τ_p and ω_p represent the composite vegetation volume (leaves, stalks, branches, trunks) modeled as a homogeneous medium. Nadir vegetation opacity is related to the total columnar vegetation water content $W(\mathrm{kg}\cdot\mathrm{m}^{-2})$ by

$$\tau_p = b_p W. \quad (2)$$

The coefficient b_p depends on vegetation type [2], [15]. Due to the predominantly vertical orientation of vegetation stem and

trunk structure, b_v tends to be larger than b_h for most vegetation types [16]. The vegetation water content W is the average value over the 1-km pixel since no attempt is made to model fractional vegetation cover within a 1-km pixel. In addition, the polarization dependence of ω_p is ignored.

Roughness influences the sensor response through the root mean squared (RMS) surface height s, the horizontal correlation length l, and the roughness spectrum. The roughness effects can be computed using a theoretical model such as the integral equation model [17] or empirically using the $h-Q$ model [18]. At L-band the polarization mixing parameter Q is small, and the roughness effect using this model can be approximated as

$$r_p = r_{sp}\exp(-h). \tag{3}$$

The parameter h is assumed to be linearly related to the RMS surface height s, and r_{sp} is the reflectivity of the equivalent smooth soil surface, determined by the soil dielectric constant ε using the Fresnel equations.

At the 6 A.M. overpass time, the vegetation temperature T_c is assumed to be equal to the soil surface skin temperature T_0. The soil microwave effective temperature T_s is estimated as the weighted average of the surface skin temperature and the 5-cm temperature T_5. The weighting is actually a function of soil moisture, since the penetration depth varies with moisture, but, for simplicity, the direct average is used here

$$T_s = \frac{(T_0 + T_5)}{2}. \tag{4}$$

For a water surface the observed brightness temperature is

$$T_{Bp} = T_0(1 - r_p) \tag{5}$$

where r_p is the Fresnel smooth surface reflectivity for fresh water at temperature T_0 [19]. Wind-induced surface roughness is neglected.

In general, the key assumptions implicit in (1)–(5) (e.g., the neglect of multiple vegetation scattering and reflection, the neglect of polarization differences in ω, the linear relationship between τ_p and W, and the linear relationship between h and s) have been supported by airborne and tower campaigns over grass, shrubland, and agricultural crop land cover types [1], [2], [20]. However, much less evidence is available to support their suitability over forested regions. Consequently, predictions made by the MEM are likely more accurate over lightly vegetated western and central portions of the Red-Arkansas River Basin than over forested areas along the eastern edge of the OSSE domain.

A. Soil Dielectric Mixing Model

The real part of the soil dielectric constant, ε, and the Fresnel smooth surface reflectivities, r_{sv} and r_{sh}, are computed as functions of soil moisture for the distribution of soil textures (sand and clay fractions, S and C) found over the basin. The computation is done using the empirical relations of Dobson *et al.* [21] based on the data of Hallikainen *et al.* [22]. Parameters required in these models include the soil bulk density and the specific density of solid soil particles. A value of 2.66 g $\cdot$ cm^{-3} was assumed for soil specific density. Bulk density values were calculated from specific density and porosity values estimated from S and C using the regression relationship presented in [23].

B. Parameterization and Dataset Generation

Due to large uncertainties in the correct field-scale emission parameterization for many land surface types, it is difficult to unambiguously specify vegetation and surface parameters required by (1). While emission parameters are selected based on best available published information and constitutive relationships, some level of uncertainty and error is inevitable. However, it is important to note that the purpose of this study is not to perfectly mimic actual land surface emission for the study period, but rather to investigate—as accurately as possible—key aspects of error propagation in Hydros soil moisture retrievals given our current level of understanding concerning large-scale patterns in land surface emission parameters. Therefore, a reasonable level of error in the parameters described in this section is both expected and tolerable.

Vegetation and surface roughness parameters assigned to land cover types are listed in Table II. Surface roughness parameters s and h are assigned to the different land cover classes in the basin using values for categories of bare soils, crops, and grassland given in [20]. A factor of 0.1 (cm^{-1}) provides an approximate scaling between s (cm) and h at L-band [20], [24]. Listed values are intended to be representative of mean magnitudes for the different land cover classes. For example, cropped surfaces are assigned higher mean roughness than pasture or uncultivated surfaces. Due to a lack of reliable published information concerning s values for trees, they are set to a baseline value of 1 cm. Nonpolarized values of the vegetation parameters ω and b are also assigned based on land cover types, using representative values for herbaceous and woody vegetation listed in [1]. Based loosely on polarization results presented in [16], values for vertically polarized (horizontally polarized) b-parameters are estimated by increasing (decreasing) the baseline nonpolarized b values by 10% for grasses/crop/shrub landcovers and 20% for forested areas (Table II). The canopy vegetation water content W_c is obtained from the NDVI dataset using the relationship [20]

$$W_c = -0.3215(\mathrm{NDVI}) + 1.9134(\mathrm{NDVI})^2. \tag{6}$$

This relationship was derived for grassland/crop conditions using NDVI observations derived from 30-m Landsat Thematic Mapper data over south-central Oklahoma. Here we extrapolate its use to all land cover types using 1-km AVHRR observation of the Red-Arkansas River basin between May 13 and July 7, 1995. While this is clearly stretching (6) somewhat beyond its original formulation, the resulting 1-km W_c fields represent the best available representation of spatial patterns of W_c variability required for the analysis. For NDVI < 0.168, (6) predicts W_c values less than zero, in which case W_c is set to zero. Since the NDVI is sensitive to vegetation greenness, its signal comes primarily from the foliar canopy and not the woody components of the vegetation. In contrast, the microwave signals include contributions from the water content of the trunks and branches

as well as the leaves. Hence, for the simulations, a woody component fraction f_T is used to scale the foliar water content derived from (6) to a total columnar water content, W

$$W = \frac{W_c}{(1 - f_T)}. \tag{7}$$

The values assumed for f_T are shown in Table II. They are rough intuitive estimates based on basic tree and shrub physiology and not derived from any published source.

Simulated brightness temperatures T_{Bv} and T_{Bh} were computed from (1)–(7) using the h, ω, b (or b_v and b_h), and f_T parameters in Table II for each vegetation class. Cases for unpolarized and polarized b values were run separately to examine the effects of polarization and uncertainty in b values on the retrievals. To examine the impact of larger amounts of vegetation, the microwave computations of T_{Bv} and T_{Bh} were carried out for two levels of vegetation water content, i.e., $1W$ and $3W$, where W is the value obtained from (7).

VI. Orbit and Sensor Modeling

A. Aggregation

To facilitate Hydros data processing, radiometer data will be registered to earth-fixed grids of dimensions appropriate to the sensor footprint resolutions and sampling. For the radiometer simulations we define two nested grids of dimensions 1 and 36 km. The 1-km grid represents the resolution of the underlying geophysical fields, and the 36-km grid (a multiple of three for later nesting of 3-km radar measurements) represents the ~40-km radiometer footprint resolution; 36-km radiometer observations were computed by linear averaging of the simulated 1-km T_{Bv} and T_{Bh} in each 36-km grid cell. Precise Hydros orbital sampling patterns were not simulated. Instead a Hydros overpass covering the simulation domain was assumed every day at 6 A.M. and 6 P.M. local times. This approach is sufficient to illustrate the main features of resolution degradation, instrument measurement error (discussed in the next section), and soil moisture algorithms. In this paper, only the results from the 6 A.M. simulations are discussed. The LSM outputs (T_0, T_5, m_v) and surface characteristics parameters (W, S, C, h, ω, b_v, b_h) were also linearly averaged from 1- to 36-km grid cells. Since values for m_v, S, and C are undefined for water surfaces, inland water pixels were neglected when aggregating these parameters to 36 km.

