8-15-2012

# Hydrologic Data Assimilation with a Hillslope-Scale-Resolving Model and L Band Radar Observations: Synthetic Experiments with the Ensemble Kalman Filter

Alejandro N. Flores
*Boise State University*

Rafael L. Bras
*Georgia Institute of Technology*

Dara Entekhabi
*Massachusetts Institute of Technology*

 DOI: 10.1029/2011WR011500

# Hydrologic data assimilation with a hillslope-scale-resolving model and L band radar observations: Synthetic experiments with the ensemble Kalman filter

Alejandro N. Flores,[1] Rafael L. Bras,[2] and Dara Entekhabi[3]

Received 10 October 2011; revised 4 June 2012; accepted 19 June 2012; published 15 August 2012.

[1] Soil moisture information is critical for applications like landslide susceptibility analysis and military trafficability assessment. Existing technologies cannot observe soil moisture at spatial scales of hillslopes (e.g., $10^0$ to $10^2$ m) and over large areas (e.g., $10^2$ to $10^5$ km$^2$) with sufficiently high temporal coverage (e.g., days). Physics-based hydrologic models can simulate soil moisture at the necessary spatial and temporal scales, albeit with error. We develop and test a data assimilation framework based on the ensemble Kalman filter for constraining uncertain simulated high-resolution soil moisture fields to anticipated remote sensing products, specifically NASA's Soil Moisture Active-Passive (SMAP) mission, which will provide global L band microwave observation approximately every 2–3 days. The framework directly assimilates SMAP synthetic 3 km radar backscatter observations to update hillslope-scale bare soil moisture estimates from a physics-based model. Downscaling from 3 km observations to hillslope scales is achieved through the data assimilation algorithm. Assimilation reduces bias in near-surface soil moisture (e.g., top 10 cm) by approximately 0.05 m$^3$/m$^3$ and expected root-mean-square errors by at least 60% in much of the watershed, relative to an open loop simulation. However, near-surface moisture estimates in channel and valley bottoms do not improve, and estimates of profile-integrated moisture throughout the watershed do not substantially improve. We discuss the implications of this work, focusing on ongoing efforts to improve soil moisture estimation in the entire soil profile through joint assimilation of other satellite (e.g., vegetation) and in situ soil moisture measurements.

**Citation:** Flores, A. N., R. L. Bras, and D. Entekhabi (2012), Hydrologic data assimilation with a hillslope-scale-resolving model and L band radar observations: Synthetic experiments with the ensemble Kalman filter, *Water Resour. Res.*, *48*, W08509, doi:10.1029/2011WR011500.

## 1. Introduction

[2] There is currently no technological platform that allows for the observation of soil moisture at hillslope scales (e.g., $10^0$ to $10^2$ m), over large areas (e.g., $10^2$ to $10^5$ km$^2$) with sufficiently high temporal coverage (e.g., 1–3 days). Despite this gap in the hydrologic observing system, in many intensively monitored watersheds throughout the world soil moisture has been shown to vary at hillslope scales [*Seyfried and Wilcox*, 1995; *Grayson et al.*, 1997; *Jencso et al.*, 2009; *Kim*, 2009; *Williams et al.*, 2008]. At the same time, knowledge of soil moisture at spatial scales of hillslopes is critically important to a number of applications like slope stability assessment [*Simoni et al.*, 2008], soil erosion and slope stability in burnt areas [*Bovolo et al.*, 2009], and soil carbon assessment in complex terrain [*Gessler et al.*, 2000].

[3] Existing and emerging technologies allow for measurement of soil moisture across spatial scales ranging from effectively points (e.g., $10^{-2}$ m) to individual hillslopes (e.g., $10^2$ m). Some of the most promising new techniques include electrical resistivity tomography [e.g., *Brunet et al.*, 2010], distributed temperature sensing [*Selker et al.*, 2006; *Tyler et al.*, 2009], measurement of neutron thermalization due to interactions of primary cosmic rays with atmospheric nuclei [*Zreda et al.*, 2008], and noise characteristics of ground-based Global Positioning System (GPS) receivers [*Larson et al.*, 2010]. Despite these advances, ground-based observation of soil moisture over large watersheds at hillslope scales is unlikely in the foreseeable future.

[4] Space-based satellites are the best platform from which soil moisture may be observed globally and over large areas. The L band microwave region of the electromagnetic spectrum (i.e., 1–3 GHz) is particularly useful for measurement of soil moisture because soil electrical properties particularly sensitive to moisture content [*Njoku and Kong*, 1977; *Ulaby et al.*, 1986]. As a result, L band microwave

[1]Department of Geosciences, Boise State University, Boise, Idaho, USA.
[2]Department of Civil and Environmental Engineering and Department of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA.
[3]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Corresponding author: A. N. Flores, Department of Geosciences, Boise State University, 1910 University Dr., MS-1535, Boise, ID 83725-1535, USA. (lejoflores@boisestate.edu)


0043-1397/12/2011WR011500

observation has proven a powerful tool for global observation of soil moisture [e.g., *Njoku and Entekhabi*, 1996; *Crow et al.*, 2001; *Kerr et al.*, 2001; *Paloscia et al.*, 2006; *Entekhabi et al.*, 2010]. Passive L band microwave radiometry technology is vital to both the European Space Agency's Soil Moisture and Ocean Salinity (SMOS) mission [*Wigneron et al.*, 2000; *Pellarin et al.*, 2003] and NASA's Soil Moisture Active-Passive (SMAP) mission [*Entekhabi et al.*, 2010]. Passive L band microwave observations are, however, associated with coarse spatial resolutions—approximately 50 km for SMOS and 40 km for SMAP. The SMAP mission also uses an imaging radar operating at a central frequency of 1.26 GHz. The principal advantage of active remote sensing of soil moisture is a considerable improvement in the resolution of radar products; the SMAP mission calls for a 3 km soil moisture product [*Entekhabi et al.*, 2010]. Like passive L band microwave observations, L band radar observations are sensitive to the soil dielectric constant and therefore soil moisture [*Dobson and Ulaby*, 1986; *Wang et al.*, 1986; *Engman*, 1991; *Evans et al.*, 1992]. Relative to passive observation, however, soil moisture retrieval algorithms for radar observations are more complex [e.g., *Eom and Boener*, 1986; *Fung et al.*, 1992; *Fung*, 1994; *Tabatabaeenejad and Moghaddam*, 2006]. Radar observations are also more sensitive to roughness at the surface [*Fung et al.*, 1992] and water contained in vegetation, which attenuates and scatters transmitted microwave radiation [*Dubois et al.*, 1995; *Njoku et al.*, 2002].