B. Measurement Error

The radiometer measurement relative error includes instrument noise (that limits the measurement precision) and calibration relative error. The measurement precision (i.e., ΔT for the radiometer) is inversely proportional to the square root of the bandwidth and the sample integration time. For the Hydros orbit altitude and antenna size, mapping coverage at the surface requires a reflector rotation rate of 14.6 rpm and radiometer integration time of 42 ms. The integration time is effectively doubled if both fore- and aft-looking measurements of the same area are combined. The radiometer bandwidth of 25 MHz and system noise temperature of 590 K result in a single-look precision of 0.58 K, and combined fore and aft (two looks) precision of 0.4 K. The relative calibration error, including antenna effects, is estimated as 0.5 K. The root-sum square of the calibration stability and the two-look measurement precision results in a total measurement relative error of 0.64 K (Table I). Accounting also for the reduction in noise by averaging the samples into 36-km grid cells, a reasonable range of relative error to consider is 1 K (high case) to 0.5 K (low case). The high case allows for some additional uncertainty in the calibration stability of the mesh antenna. To simulate the measurement error, spatially independent Gaussian noise with a standard deviation of 1 K was added to the 36-km brightness temperature retrievals. This was done independently for both radiometer polarizations and for each day of the simulation.

C. Parameter Error

In operational settings, large uncertainty may exist in footprint-scale parameters used to invert observations of brightness temperature into soil moisture estimates. The impact of this uncertainty was addressed by adding synthetic noise to the 36-km aggregated fields of b_v and b_h used in the retrievals. To simulate retrieval parameter uncertainty, Gaussian error with a standard deviation of 0.02 was added after the 1-km b_v and b_h fields used in the original forward MEM modeling had been aggregated up to 36 km. For simplicity of analysis, the same perturbation was applied to both b_v and b_h for a given pixel.

The radiometer retrievals also require knowledge of effective surface temperature and various ancillary land surface parameters. It is assumed that the 6 A.M. effective temperature (corresponding to the local time of the morning Hydros overpass) can be estimated from observations or operational forecast models to within a root mean square accuracy of 1.5 K. Hence, the computed and aggregated T_s data fields were perturbed with 1.5-K Gaussian random noise. Among other sources, this induced error is meant to capture uncertainties in the simplified model in (4) used to estimate T_s from the surface skin and soil temperatures.

VII. Soil Moisture Retrieval Algorithms

In the next step of the simulation, the synthetic 36-km T_B observations described above were inverted into soil moisture products using a number of different retrieval models. The intent was to examine the sensitivities of the different algorithms to specific errors and assumptions. Three models using the radiometer data to produce a 36-km soil moisture product, designated "A", "B", and "C", were used. Retrievals were performed using the realistic vegetation water content (W) fields in the Red-Arkansas Basin (Case $1W$) as well as the synthetic case of artificially tripling the vegetation density (Case $3W$). The three algorithms differ primarily in the amount of ancillary soil and vegetation information required to retrieve soil moisture, with algorithm A requiring the most ancillary information and algorithm C the least.

A. Single-Polarization Algorithm

The first retrieval algorithm (A) is based on the algorithm introduced in [25]. It uses a single-channel T_{Bh} measurement and neglects differences between canopy and soil temperature by as-

suming $T_c = T_s$ in (1). Given known ancillary values of h, ω, T_s, b_h, W, and θ, this assumption allows (1) to be solved for r_{sh} based on a single-channel T_{Bh} observation (here synthetically derived using the MEM). Footprint-scale (36-km) values of h, ω, W, S, and C are obtained by averaging values in Table II according to the aggregation rules presented in Section VI-A. As discussed in Section VI-C, noisy values of T_s and b_h are used in the retrieval. Reflectivity retrievals are converted into soil moisture via the Fresnel equations and the Dobson soil dielectic mixing model [21].

B. Dual-Polarization Iterative Algorithm

The second algorithm (B) is based on the retrieval strategy introduced by [26] and utilizes both T_{Bh} and T_{Bv} measurements (synthetically generated using the MEM approach in Section V). The soil moisture m_v and vegetation water content W are given initial estimates that are adjusted iteratively in computations of T_{Bh} and T_{Bv} until the difference between the computed (using (1) with $T_c = T_s$ and the Dobson soil dielectric mixing model [21]) and observed dual-polarization brightness temperatures is minimized in a least squares sense. Experience with the approach has demonstrated that such an optimization is generally well-posed and exhibits little sensitivity to the values of m_v and W used to initialize the iterative search. No knowledge of W is required since both m_v and W are simultaneously retrieved. Footprint-scale (36-km) values of h, ω, S, and C used in the retrieval are obtained by averaging values in Table II according to the aggregation rules presented in Section VI-A. As discussed in Section VI-C, noisy values of T_s, b_h, and b_h are used in the retrieval.

C. Multipolarization Regression Algorithm

The third algorithm (C) is based on utilizing both T_{Bh} and T_{Bv} measurements to estimate a combined roughness and vegetation correction factor which is then used to calculate an H-polarization bare-soil reflectivity and, eventually, soil moisture. This algorithm was developed to investigate the extent to which soil moisture could potentially be estimated without the use of roughness and vegetation ancillary data, although, in practice, regression algorithms must be used with caution since they are tuned to specific datasets.

The approach can be illustrated by writing (1) in a more detailed form that separates out contributions from the vegetated (f_c) and bare soil $(1 - f_c)$ fractions of the footprint

$$T_{Bp} = f_c \left\{ T_s e_p \exp\left(-\frac{\tau_p}{\cos\theta}\right) + T_c(1-\omega_p)\left[1-\exp\left(-\frac{\tau_p}{\cos\theta}\right)\right] \times \left[1 + r_p \exp\left(-\frac{\tau_p}{\cos\theta}\right)\right]\right\} + (1 - f_c) T_s e_p. \quad (8)$$

The parameters ω_p and τ_p now refer to just the vegetated portion of the footprints, rather than to averaged values over the entire footprint as in (1). Assuming $T_c = T_s$ and inserting (3), (8) can be rewritten as

$$1 - VC_p - e_p^t = r_{sp} CF_p \quad (9)$$

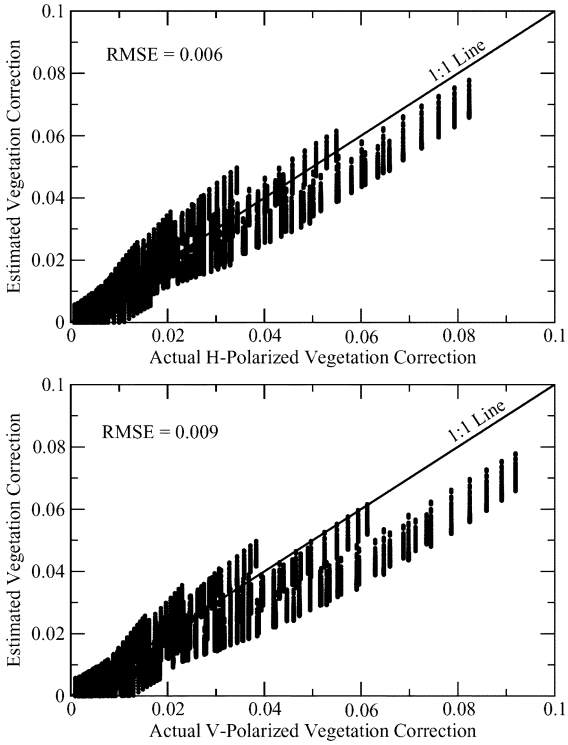


Fig. 4. Actual versus calculated, via (13), vegetation correction factors (VC).

where CF_p is a correction parameter for the combined effect of the surface roughness and vegetation and defined as

$$CF_p = \exp(-h)\left[1 - f_c + f_c \exp\left(-\frac{\tau_p}{\cos\theta}\right) \times \left(1 - (1-\omega_p)\left[1 - \exp\left(-\frac{\tau_p}{\cos\theta}\right)\right]\right)\right]. \quad (10)$$

VC_p is a vegetation correction factor defined as

$$VC_p = f_c \omega_p \left[1 - \exp\left(-\frac{\tau_p}{\cos\theta}\right)\right] \quad (11)$$

and the total emissivity e_p^t is

$$e_p^t = \frac{T_{Bp}}{T_s}. \quad (12)$$

The retrieval approach is based on combining measurements of e_p^t with regression models for CF_p and VC_p to solve for r_{sp} via (9).