[5] Progress in the observation of soil moisture has also been accompanied by improvements in physically based modeling of soil moisture. Integrated hydrologic models have improved process representation and computational efficiency, allowing simulation of watersheds ranging in scale from $10^1$ km$^2$ to $10^4$ km$^2$ [e.g., *Maxwell and Miller*, 2005; *Kollet et al.*, 2010; *Qu and Duffy*, 2007; *Kumar et al.*, 2009; *Ivanov et al.*, 2004a, 2004b, 2008a, 2008b]. A key strength of these integrated hydrologic models is that they synthesize diverse sources of data to represent topography, soils, vegetation, and environmental forcings in the area being simulated. But uncertainties in these input data, however, ultimately result in uncertainties in the predicted hydrologic states and fluxes.

[6] Leveraging the strengths of both hydrologic observation and modeling to combat these uncertainties, data assimilation methods—the algorithmic fusion of observations and models—have proven important for estimating soil moisture [e.g., *Margulis et al.*, 2002; *Reichle et al.*, 2002, 2008; *Crow and Wood*, 2003; *Dunne and Entekhabi*, 2005, 2006; *Merlin et al.*, 2006; *Kashif Gill et al.*, 2007; *Parada and Liang*, 2008; *Crow and Ryu*, 2009; *Pan and Wood*, 2010]. *Reichle et al.* [2001] importantly demonstrated that data assimilation can be used to effectively disaggregate passive microwave data from 40 km to 1 km scales. Similarly, *Merlin et al.* [2006] improved estimation of near-surface soil moisture through assimilation of disaggregated synthetic SMOS observations into a soil-vegetation-atmosphere transfer model. *Parada and Liang* [2008] found that assimilation of even coarse (25 km) passive microwave observations yielded improvement in near-surface soil moisture estimation within a macroscale hydrologic model. With the notable exceptions of *van Loon and Troch* [2002] and *Camporese et al.* [2010], however, these studies have mostly targeted the macroscale (i.e., $10^3$–$10^4$ m) because of the influence of soil moisture on surface energy fluxes estimates, to which weather forecasting models are sensitive.

[7] The purpose of this study is to construct and test a data assimilation system for estimating soil moisture at hillslope scales using an integrated hydrologic model with synthetic observations roughly consistent with those expected from SMAP. We use the tRIBS-VEGGIE model to simulate soil moisture conditions at hillslope scales within the Walnut Gulch Experimental Watershed in southeast Arizona, USA, during a synthetic 27 day period in August. An ensemble Kalman filter (EnKF) algorithm is used to assimilate synthetic SMAP radar measurements with a 3 km spatial resolution. The remainder of the paper is organized as follows: (1) an outline the methods, (2) an overview of the synthetic experiment framework, (3) results and interpretation, and (4) discussion, implications, and conclusions. Although the tRIBS-VEGGIE model is the focus here, we intend this study to be broadly illustrative of the potential use of data assimilation for improving hydrologic estimation at hillslope scales with models and remote sensing data.

## 2. Methods

[8] The lack of spaceborne L band radar observations adequate for observing soil moisture approximately every 2–3 days necessitates an Observing System Synthetic Experiment (OSSE) approach. Within the framework of the OSSE presented here, (1) a true distribution of bare soil moisture is simulated by the hydrologic model, (2) an observational model is used to synthesize L band microwave radar backscatter consistent with the 3 km spatial resolution and 3 day temporal revisit of the SMAP satellite, and (3) the synthetic backscatter observations are perturbed with noise that is consistent with SMAP radar and used to update simulated tRIBS-VEGGIE soil moisture states using the ensemble Kalman filter (Figure 1). The particular set of experiments presented here represent a multiple truth OSSE in which the workflow shown in Figure 1 is replicated four times with different sets of observations and results are considered in aggregate. Benefits and shortcomings of this multiple-truth OSSE approach are outlined in the discussion.

[9] Here we present (1) a brief discussion of the hydrologic model used in these experiments, (2) an outline of the L band microwave radar scattering model, (3) a review of the ensemble Kalman Filter algorithm used, (4) a review of the treatment of uncertainty in the model parameters and environmental forcings, and (5) a description of the experiment setting, details, and simplifying assumptions.

### 2.1. The tRIBS-VEGGIE Model

[10] The ecohydrology model used in this data assimilation study is the tRIBS-VEGGIE model [*Ivanov et al.*, 2004a, 2004b, 2007, 2008a, 2008b]. The tRIBS-VEGGIE model is a spatially distributed mass, energy, and carbon balance-resolving model that takes as input precipitation and meteorological forcings, as well as the topographic and soil boundary conditions. A full treatment of the tRIBS-VEGGIE model is beyond the present scope of work and the reader is directed to the work of *Ivanov et al.* [2004a, 2004b, 2007, 2008a, 2008b].