Neglecting polarization differences in τ and ω, a set of emission signals were simulated using (8) with f_c varied between 0.2 and 1, τ between 0.067 and 1, and ω between 0.03 and 0.18. Using this simulated set, a regression analysis was performed to derive the best functional relationship between VC (now assumed to be polarization independent) and multipolarized total emissivity

$$VC \approx -5.656 - 2.377 \cdot \left(e_v^t - e_h^t\right) + 6.601 \cdot \frac{e_v^t}{e_h^t} - 16.658 \cdot \sqrt{e_h^t} + 15.749 \cdot \sqrt{e_v^t}. \quad (13)$$

Fig. 4 plots scatterplots of VC values estimated using (13) and true VC_h and VC_v values calculated using (11). By using only

a single VC value from (13), small polarization differences between VC_h and VC_v (and thus between b_v and b_h) are implicitly neglected in this stage of the retrieval.

In addition, using the Fresnel equations, a set of simulated r_{sv} and r_{sh} values were derived at the 40° Hydros look angle using volumetric soil moisture values between 2% and 50%, and the power law relationship

$$r_{sh} = 1.013\, r_{sv}^{0.598} \tag{14}$$

was fit to this simulated data. Values of r_{sv} estimated via (14) have an RMSE of 0.000 64. By combining (9) and (14), the effect of surface dielectric properties in the measurements is canceled out and a correction index factor CF_{index} can be defined as

$$CF_{\text{index}} = 0.987 \cdot \frac{1 - VC - e_h^t}{\left(1 - VC - e_v^t\right)^{0.598}} = \frac{CF_h}{CF_v^{0.598}}. \tag{15}$$

Since the forward MEM simulation is based on $b_v > b_h$ with different b_v and b_h values for various landcover types, a basin-wide average value of 1.25 is assumed for the ratio b_v/b_h. Based on this assumption, it was found that CF_h can be estimated by

$$CF_h = \exp\left(5.73 + 8.11 \log(CF_{\text{index}}) - 5.37\, CF_{\text{index}}\right). \tag{16}$$

The expression (16) works well for small to moderate optical thickness conditions with $\exp(-\tau_p/\cos\theta) > 0.5$; however, estimation error increases for $\exp(-\tau_p/\cos\theta) < 0.5$. Overall, the absolute RMSE for CF_h values estimated via the regression fit in (16) is 0.029.

For the inversion, e_h^t and e_v^t are first derived using TB_v and TB_h observations combined with ancillary T_s data. Then, (13) is used to estimate VC, (15) is used to estimate CF_{index}, and (16) is used to estimate CF_h. Finally, r_{sh} is estimated using (9) and converted into soil moisture using the Fresnel equations and Dobson mixing model [21]. As mentioned earlier, this algorithm requires no ancillary vegetation or surface roughness information.

For all three retrieval approaches, it is important to note that a number of key retrieval assumptions (e.g., neglecting polarization impacts on ω, the linear relationship between W and τ_p, the linear relationship between h and s, and neglecting multiple vegetation scattering and reflection) are common to both the forward and retrieval models. Consequently, their impact on retrieval errors is not captured. Instead, within this particular OSSE, retrieval errors arise primarily from four sources: aggregation impacts associated with the loss of sub-footprint-scale land surface heterogeneity (Section VI-A), observation error added to T_{Bh}, T_{Bv}, and T_s (Section VI-B), parameter error added to b_v and b_h (Section VI-C), and assumptions employed in retrieval modeling not utilized in the forward MEM portion of the OSSE (i.e., $T_c = T_s$ in algorithms A, B, and C and $b_v = b_h$ in the regression (14) for algorithm C). In addition, the accuracy of the regression expressions employed by algorithm C (e.g., see Fig. 4) impacts its retrieval accuracy.

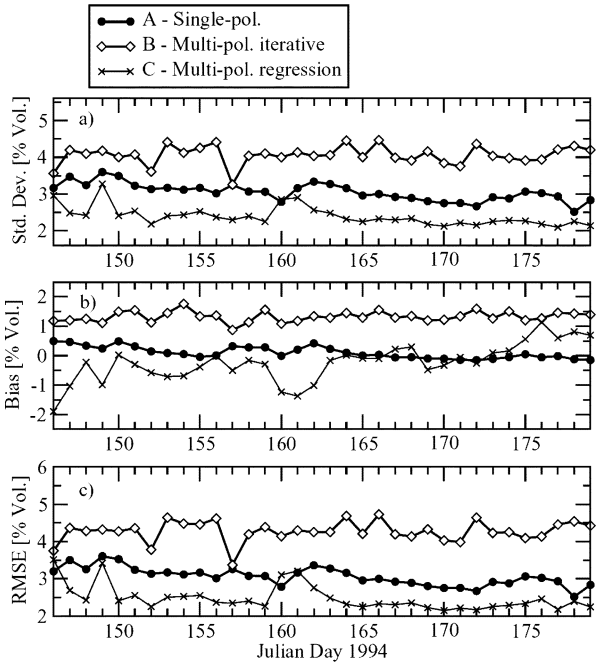


Fig. 5. Time series of (a) mean and (b) standard deviation of retrieved minus benchmark 36-km soil moisture values. Also shown are (c) total RMSE for 36-km soil moisture retrievals.

VIII. Retrieval Results and Error Sources

Fig. 1 illustrates the OSSE procedure using simulated 1-km soil moisture images generated by TOPLATS (Section IV). This imagery is used to create 1-km T_B images (Section V), which are first aggregated and then randomly perturbed with added instrument noise (Section VI) prior to their insertion into a retrieval algorithm (Section VII). Time series for the mean and standard deviation of differences between 36-km soil moisture retrievals and benchmark 36-km soil moisture fields for the $1W$ case are shown in Fig. 5(a) and (b). Benchmark fields are derived via direct aggregation of the original 1-km soil moisture field produced by the LSM, and differences are defined as retrieved minus benchmark values. The mean (bias) values shown in Fig. 5(b) are the overall basin-wide biases in the soil moisture retrievals. The standard deviation values plotted in Fig. 5(a) are equivalent to the bias-removed RMS error (RMSE) for individual 36-km soil moisture retrievals within the basin, as in (17), shown at the bottom of the page, where n_{total} is the total number of 36-km pixels in the OSSE domain for a given day (i.e., a single Hydros overpass). Fig. 5(c) shows total RMSE

$$\begin{aligned}\text{RMSE} &= \sqrt{n_{\text{total}}^{-1} \sum (\text{Retrieved} - \text{Benchmark})^2} \\ &= \sqrt{\text{Bias}^2 + \text{Std. Dev.}^2}\end{aligned} \tag{18}$$

derived by pooling results from all 36-km pixels in the OSSE domain. Total soil moisture RMSE for all three algorithms is constrained within a range of between 2.0% and 4.5% volumetric

$$\text{Std. Dev.} = \sqrt{n_{\text{total}}^{-1} \sum (\text{Retrieved} - \text{Benchmark})^2 - n_{\text{total}}^{-2} \left(\sum \text{Retrieved} - \sum \text{Benchmark}\right)^2} \tag{17}$$

for all days. The multipolarization regression algorithm (C) has the lowest retrieval standard deviation and RMSE followed by the single-polarization algorithm (A). The superior performance of algorithm A relative to the multipolarization iterative approach (B) likely reflects the fact the algorithm A is given access to vegetation water content information (W) information that is withheld from algorithm B. Retrieval biases are generally small for all algorithms, although algorithm B does exhibit a persistent positive bias of approximately 1% volumetric and algorithm C exhibits a temporally varying bias that is negative for early (wet) portions of the OSSE and gradually becomes more positive as the OSSE domain dries out toward the end of June (i.e., after Julian day 165). However, with the exception of bias results for algorithm C, relatively little temporal variability is observed in soil moisture retrieval error statistics despite the presence of significant temporal variability in overall soil moisture conditions within the basin (see Fig. 3).

Rerunning the OSSE for the case of no b polarization effects (i.e., $b_v = b_h = b$ in both the forward MEM and retrieval algorithms) and no added b noise, reduces absolute RMSE errors in Fig. 5 for algorithm B by about 1% to 2% volumetric and algorithm A by 0.5% to 1% volumetric. The smaller impact for the single-polarization algorithm A occurs because the positive impact of removing b_h noise on retrieval accuracy is partially offset by the effective increase in vegetation optical depth when moving from b_h to b (i.e., $Wb_h < Wb$; see Table II).

A. Impact of Vegetation

Spatial patterns in retrieval errors appear to be driven primarily by the distribution of vegetation within the OSSE domain. For the original case of polarized dependency and uncertainty in b, Fig. 6 plots retrieved soil moisture RMSE [Fig. 6(c)] and bias [Fig. 6(b)] stratified by the mean vegetation water content (W) contained within each 36-km pixel. Statistics are pooled values for all 36-km pixels in each scene and all days of the OSSE. Despite relatively low errors for the domain as a whole (Fig. 5), significant errors are present in algorithm A and B soil moisture retrievals for the relatively small number of 36-km pixels where $W > 1.5$ kg $\cdot$ m^{-2}. A large fraction of this error is associated with a positive bias in retrievals for heavily vegetated pixels [Fig. 6(b)]. Unfortunately, there are too few densely vegetated pixels within the study area to study this retrieval aspect in detail for the baseline $1W$ case [Fig. 6(d)]. To remedy this, the W values in the basin were artificially scaled up by a value of three and the OSSE repeated in whole. Depending on the amount of vegetation present, this artificial scaling of W increases 36-km T_{Bh} values by between 0–30 K and 36-km T_{Bv} values between 0–20 K. Soil moisture retrieval results for this artificial $3W$ case are presented in Fig. 7. Substantial errors in A and B retrievals are noted for W values exceeding 3 kg $\cdot$ m^{-2} and in C retrievals for W values exceeding 4 kg $\cdot$ m^{-2}. As in Fig. 6, a significant fraction of this error is due to the impact of positively biased soil moisture retrievals at high W. The retrieval bias is dependent on the landscape heterogeneity within the radiometer footprint caused by inland water and vegetation, as discussed below.

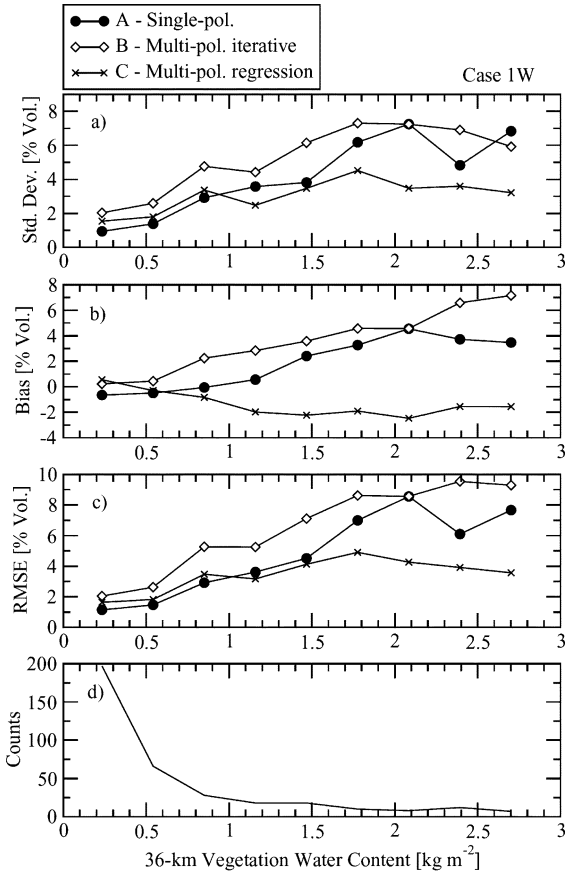


Fig. 6. For the $1W$ case, plots of retrieved 36-km soil moisture (a) standard deviation, (b) bias, and (c) RMSE stratified by 36-km vegetation water content values and (d) the histogram for 36-km vegetation water content values within the OSSE domain.

B. Impact of Inland Water

Within the Red-Arkansas river basin, heavily vegetated areas of the basin exhibit the greatest amount of spatial variability in canopy opacity due to mixed grassland/forest regions and a relatively high areal fraction of inland water in the densely vegetated eastern portion of the basin. In particular, emission from inland water can confound soil moisture retrievals by contaminating radiometer observations with surface signals containing no soil moisture information [8]. In the retrieval algorithms discussed in Section VII, soil moisture estimates were derived assuming that all observed microwave emission originates from land. The impact of neglected inland water on soil moisture estimates is illustrated by repeating Fig. 7(b) for the case of screening all 36-km pixels containing any 1-km pixels classified as inland water. Fig. 8 illustrates the impact of such stringent masking by plotting soil moisture retrieval biases for the A, B, and C algorithms (Case $3W$) stratified by 36-km W values. Screening 36-km pixels with at least some inland water substantially reduces positive biases in algorithm B retrievals for 36-km pixels with W greater than 2 kg$\cdot$m^{-2}. Although less dramatic and limited only to W values below about 5 kg$\cdot$m^{-2}, improvements in A and C retrievals are also realized. Operationally, such stringent screening is unappealing since it prevents the retrieval of soil moisture in approximately one-third of the 36-km pixels in the basin. An inland-water correction approach suitable for

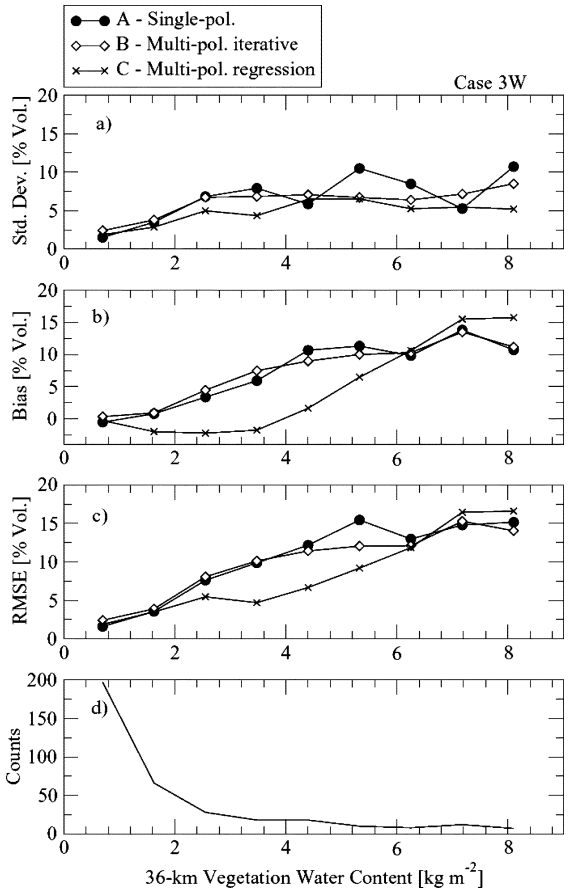


Fig. 7. For the $3W$ case, plots of retrieved 36-km soil moisture (a) standard deviation, (b) bias, and (c) RMSE stratified by 36-km vegetation water content values and (d) the histogram for 36-km vegetation water content values within the OSSE domain.