[11] The tRIBS-VEGGIE model simulates the spatial distribution of soil moisture by solving the one-dimensional

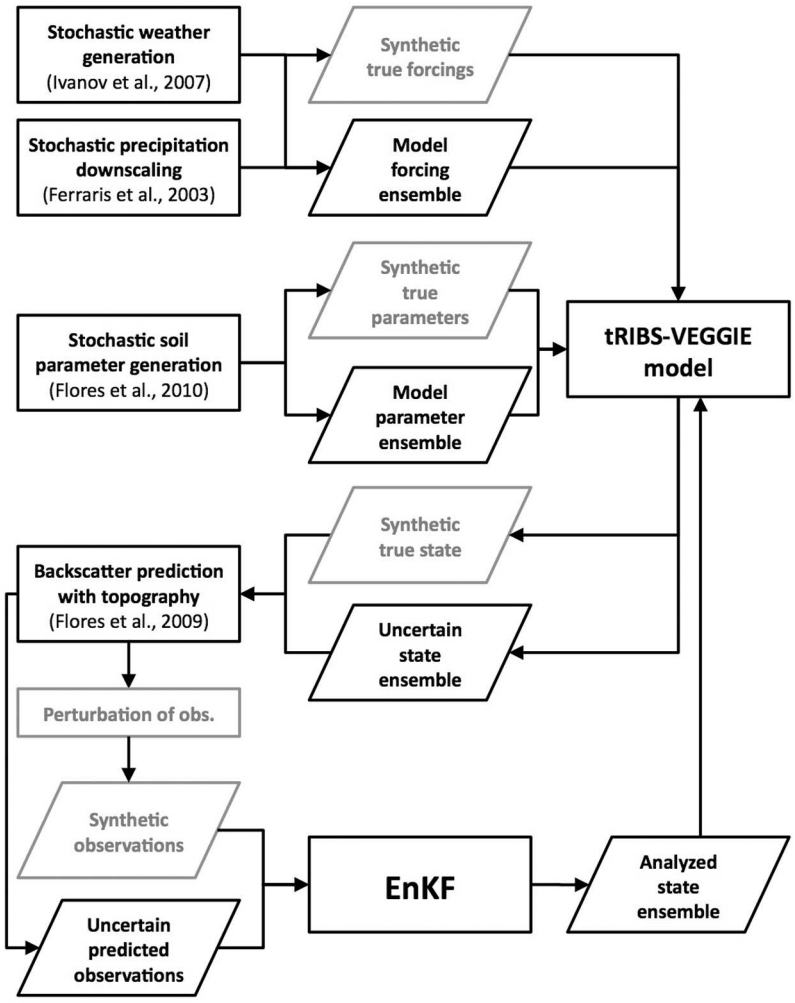


**Figure 1.** A flowchart depicting the workflow for one OSSE. Gray boxes show the steps followed to produce the synthetic true states and observations. Black boxes show the EnKF loop.

Richards equation for a sloped surface, allowing lateral gravitational drainage in the unsaturated zone in the direction of steepest topographic descent and driven by the local topographic slope. The bottom of the soil column is set as a free drainage boundary condition, consistent with the assumption of significant depth to the saturated zone in the semiarid environment for which the model is currently most applicable.

[12] Inputs to the tRIBS-VEGGIE model can be categorized into four kinds of data: (1) hourly hydrometeorological forcings, (2) soil hydraulic and thermal properties, (3) vegetation parameters, and (4) a static elevation field representing watershed topography. Hydrometeorological forcings for tRIBS-VEGGIE include hourly (1) precipitation, (2) sky fractional cover or incoming solar radiation, (3) air temperature, (4) dew temperature, and (5) wind speed. In this study, hydrometeorological forcings to tRIBS-VEGGIE were generated by a stochastic weather generator [*Ivanov et al.*, 2007]. Soil hydraulic and thermal parameters are similar to those common to many land surface models [e.g., *Liang et al.*, 1994; *Chen et al.*, 1996; *Koster and Suarez*, 1996; *Peters-Lidard et al.*, 1997; *Oleson et al.*, 2004] and available in published soil databases such as the STATSGO or SSURGO products. Soil parameters required for a water and energy balance solution include (1) saturated hydraulic conductivity, (2) saturation moisture content, (3) residual moisture content, (4) Brooks-Corey parameters pore distribution index and air entry pressure, (5) specific volumetric heat capacity, and (6) thermal conductivity.

[13] The study domain is the Walnut Gulch Experimental Watershed (WGEW) in southeast Arizona, USA. Established by the U.S. Department of Agriculture (USDA) in 1958, WGEW is approximately 150 $km^2$ in area, and is located at a transition between the Chihuahuan and Sonoran deserts [*Moran et al.*, 2008]. The semiarid watershed has a mean annual temperature of 17.7°C, while mean annual precipitation is 312 mm [*Moran et al.*, 2008]. The North American Monsoon System (NAMS) brings approximately 60% of the mean annual precipitation in the months of July through September, and summertime rainfall events tend to be localized and of high intensity [*Moran et al.*, 2008].

[14] Watershed topography is represented as a network of Voronoi polygons derived from a TIN representation of input static digital elevation models (DEMs). Recommendations of *Vivoni et al.* [2004] were followed to obtain a TIN mesh for tRIBS-VEGGIE from a U.S. Geological Survey 30 m Digital Elevation Model for the WGEW area using the ArcInfo Geographic Information System (GIS) package developed

**Table 1.** Characteristics of tRIBS-VEGGIE Computational Mesh

| Property | Value |
|---|---|
| Watershed area (km$^2$) | 148 |
| Number of pixels $N_p$ | 19,447 |
| Effective resolution[a] (m) | 87.2 |
| Number of soil layers | 10 |
| Number of soil textures | 27 |
| Minimum/maximum/ mean elevation (m) | 1222/1933/1420 |

[a]Effective resolution is defined as the square root of the watershed area divided by $N_p$.

by ESRI. Attempts to delineate floodplain features following the techniques used by *Vivoni et al.* [2004] revealed little discernable floodplain structure, and the final TIN mesh is composed of 19,447 computational nodes. Each tRIBS-VEGGIE pixel is formulated with 10 finite element soil layers with thicknesses that increase exponentially with depth. Layer thickness increases exponentially from 25 mm at the atmosphere-adjacent layer to 845 mm in the bottom layer. Total thickness of the soil column is 3155 mm. Properties of the computational mesh are summarized in Table 1.