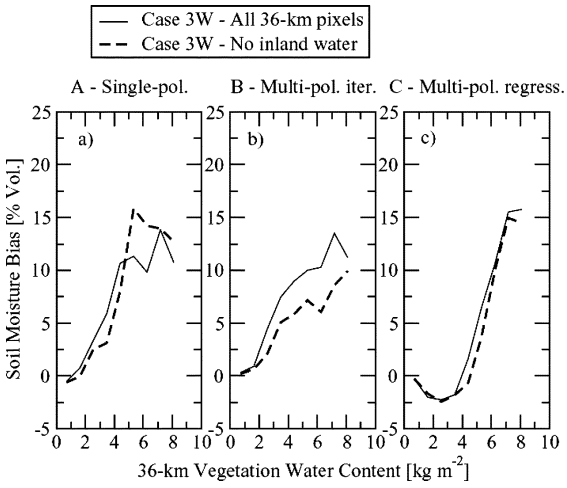


Fig. 8. For all three retrieval algorithms, soil moisture retrieval biases, stratified by 36-km W levels, for all 36-km pixels (solid line) and only those 36-km pixels without any inland water (dashed line).

operational soil moisture retrievals will be presented later (see Section IX-A).

C. Impact of Land Surface Heterogeneity

Large residual biases in Fig. 8(a) remain in algorithm A soil moisture retrievals even when inland water is completely

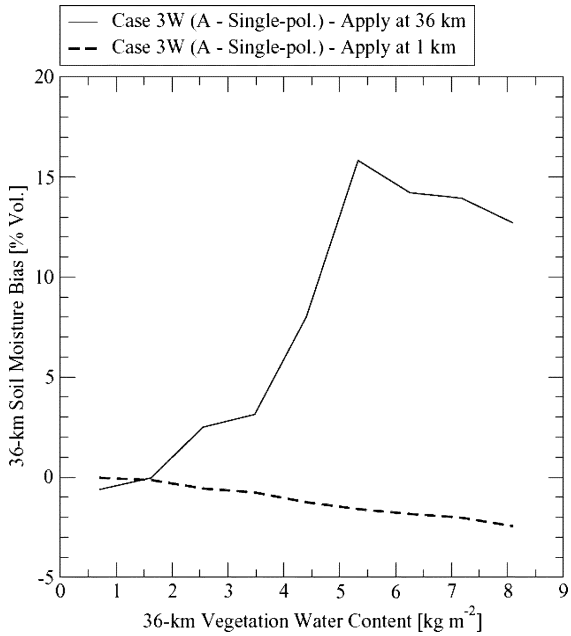


Fig. 9. Difference in soil moisture retrieval biases when applying algorithm A to 1-km (dashed line) and 36-km (solid line) geophysical fields.

screened. Sensitivity results (not shown) demonstrate that neither the inclusion of mean-zero noise in T_s and T_B observations nor the assumption that $T_s = T_c$ contributes significantly to this positive bias. Instead, the bias appears due to aggregation effects associated with the poor spatial resolution of radiometer (36-km) observations relative to the baseline resolution of the LSM simulations underlying the OSSE (1-km). These aggregation impacts arise from the nonlinearity of the soil moisture retrieval process with respect to various land surface parameters and the inability of a spaceborne radiometer to capture the net impact of sub-footprint-scale correlations between T_B and land surface parameters. Fig. 9 quantifies these impacts on algorithm A retrievals by comparing bias results for the application of algorithm A to simulated 1-km T_s and T_B measurements (with 1.5 and 1 K added Gaussian noise, respectively) and the subsequent aggregation of 1-km soil moisture retrievals to 36-km, versus results for the baseline OSSE approach of applying retrieval algorithms to simulated 36-km T_B observations and aggregated fields of ancillary data. The large positive bias observed at high W in Figs. 7(b) and 8(a) is eliminated by application of the retrieval approach at a finer spatial resolution (1 versus 36 km).

For $W < 6\,\mathrm{kg \cdot m^{-2}}$, it is telling that aggregation impacts appear largest in algorithm A results (see Figs. 8 and 9) relative to other approaches since it is the only algorithm to rely on directly averaged W values to perform 36-km soil moisture retrievals. Fig. 10(a) plots the bias in 36-km soil moisture retrievals versus the 1-km variability in W within each 36-km pixel for algorithms A and B. Large biases in algorithm A retrievals, and large differences between A and B results, occur primarily in pixels exhibiting high W variability. Differences between algorithms A and B stem solely from the manner in which 36-km W values are calculated. Fig. 10(b) plots biases (calculated relative to the direct aggregation of 1-km W values) in 36-km W values used by both algorithms. Algorithm A is forced to use a direct linear

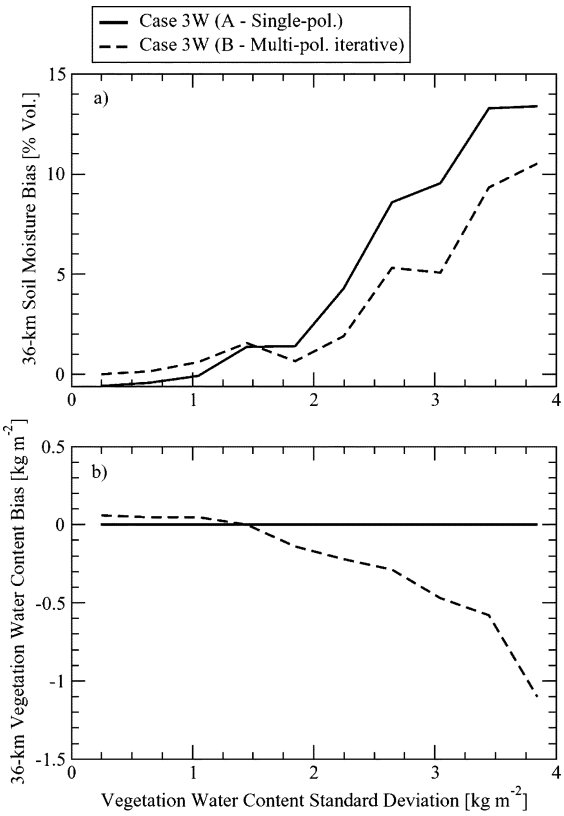


Fig. 10. For algorithms A (solid line) and B (dashed line), retrieval biases in (a) soil moisture and (b) W stratified by levels of 1-km W variability within each 36-km pixel.

average of 1-km W values. In contrast, algorithm B is allowed to solve for effective 36-km W values based on optimal fitting to 36-km T_{Bh} and T_{Bv} observations. For pixels with high W variability, algorithm B retrieves 36-km W values that are consistently biased low relative to those derived from direct aggregation of 1-km W fields. However, using lower W values means that algorithm B calculates lower surface reflectivities, and soil moistures, relative to algorithm A. This partially corrects algorithm B results for the large positive bias seen in A for highly variable pixels. That is, the iterative B algorithm solves for effective values of 36-km W that offer at least a partial correction to biases arising from aggregation impacts. These impacts are further examined in Section IX-B.