### 2.2. The Integral Equation Model for L Band Radar Backscatter

[15] The forward geophysical model used to simulate L band microwave radar backscatter is the integral equation model (IEM) [*Fung*, 1994; *Altese et al.*, 1996; *Chen et al.*, 1995; *Hoeben and Troch*, 2000]. *Altese et al.* [1996] previously used the IEM with a linearized version of the Richards equation and an extended Kalman filter to estimate soil moisture. The IEM allows for modeling backscatter with information about (1) the radar sensor, (2) the roughness features of the area, and (3) the near-surface moisture conditions. Owing to its sensitivity to roughness, the IEM is most applicable for sparsely vegetated and/or unvegetated surfaces [*Dubois et al.*, 1995]. As such, we assume bare soil conditions. Although this assumption is an important departure from reality, it eliminates the need to address volume scattering from water stored in vegetation. Scattering models that account for effects of vegetation water content on backscatter have been developed, but require additional parameters related to the type and water content of the vegetation canopy [e.g., *Bindlish and Barros*, 2001].

[16] The like-polarized dimensionless backscatter coefficient, $\sigma^0_{pp}$, is expressed as

$$\sigma^0_{pp} = \frac{k_w^2}{2} e^{-2k_z^2 s} \sum_{n=1}^{\infty} s^{2n} \left| I^n_{pp} \right|^2 \frac{W^n(-2k_x, 0)}{n!}, \tag{1}$$

where the wave number ($k_w$) is equal to $2\pi/\lambda$ ($\lambda$ is the wavelength), $s$ is the surface roughness parameter, $k_z = k_w \cos\theta_l$, $k_x = k_w \sin\theta_l$, $\theta_l$ is the local incidence angle [radians], and $W(\bullet)$ is the power spectrum of the surface roughness. In equation (1) $\sigma^0_{pp}$ represents the ratio of backscattered energy; backscatter is typically presented in units of decibels, such that $\sigma^0_{pp}[\mathrm{dB}] = 10\log_{10}\sigma^0_{pp}$. The assumed radar frequency is 1.26 GHz, corresponding to a wavelength of approximately 24 cm. The subscript $pp$ denotes the copolarization state (i.e., $hh$ or $vv$ for horizontally or vertically copolarized states, respectively). The function $I_{pp}$ is defined as,

$$I^n_{pp} = (2k_z)^n f_{pp} e^{-s^2 k_z^2} + \frac{k_z^n}{2}\left(F_{pp}(-k_x, 0) + F_{pp}(k_x, 0)\right) \tag{2}$$

The copolarized Kirchoff coefficients ($f_{vv}$ and $f_{hh}$) are defined as

$$\begin{aligned} f_{vv} &= 2R_{vv}/\cos\theta_l \\ f_{hh} &= -2R_{hh}/\cos\theta_l \end{aligned} \tag{3}$$

and the copolarized complementary field coefficients ($F_{vv}$ and $F_{hh}$) defined as

$$F_{vv}(-k_x, 0) + F_{vv}(k_x, 0) = \frac{2\sin^2\theta_l}{\cos\theta_l}\left[\left(1 - \frac{\varepsilon\cos^2\theta_l}{\varepsilon - \sin^2\theta_l}\right)(1 - R_{vv})^2 + \left(1 - \frac{1}{\varepsilon}\right)(1 + R_{vv})^2\right] \tag{4}$$

$$F_{hh}(-k_x, 0) + F_{hh}(k_x, 0) = \frac{2\sin^2\theta_l}{\cos\theta_l}\left[4R_{hh} - \left(1 - \frac{1}{\varepsilon}\right)(1 + R_{hh})^2\right] \tag{5}$$

where $\varepsilon$ is the effective soil dielectric constant (dimensionless). The soil dielectric model used in this study is an empirical model relating the effective dielectric constant $\varepsilon_{\mathrm{eff}}$ to volumetric soil moisture $\theta_v$ (m$^3$/m$^3$) [*Topp et al.*, 1980] as follows:

$$\varepsilon_{\mathrm{eff}} = 3.03 + 9.3\,\theta_v + 146\,\theta_v^2 - 76.7\,\theta_v^3. \tag{6}$$

The surface roughness spectrum, $W(u, v)$, is related to the surface roughness correlation function, $\rho(\zeta, \xi)$, through its Fourier transform,

$$W^n(u, v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho^n(\xi, \zeta) e^{-ju\xi - jv\zeta} d\xi d\zeta \tag{7}$$

For simplicity, we assume an exponential function roughness correlation function as in *Fung et al.* [1992], *Chen et al.* [1995], and *Altese et al.* [1996], which is a function of only the roughness correlation length $L$. Noting that the roughness parameters of the IEM are often normalized by $k_w$, the $n$th power of the roughness spectrum is then

$$W^n(k_w) = \left(\frac{L}{n}\right)^2 \left[1 + \left(\frac{k_w L}{n}\right)^2\right]^{-1.5} \tag{8}$$

Topography plays an important role in soil moisture remote sensing, particularly at the 1–3 km resolution of the SMAP radar [*Mätzler and Standley*, 2000; *Kerr et al.*, 2003; *Mialon et al.*, 2008; *Sandells et al.*, 2008; *Flores et al.*, 2009]. *Mätzler and Standley* [2000] and *Flores et al.* [2009] provide a framework through which the local incidence angle and the aggregation of hillslope-scale elements to the viewing resolution can be calculated from digital elevation data. At a tRIBS-VEGGIE pixel at which backscatter is modeled the local incidence angle is a function of the azimuth angle ($\zeta_S$, rad) and zenith angle ($\delta_S$, rad) from the pixel to the

satellite, and the local topographic slope ($\alpha_\nabla$) and aspect ($\zeta_\nabla$),

$$\cos\theta_l = \cos\alpha_\nabla \cos\delta_S + \sin\alpha_\nabla \sin\delta_S \cos(\zeta_S - \zeta_\nabla), \quad (9)$$

and the rotation of the linear polarization by an angle, $\phi$, is calculated as

$$\sin\phi = \sin(\zeta_S - \zeta_\nabla)\sin\alpha_\nabla / \sin\theta_l. \quad (10)$$

Local (i.e., hillslope-scale) incidence and polarization rotation angles are used to compute reflectivity in both horizontally and vertically copolarized states through the Fresnel equations. Hillslope-scale backscatter is computed based on this local reflectivity and aggregated to a regular 3 km grid, accounting for effects of topographic slope and aspect [see *Flores et al.*, 2009; *Mätzler and Standley*, 2000].

### 2.3. The Ensemble Kalman Filter

[17] The ensemble Kalman filter (EnKF) is an extension of the Kalman filter that allows for the use of nonlinear system models by approximating the prior state error statistics using Monte Carlo simulation [*Evensen*, 1994, 2003, 2004]. It is important to recognize that because the state error covariance matrix is approximated through Monte Carlo simulation, the state estimate will be inherently suboptimal. Despite departures from the requirements of the Kalman Filter, the EnKF has seen broad applicability in hydrology, meteorology, and oceanography.

[18] Several update (or "analysis") schemes have been developed to deal with problems frequently encountered in geophysical data assimilation. This study uses the analysis scheme outlined in *Evensen* [2004, section 7.4.2]. Let $x_k(t)$ denote the $n_s$-dimensional tRIBS-VEGGIE model state vector (i.e., all soil moisture pixels within the areal extent of each watershed and all vertical soil moisture layers at each pixel) for the $k$th of $K$ members of an ensemble of model states simulated at a time $t$ when a SMAP observation is available. More explicitly, the vector $x_k(t|t-1)$ is a replicate of a $K$-member unconditional forecast of soil moisture, given a $K$ replicate ensemble that was previously conditioned on observations at time $t-1$, $x_k(t-1|t-1)$. Each replicate is obtained from,

$$x_k(t|t-1) = f(x_k(t-1|t-1), u_k, \phi_k) \quad (11)$$

where the function $f(\bullet)$ represents the tRIBS-VEGGIE model, $u_k$ is the $k$th realization of the uncertain hydrometeorological forcings between time $t-1$ and $t$, and $\phi_k$ is the $k$th realization of the uncertain soil hydraulic and thermal properties. Dropping the explicit dependence on time $t$, and denoting each vector of the unconditional ensemble forecast with a superscript '$f$' ("forecast"), we collect the $K$ replicates of the forecast ensemble into the $n_s \times K$ matrix $\mathbf{X}^f$,

$$\mathbf{X}^f = \begin{bmatrix} x_1^f & x_2^f & x_3^f & \dots & x_K^f \end{bmatrix} \quad (12)$$

The $n_s \times 1$ sample mean vector of this forecast ensemble, $\bar{x}^f$, is computed and used to form the $n_s \times K$ matrix, $\bar{\mathbf{X}}^f$, which contains $K$ copies of $\bar{x}^f$.

[19] The radar backscatter model outlined above transforms the input state $x_k^f$ into an $m \times 1$ vector of predicted observations

$$\mathbf{z}_k^f = H\left(x_k^f\right) \quad (13)$$

In data assimilation applications where the dimensionality of the state vector is large, computing the full $n_s \times n_s$ estimate of the state error covariance is difficult or intractable. For this reason, a series of so-called square-root analysis schemes, which instead only require construction of a matrix containing an ensemble of perturbations from the ensemble mean, have become popular. This analysis scheme is summarized here, with slight modifications from the notation of *Evensen* [2004].

[20] The ensemble of predicted observation perturbations, $\mathbf{S}$, is

$$\mathbf{S} = H\left(\mathbf{X}^f - \bar{\mathbf{X}}^f\right) \approx H\left(\mathbf{X}^f\right) - H\left(\bar{\mathbf{X}}^f\right) \quad (14)$$

in which $H(\bullet)$, $\mathbf{X}^f$, and $\bar{\mathbf{X}}^f$ have been previously defined. It is important to note that this approximation to $\mathbf{S}$, although necessary, assumes that backscatter varies with soil moisture in only a weakly nonlinear fashion. The $m \times K$ matrix $\mathbf{S}$ is factored using the singular value decomposition (SVD) into

$$\mathbf{U}_0 \mathbf{\Sigma}_0 \mathbf{V}_0^T = \mathbf{S} \quad (15)$$

An intermediate matrix $\mathbf{X_0}$ is then constructed as

$$\mathbf{X}_0 = \mathbf{\Sigma}_0^+ \mathbf{U}_0^T \mathbf{E} \quad (16)$$

where the matrix $\mathbf{\Sigma}_0^+$ is a diagonal matrix in which diag $(\mathbf{\Sigma}_0^+) = (\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_{K-1}^{-1}, 0)$, the elements of which are the reciprocal of the corresponding diagonal elements of the matrix $\mathbf{\Sigma}_0$. The SVD is performed on the matrix $\mathbf{X}_0$,

$$\mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{V}_1^T = \mathbf{X}_0 \quad (17)$$

and an intermediate matrix $\mathbf{X}_1$, computed as

$$\mathbf{X}_1 = \mathbf{U}_0 \mathbf{\Sigma}_0^{+T} \mathbf{U}_1 \quad (18)$$

The $n_s \times 1$ analyzed ensemble mean vector ($\bar{x}^a$) is then computed via the following five-step process of matrix-vector multiplications:

$$\mathbf{y}_0 = \mathbf{X}_1^T\left(\mathbf{z} - \mathbf{H}\mathbf{x}^f\right), \quad (19a)$$

$$\mathbf{y}_2 = \left(\mathbf{I} + \mathbf{\Sigma}_1^2\right)^{-1} \mathbf{y}_0, \quad (19b)$$

$$\mathbf{y}_3 = \mathbf{X}_1 \mathbf{y}_2, \quad (19c)$$

$$\mathbf{y}_4 = \mathbf{S}^T \mathbf{y}_3, \quad (19d)$$

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \left(\mathbf{X}^f - \bar{\mathbf{X}}\right)\mathbf{y}_4, \quad (19e)$$