IX. Operational Correction Strategies

Results in Figs. 7–10 suggest that some type of correction strategy will be required for the dual impact of inland water and vegetation variability to guarantee adequate accuracy for soil moisture retrievals in heterogeneous and/or heavily vegetated pixels (i.e., $W > 3\,\mathrm{kg \cdot m^{-2}}$). The potential of two operational strategies is examined here.

A. Inland Water Correction

The impact of inland water can be corrected by modifying T_B observations for the impact of inland water and recalculating

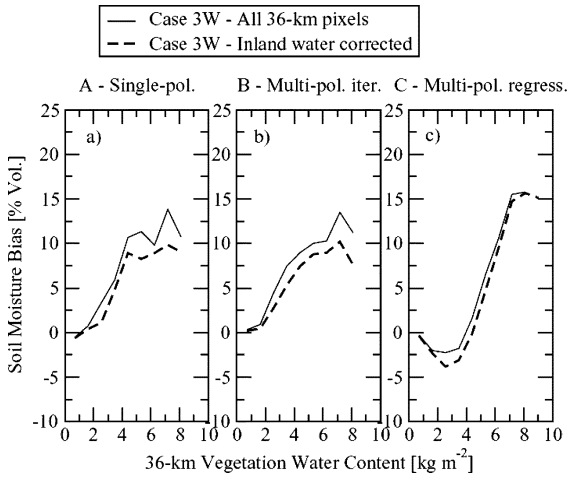


Fig. 11. For all three retrieval algorithms, soil moisture retrieval biases, stratified by 36-km W levels, for the baseline retrieval case (solid line) and the inland water correction strategy outlined in (20) (dashed line).

soil moisture. Actual 36-km T_B observations are based on a linear average of emission from both water and land surfaces

$$T_B = f_w T_{B,\mathrm{water}} + (1 - f_w) T_{B,\mathrm{land}} \tag{19}$$

where f_w is the fraction of the footprint covered by water. This fraction can be estimated from either high-resolution visible/infrared imagery or high-resolution (3-km) mapping of the footprint using the active radar capability of the Hydros sensor. Here a separate f_w is determined for each 36-km pixel in the OSSE domain based on the 1-km land cover classification described in Section IV. Values for f_w within 36-km pixels in the domain range between zero to near 0.40. Given knowledge of f_w and (19), emission from the land surface can be estimated as

$$T_{B,\mathrm{land}} = \frac{(T_B - f_w T_{B,\mathrm{water}})}{(1 - f_w)}. \tag{20}$$

Assuming f_w and T_0 are available from ancillary sources, V- and H-polarized $T_{B,\mathrm{land}}$ can be estimated via (5) and (20) and used in lieu of T_B observations to derive 36-km soil moisture products. Doing so leads to reductions in soil moisture biases for all three retrieval algorithms (Fig. 11). These results are similar to results based on stringent screening of inland water in Fig. 8, except this approach has the large added benefit of not reducing the areal extent of soil moisture retrievals.

B. Alternative Aggregation Techniques

Even after correction for inland water, soil moisture retrieval errors remain significant for algorithm A in the case of artificially enhanced W levels (case $3W$). Residual error is due to aggregation impacts associated with the retrieval procedure. For instance, neglecting the single-scattering albedo, reflectivity retrievals for algorithm A simplify to

$$r_{sh} = \left(1 - \frac{T_{Bh}}{T_s}\right) \exp(2 b_h W \sec\theta + h). \tag{21}$$

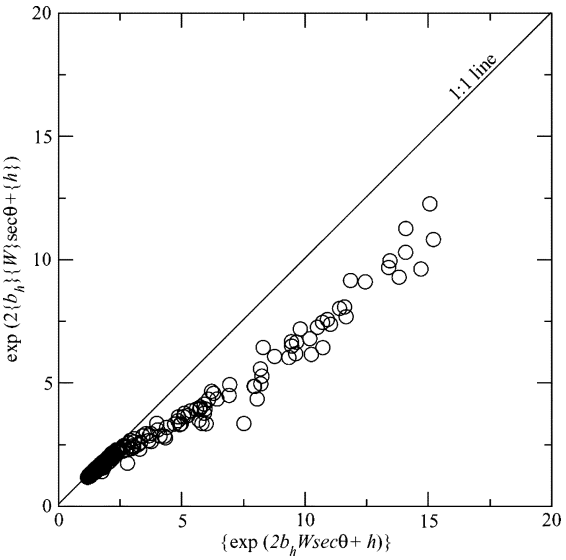


Fig. 12. Impact of alternative aggregation for soil (h) and vegetation (W and b_h) parameters on the calculation of 36-km $\exp(2b_h W \sec\theta + h)$.

Applying (21) at 1 km (i.e., the resolution of vegetation and LSM products) and aggregating results to 36 km yields the benchmark case

$$r_{sh,36\text{ km}} = \left\{ \left(1 - \frac{T_{Bh}}{T_s}\right) \exp(2b_h W \sec\theta + h) \right\} \quad (22)$$

or

$$r_{sh,36\text{ km}} = \left\{1 - \frac{T_{Bh}}{T_s}\right\} \{\exp(2b_h W \sec\theta + h)\} + \text{Cov}\left[1 - \frac{T_{Bh}}{T_s}, \exp(2b_h W \sec\theta + h)\right] \quad (23)$$

where brackets { } indicate spatial averaging. In contrast, first averaging and then applying algorithm A (as in Section VII-A) yields

$$r'_{sh,36\text{ km}} = \left(1 - \frac{\{T_{Bh}\}}{\{T_s\}}\right) \exp\left(2\{b_h\}\{W\}\sec\theta + \{h\}\right) \quad (24)$$

or, neglecting T_s variability,

$$r'_{sh,36\text{ km}} = \left\{1 - \frac{T_{Bh}}{T_s}\right\} \exp\left(2\{b_h\}\{W\}\sec\theta + \{h\}\right). \quad (25)$$

Because soil moisture is (nearly) directly proportional to reflectivity, soil moisture bias results from Figs. 9 demonstrate that, on average, $r'_{sh,36\text{ km}} > r_{sh,36\text{ km}}$.

Intercomparison of (23) and (25) suggests that one possible way to correct the bias in $r'_{sh,36\text{ km}}$ retrievals is to utilize an alternative aggregation scheme whereby 1-km b, W, and h values are transformed via $\exp(2b_h W \sec\theta + h)$ and aggregated as a single quantity. However, due to the positive concavity of the exponential function, plotted results in Fig. 12 demonstrate a clear tendency for

$$\{\exp(2b_h W \sec\theta + h)\} > \exp\left(2\{b_h\}\{W\}\sec\theta + \{h\}\right) \quad (26)$$

within the OSSE domain. Consequently, utilizing an alternative aggregation scheme for b_h, W, and h in the calculation of $r'_{sh\ 36\text{ km}}$ via (25) actually increases the positive bias in reflectivity retrievals. Intercomparison of (23) and (25) in light of the inequality in (26) reveals that the covariance term in (23)—$\text{Cov}[1 - T_{Bh}/T_s, \exp(2b_h W \sec\theta + h)]$—must be negative and the dominant source of positive bias in retrievals utilizing (25) relative to (23). The covariance term represents the impact of sub-footprint-scale correlation between static land surface characteristics and the observed brightness temperature. Such correlation is imposed on the simulations by the 1-km forward microwave emission model and reflects the tendency for areas with high W (and/or h) to be associated with low observed reflectivity $(1 - T_{Bh}/T_s)$. Because it is partially based on unresolvable sub-footprint-scale variations in T_B, estimation of this term is difficult even if 1-km fields of W, h, and b_h are known. Within this OSSE, variability in W dominates overall variability in the term $2b_h W \sec\theta + h$. Consequently, the inability to resolve this sub-grid-scale correlation between this term and T_B appears to have the largest impact in cases of dense and/or highly variable vegetation cover.