A subsequent sequence of matrix operations is performed to compute the analyzed ensemble of state perturbations, $\mathbf{X}'^a$. First, an intermediate matrix $\mathbf{X}_2$ is computed as

$$\mathbf{X}_2 = \left(\mathbf{I} + \boldsymbol{\Sigma}_1^2\right)^{-1/2} \mathbf{X}_1^T \mathbf{S} \tag{20}$$

and factored via the SVD to obtain

$$\mathbf{U}_2 \boldsymbol{\Sigma}_2 \mathbf{V}_2^T = \mathbf{X}_2. \tag{21}$$

Finally, the matrix of analyzed ensemble perturbations is computed as

$$\mathbf{X}'^a = \left(\mathbf{X}^f - \bar{\mathbf{X}}\right) \mathbf{V}_2 \sqrt{\mathbf{I} - \boldsymbol{\Sigma}_2^T \boldsymbol{\Sigma}_2}\, \boldsymbol{\Theta}^T \tag{22}$$

where the random orthonormal matrix $\boldsymbol{\Theta}$ is composed of the right singular vectors from a SVD performed on a random $K \times K$ matrix. The purpose of the matrix $\boldsymbol{\Theta}$ is to spread the variance reduction across the ensemble members. Recent studies suggest that this random rotation can potentially induce a bias in the analysis ensemble [*Wang et al.*, 2004; *Sakov and Oke*, 2008]. While *Livings et al.* [2008] provides some diagnostic tools to assess the degree to which this rotation induces bias, we have not addressed this issue for the OSSEs presented here and leave this to future study.

[21] The square root analysis used in this study is ideally suited to $m \gg K$. In the numerical experiments described below, $m$ is much less than $K$. The employed analysis algorithm used here, therefore, is more computationally expensive than other square root analysis algorithms. The majority of the cost in our numerical experiments is associated with integration of the tRIBS-VEGGIE model and the increased cost of the analysis scheme used is miniscule.

### 2.4. Treatment of Uncertainty in Model Parameters and Forcings

[22] We assume uncertainty in the simulated soil moisture state arises from two principal sources: (1) imperfect characterization of the soil hydraulic and thermal properties and (2) errors in and sparsity of the measurement of the hydrometeorologic forcings of the tRIBS-VEGGIE model. We briefly describe here how these sources of uncertainty are treated in the developed data assimilation framework.

#### 2.4.1. Development of Synthetic True and Ensemble Model Forcings

[23] The hydrometeorological forcings for this synthetic study were designed to broadly replicate experiment conditions that might be encountered in an application where, for example, hydrometeorological forcings are remote sensing observations or downscaled outputs of numerical weather prediction or climate models. As such, we assume there are macroscale or synoptic-scale hydrometeorological conditions that are input to a stochastic downscaling scheme to develop forcings for a watershed model. In practice, we use a modified version of the weather generator of *Ivanov et al.* [2007] to generate both true and uncertain hydrometeorological forcings for this study. Modifications to the weather generator of *Ivanov et al.* [2007] include (1) a Bartlett-Lewis model to generate hourly precipitation forcings [*Islam et al.*, 1990] and (2) a stochastic multiplicative cascade model to disaggregate hourly rainfall forcings in space to a 4 km grid [*Ferraris et al.*, 2003]. Note that this is coarser than the 3 km scale observations, but consistent with existing weather radar rainfall products and contemporary numerical weather prediction models. Outputs of the weather generator are hourly rainfall; air temperature; dew point temperature; wind speed; cloud fraction; incoming direct beam solar radiation; diffuse solar radiation; and thermal radiation. Parameters for the Bartlett-Lewis model were obtained for Tucson from *Hawk* [1992], while the remaining parameters for the weather generator were obtained from *Ivanov et al.* [2007] for Tucson. Parameters for the multiplicative cascade were taken from *Ferraris et al.* [2003] and ensure that disaggregation is mass conservative in the ensemble mean sense. Consistent with this experimental framework, the weather generator parameters are fixed throughout all OSSEs and we assume the timing of rainfall is known and does not change between synthetic truths or ensemble replicates.

[24] A master set of hourly precipitation is obtained at output from the Bartlett-Lewis rainfall generator. This precipitation time series establishes the timing of hourly rainfall for the entire experiment. The hourly precipitation forcings are perturbed with multiplicative noise to produce one synthetic true precipitation sequence and supplied as input to the weather generator to produce the remainder of the true hydrometeorological forcings. Multiplicative perturbations follow a lognormal distribution with a mean of 1, a standard deviation of 0.4, and are assumed serially uncorrelated.

[25] The true hourly rainfall realization is perturbed again with multiplicative noise with the same characteristics as described above and the resulting rainfall time series input to the weather generator to obtain the remainder of the hydrometeorological forcings and spatially distributed rainfall. This set of hydrometeorological forcings serve as the observed or forecast hydrometeorological conditions. Cumulative synthetic observed precipitation is shown in Figure 2d. Synthetic observed rainfall time series are input to the weather generator to obtain an ensemble of air temperature, dew point temperature, wind speed, cloud fraction, incoming direct beam solar radiation, diffuse solar radiation, thermal radiation, and spatially distributed rainfall with a resolution of 4 km. Although the multiplicative cascade model we use to disaggregate hourly rainfall volumes is designed to reproduce observed spatial correlation in rainfall fields, we made no effort to impose a particular spatiotemporal correlation structure on the ensemble of rainfall forcings. The relatively small size of the watershed with respect to the rainfall forcing grid (complete watershed coverage is achieved with a $4 \times 7$ grid at 4 km resolution as seen in Figure 2d) implies that, while it is reasonable to expect that rainfall is highly correlated in space, spatial correlation could not be captured in a statistically significant way at the resolution of the rainfall forcings. Moreover, based on the extent and resolution of the rainfall grid, any spatial correlation imposed to the perturbations could not be statistically distinguished from spatially uncorrelated perturbations. Future experiments with either a larger domain or a higher-resolution rainfall "product," however, would necessitate more careful consideration of the spatiotemporal statistical characteristics of stochastically generated or perturbed rainfall (e.g., following the technique proposed by *Margulis et al.* [2006]).

#### 2.4.2. Uncertainty in Soil Hydraulic and Thermal Properties

[26] Treatment of uncertainty in the soil hydraulic and thermal properties of the study area is based on the approach

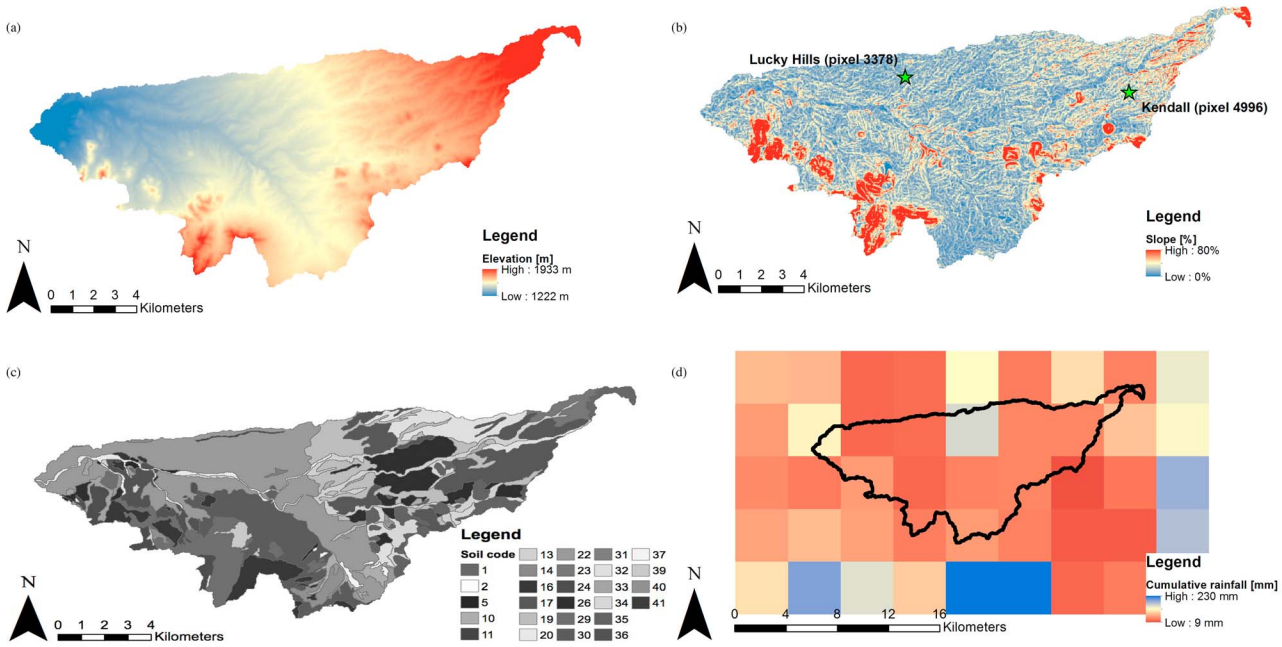


**Figure 2.** For the USDA Walnut Gulch Experimental Watershed, in southeast Arizona, USA, the spatial distribution of (a) topography, (b) topographic slope, (c) soil textural units, and (d) stochastically generated cumulative rainfall during the experiment, averaged over all four synthetic observation scenarios. Locations of intensively monitored tRIBS-VEGGIE pixels corresponding to Lucky Hills (pixel 3378) and Kendall (pixel 4996) are indicated in Figure 2b.

developed by *Flores et al.* [2010]. The authors of that study used a Latin Hypercube-based sampling scheme that also preserved empirically observed correlations between the soil hydraulic and thermal properties in the metadata database of *Schaap and Leij* [1998]. For a given soil textural class, the algorithm produces an ensemble of hydraulic and thermal properties that capture these correlations and also sample the extremes of parameter distribution. By using this technique to generate soil parameter ensembles for the tRIBS-VEGGIE model, *Flores et al.* [2010] showed that the corresponding simulated soil moisture ensembles were reproducible at ensemble sizes that were substantially smaller than a more simple sampling approach. We apply this technique to generate an ensemble of tRIBS-VEGGIE soil parameters for each soil textural unit within the WGEW.

[27] The spatial distribution of soil textural units was obtained from the Soil Survey Geographic (SSURGO) database maintained and published by the U.S. Department of Agriculture Natural Resource Conservation Service (USDA-NRCS). The SSURGO database reveals several soil textural units that are classified in more detail than the nine classes used in the Restricted Pairing RP-based soil parameter generation technique developed by *Flores et al.* [2010]. To facilitate the use of their soil parameter stochastic generation scheme, the more finely classified soil textural units in the SSURGO database for WGEW were reclassified to the conceptually nearest of the nine soil textural classes used in the approach developed by *Flores et al.* [2010]. For instance, "stony sandy loam" would be reclassified as a "sandy loam." Distinct soil units sharing a textural class but exhibiting distinct spatial coverage in the SSURGO spatial database were treated as distinct and independent when generating stochastic soil properties. Soil parameters are spatially uniform within each soil unit. The spatial distribution of soil units is shown in Figure 2c, and the soil texture associated with each soil unit is given in Table 2.

**Table 2.** Soil Codes and Textures

| Soil Code | Texture |
|---|---|
| 1 | Sandy loam |
| 2 | Sandy loam |
| 5 | Sandy loam |
| 10 | Clay |
| 11 | Loam |
| 13 | Sandy loam |
| 14 | Clay |
| 16 | Silt loam |
| 17 | Loam |
| 19 | Loamy sand |
| 20 | Loamy sand |
| 22 | Sandy loam |
| 23 | Sandy loam |
| 24 | Clay loam |
| 26 | Sandy loam |
| 29 | Silty clay loam |
| 30 | Silty clay loam |
| 31 | Silty clay loam |
| 32 | Silty clay loam |
| 33 | Sandy loam |
| 34 | Sandy loam |
| 35 | Clay |
| 36 | Loam |
| 37 | Sandy loam |
| 39 | Sandy loam |
| 40 | Silty clay loam |
| 41 | Sandy loam |

**Table 3.** Dimensions of the Data Assimilation Experiments

| Property | Value |
|---|---|
| State vector dimension $n_s$ | 194,470 |
| Number of observations $m$ | 60 |
| Number of EnKF replicates $K$ | 256 |
| Number of analysis period $N_A$ | 9 |
| Number of OSSE experiments $N_E$ | 4 |

### 2.5. Synthetic Experiment Setup and Simplifying Assumptions

[28] One distinct feature of this OSSE is that it is a multiple truth OSSE. That is, the workflow in Figure 1 is repeated four times with four sets of hydrometeorological forcings, and four sets of soil parameters that yield four realizations of synthetic truth and observations. The four 27 day realizations of the true soil moisture state were simulated using tRIBS-VEGGIE for each of the four true hydrometeorological forcings and four realizations of the soil parameters described above. Soil parameters for the four true simulations were generated using the technique described above. The four sets of hydrometeorological forcings and parameters created should yield corresponding true soil moisture and synthetic observations that are statistically unbiased with respect to each other because of the way they were generated (as an independent ensemble). This may not be true in practice, however, because of the small number of synthesized true soil moisture states and observations.