X. Discussion and Conclusion

The observing system simulation experiment described here captures the influence of land surface heterogeneity, observation noise, inversion parameter uncertainty, and retrieval assumptions on the accuracy of radiometer-only Hydros soil moisture products. Examining these error sources in a controlled numerical setting provides an opportunity to assess eventual processing and retrieval strategies designed to mitigate their impact. Nevertheless, care should be taken when equating error results presented here to accuracy expectations for actual Hydros soil moisture products. This particular OSSE provides a simplified representation of only a partial set of error sources within actual retrievals. Particular choices concerning the nature of represented error may impact the relative accuracy of various retrieval algorithms. For instance, the decision to neglect error in W estimates almost certainly reduces error in algorithm A's soil moisture retrievals but has little or no effect on algorithm B and C's results, which do not explicitly consider vegetation water content. In addition, while useful as a test-bed to study strategies for treating aggregation-based retrieval errors (see Section IX), results for the case of scaled-up vegetation density $(3W)$ should be interpreted as a worst case scenario of land surface heterogeneity and vegetation density encountered over only limited portions of the globe. Hydros soil moisture algorithms will undergo continued evolution and refinement prior to Hydros launch, based on a combination of OSSE results, further analyses, and data from ongoing airborne field campaigns. Given the generally large contribution of retrieval biases to overall RMSE (e.g., see Fig. 7), it may be possible to improve the accuracy of retrieved soil moisture via calibration of retrieval model parameters.

In this analysis, soil moisture products were simulated for three different retrieval algorithms (Section VII), each requiring a different amount of ancillary soil and vegetation information.

Within the United States Southern Great Plains, the overall impact of captured error sources on the accuracy of 36-km soil moisture retrievals is between 2.0% and 4.5% volumetric for all three retrieval algorithms (Fig. 5). The impact, however, is more acute for heavily vegetated footprints (Fig. 6).

To study these areas in greater detail, the OSSE was repeated for the case of artificially scaled W values (by a factor of three). The $3W$ case represents a worst case scenario in which land surface heterogeneity—specifically vegetation optical depth variability—is synthetically increased to illuminate the impact of sub-footprint-scale heterogeneity on retrieval accuracy. Retrieval results for the artificially scaled $3W$ case demonstrate positive biases at high W levels [Fig. 7(b)]. Despite relatively low spatial density ($<1\%$ of the basin), failure to account for the impact of inland water on soil moisture estimates contributes significantly to this bias (Fig. 8). The remaining positive bias is due to aggregation impacts associated with applying retrieval algorithms at a footprint-scale (36-km) that is much coarser than the underlying 1-km resolution of the geophysical fields used to force the OSSE (Fig. 9). Measurement noise and simplifying retrieval assumptions appear to play a lesser relative role.

Aggregation impacts are particularly strong at high W when using algorithm A, which derives 36-km soil and vegetation values based on the direct aggregation of 1-km fields. For $W < 6\,\mathrm{kg\cdot m^{-2}}$, utilizing either regression (algorithm C) or the simultaneously fitting of parameters to T_{Bv} and T_{Bh} observations (algorithm B) to derive effective footprint-scale W appears to offer a partial correction for aggregation errors (Figs. 8). Algorithm B utilizes an alternative strategy and derives effective 36-km W values based on numerical fitting to T_{Bv} and T_{Bh} observations. Negative biases in effective W derived in this manner partially counteracts the prevailing positive bias in soil moisture retrievals (Fig. 10) and may represent a preferential strategy for aggregating heterogeneous W fields up to footprint-scale resolutions. Consequently, strategies that solve for effective W values based on multipolarization observations (e.g., algorithms B and C) may have some advantages for highly heterogeneous footprints. However, more information is needed to determine how robust these approaches, particularly the empirical vegetation correction strategy in algorithm C, will be in real operational environments.

The OSSE also clarifies the potential of two operational data processing strategies to reduce the impact of land surface aggregation errors. Given knowledge of fractional water coverage, the impact of nonresolved water bodies on soil moisture retrievals can be effectively filtered using (23) (Fig. 11). However, remaining biases in soil moisture retrievals from algorithm A are problematic in the sense that they cannot be corrected via simple alternative strategies for aggregating static soil and vegetation properties up to the footprint-scale (Section IX-B and Fig. 12). More study on alternative aggregation techniques is required.

Results from this experiment will be used to modify operational data processing and soil moisture retrieval strategies for Hydros soil moisture products—particularly in areas of dense vegetation and widespread inland water. This will be done with the eventual goal of maximizing the global extent over which sufficiently accurate soil moisture retrievals are possible. Given the operational availability of higher resolution Hydros radar measurements, future work will also focus on the potential for addressing landscape heterogeneity and aggregation effects using simultaneous backscatter observations.

REFERENCES

[1] Y. Kerr and J.-P. Wigneron, "Vegetation models and observations—a review," in *Passive Microwave Remote Sensing of Land-Atmosphere Interactions*, B. Choudhury, Y. Kerr, E. Njoku, and P. Pampaloni, Eds. Utrecht, The Netherlands: VSP, 1995.

[2] T. J. Jackson and T. J. Schmugge, "Vegetation effects on the microwave emission from soils," *Remote Sens. Environ.*, vol. 36, pp. 203–212, 1991.

[3] E. Njoku, S. Hook, and A. Chehbouni, "Effects of surface heterogeneity on thermal remote sensing of land parameters," in *Scaling-Up in Hydrology Using Remote Sensing*, J. B. Stewart, E. T. Engman, R. A. Feddes, and Y. Kerr, Eds. Chichester, U.K.: Wiley, 1996.

[4] T. J. Jackson, "Multiple resolution analysis of L-band brightness temperature for soil moisture," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 151–164, Jan. 2001.

[5] M. Drusch, E. F. Wood, and C. Simmer, "Up-scaling effects in passive microwave remote sensing: ESTAR 1.4 GHz measurements during SGP'97," *Geophys. Res. Lett.*, vol. 26, pp. 879–882, 1999.

[6] J. F. Galantowicz, D. Entekhabi, and E. Njoku, "Estimation of soil-type heteorogeneity effects in the retrieval of soil moisture from radiobrightness," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 1, pp. 312–316, Jan. 2000.

[7] W. T. Crow, M. Drusch, and E. F. Wood, "An observation system simulation experiment for the impact of land surface heterogeneity on AMSR-E soil moisture retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 8, pp. 1622–1631, Aug. 2001.

[8] T. Pellarin, J.-P. Wigneron, J.-C. Calvert, and P. Waldteufel, "Global soil moisture retrieval from a synthetic L-band brightness temperature data set," *J. Geophys. Res.*, vol. 108, no. D12, 2003. DOI: 10.1029/2002JD003086.

[9] D. Entekhabi, E. Njoku, P. Houser, M. Spencer, T. Doiron, J. Smith, R. Girard, S. Belair, W. Crow, T. Jackson, Y. Kerr, J. Kimball, R. Koster, K. McDonald, P. O'Neill, T. Pulz, S. Running, J. Shi, E. Wood, and J. van Zyl, "The Hydrosphere State (HYDROS) mission concept: An earth system pathfinder for global mapping of soil moisture and land freeze/thaw," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2184–2195, Oct. 2004.

[10] E. Njoku, D. Entekhabi, T. Doiron, R. Girard, P. Houser, K. McDonald, P. O'Neill, J. Smith, and M. Spencer, "The HYDROS mission: requirements and system design," presented at the *IEEE Aerospace Conf.*, Big Sky, MT, Mar. 2004.

[11] C. D. Peters-Lidard, M. S. Zion, and E. F. Wood, "A soil-vegetation-atmosphere transfer scheme for modeling spatially variable water and energy balance processes," *J. Geophys. Res.*, vol. 102, pp. 4303–4324, 1997.

[12] W. T. Crow and E. F. Wood, "The value of coarse-scale soil moisture observations for regional surface energy balance modeling," *J. Hydrometeorol.*, vol. 3, pp. 467–482, 2002.

[13] T. R. Loveland, J. W. Merchant, D. O. Ohlen, and J. F. Brown, "Development of a land-cover characteristics database for the conterminous U. S.," *Photogramm. Eng. Remote Sens.*, vol. 57, pp. 1453–1463, 1991.

[14] J. F. Brown, T. R. Loveland, J. W. Merchant, B. C. Reed, and D. O. Ohlen, "Using multisource data for global land cover characterization: Concepts, requirements, and methods," *Photogramm. Eng. Remote Sens.*, vol. 59, pp. 977–987, 1993.

[15] A. A. Van de Griend and J.-P. Wigneron, "The b-factor as a function of frequency and canopy type at H-polarization," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 1, pp. 1–10, Jan. 2004.

[16] J.-P. Wigneron, M. Parde, P. Waldteufel, A. Chanzy, Y. Kerr, S. Schmidl, and N. Skou, "Characterizing the dependence of vegetation model parameters on crop structure, incidence angle, and polarization at L-band," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 2, pp. 416–425, Feb. 2004.

[17] A. K. Fung, Z. Li, and K. S. Chen, "Backscattering from a randomly rough dielectric surface," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 2, pp. 356–369, Mar. 1992.

[18] J. R. Wang, "Passive microwave sensing of soil moisture content: The effects of soil bulk density and surface roughness," *Remote Sens. Environ.*, vol. 13, pp. 329–344, 1983.

[19] L. A. Klein and C. T. Swift, "An improved model for the dielectric constant of sea water," *IEEE Trans. Antennas Propagat.*, vol. AP-25, pp. 104–111, 1977.
[20] T. J. Jackson, D. M. Le Vine, A. Y. Hsu, A. Oldak, P. J. Starks, C. T. Swift, J. Isham, and M. Haken, "Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains Hydrology Experiment," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2136–2151, Sep. 1999.
[21] M. C. Dobson, F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes, "Microwave dielectric behavior of wet soil—Part II: Dielectric mixing models," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-23, no. 1, pp. 35–46, Jan. 1985.
[22] M. T. Hallikainen, F. T. Ulaby, M. C. Dobson, M. A. El-Rayes, and L. Wu, "Microwave dielectric behavior of wet soil, Part I: Empirical models and experimental observations," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-23, no. 1, pp. 25–34, Jan. 1985.
[23] B. J. Cosby, G. M. Hornberger, R. B. Clapp, and T. R. Ginn, "A statistical exploration of the relationships of soil water characteristics to the physical properties of soils," *Water Resour. Res.*, vol. 20, pp. 682–690, 1984.
[24] T. J. Jackson, H. McNairn, M. A. Weltz, B. Brisco, and R. Brown, "First order surface roughness correction of active microwave observations for estimating soil moisture," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 1065–1069, Jul. 1997.
[25] T. J. Jackson, "Measuring surface soil moisture using passive microwave remote sensing," *Hydrol. Process.*, vol. 7, pp. 139–152, 1993.
[26] E. Njoku and L. Li, "Retrieval of land surface parameters using passive microwave measurements at 6–18 GHz," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 79–93, Jan. 1999.

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

He is currently a Research Scientist with the Hydrology and Remote Sensing Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, MD. His research involves the application of land surface modeling and remote sensing technology to hydrology and agriculture.

Steven Tsz K. Chan (M'02–SM'03) received the B.Sc. degree in electrical engineering from Portland State University, Portland, OR, in 1993, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, in 1995 and 1998, respectively.

In 1999, he joined the National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL), Pasadena, CA. At JPL, his research focuses on microwave remote sensing of soil moisture. Specifically, he is responsible for algorithm development of the Soil Moisture Data Product for the Advanced Microwave Scanning Radiometer (AMSR) onboard the NASA EOS Aqua satellite. More recently, he has been involved in science studies for the NASA Hydrosphere State (HYDROS) mission. He is a Scientist in the Water and Carbon Cycles group at JPL.

Dara Entekhabi (M'04–SM'04) received the Ph.D. degree in civil engineering from the Massachusetts Institute of Technology, Cambridge, in 1990.

He is currently a Professor in the Department of Civil and Environmental Engineering and in the Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge. He is the Hydros Principal Investigator.

Paul R. Houser received the Ph.D. degree in hydrology and water resources from the University of Arizona, Tucson, in 1996.

He is currently Head of the Hydrologic Sciences Branch, National Aeronautics and Space Administration Goddard Space Flight Center (GSFC), Greenbelt, MD. He is the Hydros GSFC Project Scientist.

Ann Y. Hsu received the Ph.D. degree from the University of Georgia, Athens.

She is currently with the Agricultural Research Service, Hydrology and Remote Sensing Laboratory, U.S. Department of Agriculture, Beltsville, MD, as a remote sensing specialist. She has been involves in studies that used remote sensing data, ranging from visible/near infrared to microwave, to derive land surface parameters.

Thomas J. Jackson (A'86–SM'96–F'02) received the Ph.D. degree in civil engineering from the University of Maryland, College Park, in 1976.

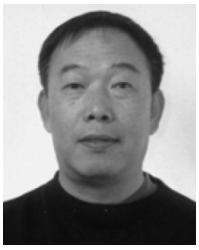
He is currently a Hydrologist with the U.S. Department of Agriculture's Agricultural Research Service Hydrology and Remote Sensing Laboratory, Beltsville, MD. His research involves the application and development of remote sensing technology in hydrology and agriculture.

Eni G. Njoku (M'77–SM'83–F'95) received the B.A. degree in natural and electrical sciences from Cambridge University, Cambridge, U.K., in 1972, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1974 and 1976.

He is currently a Principal Scientist with the Jet Propulsion Laboratory (JPL), Pasadena, CA, and is the JPL Project Scientist for the Hydros Mission. His primary interests are in the use of passive and active microwave remote sensing for hydrology and climate applications. His research involves studies of microwave interactions with land surfaces and retrieval algorithm development.

Peggy E. O'Neill (A'85–SM'03) received the B.S. degree (summa cum laude with university honors) in geography from Northern Illinois University, Chicago, in 1976, the M.A. degree in geography from the University of California, Santa Barbara, in 1979, and the Ph.D. degree in civil and environmental engineering from Cornell University, Ithaca, NY.

She is currently a Physical Scientist in the Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD, where she conducts research in soil moisture retrieval and land surface hydrology, primarily through microwave remote sensing techniques.



Jiancheng Shi (M'95–SM'02) received the B.A. degree from the University of Lanzhou, Lanzhou, China, and the M.A. and Ph.D. degrees in geography from the University of California, Santa Barbara (UCSB), in 1982, 1987, and 1991, respectively.

He then joined the Institute for Computational Erath System Sciences at UCSB as a Research Scientist. He is currently an Adjunct Researcher Scientist at the Institute of Remote Sensing and Applications in Beijing, Chinese Academy since 1998 and an Adjunct Professor at Beijing Normal University, China since 1999. His research interests are microwave modeling snow and soil signatures, image processing and analysis, and inversion models for retrieving physical parameters from remote sensing data. He is a Co-Principal Investigator of the Hydros mission.

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

He is currently a Research Scientist with the Goddard Earth Science and Technology Center, University of Maryland Baltimore County, Baltimore. His research interests include land surface soil moisture and vegetation observation, modeling, and data assimilation for global water, energy, and carbon cycle studies.