[29] The OSSEs were set during a hypothetical 27 day period in August. Dimensions of the data assimilation experiments are given in Table 3. Four time sequences of horizontally and vertically copolarized backscatter observations on a 3 × 3 km grid over the watershed were generated at 72 h (3 day) intervals based on the synthetic true soil moisture states at 0900 local time using the observing system described in section 2.2. We assume the satellite azimuth angle ($\zeta_S$) equals 150°, consistent with an ascending limb traverse of the site, and the off-nadir look angle $\delta_S$ equals 40°. Allowing for the presence of radar pixels not falling completely within WGEW, the spatial resampling of synthetic observations to the 3 km square grid reveals that WGEW is intersected by 30 pixels. Hence, the synthetic radar backscatter observations of WGEW based on the true states in both the horizontally and vertically copolarized states produces 60 backscatter measurements (30 in each copolarized state) every 72 h. Each set of observations was perturbed with multiplicative noise to simulate the known speckle of active observations with a variance of 0.5 dB equal to the relative accuracy of the SMAP radar (e.g.,

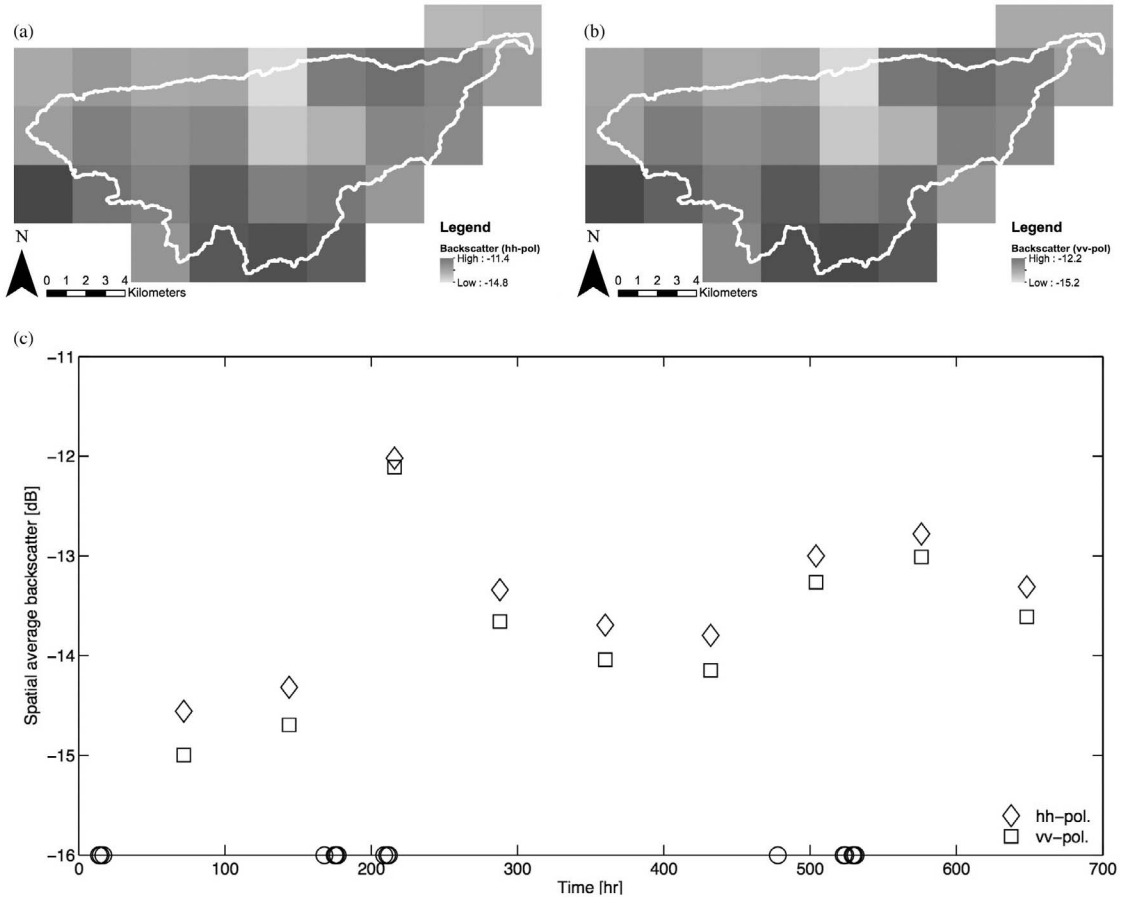


**Figure 3.** Synthetic L band radar backscatter observations averaged temporally over the duration of the experiment in (a) *hh* polarization and (b) *vv* polarization and (c) spatially averaged in both copolarized states as a function of time. In Figure 3c open circles on the ordinate axis correspond to the occurrence of rainfall events.

Table 1) [*Entekhabi et al.*, 2010]. There was no correlation assumed between perturbations to *hh* and *vv* copolarized observations The spatial distribution of backscatter observations in the *hh* and *vv* copolarized states are shown in Figures 3a and 3b, respectively. The spatial average backscatter observations at each of the nine analysis times are also shown in Figure 3c.

[30] For each of the four candidate sequences of observations, an EnKF experiment was performed to estimate soil moisture with an ensemble size of 256 replicates. For comparison we performed an open loop (OL) ensemble simulation (i.e., unconstrained to observations) during the study period. The OL ensemble consisted of 1024 simulations obtained by subjecting tRIBS-VEGGIE to uncertain hydrometeorological forcings and soil parameters. Forcings for each EnKF and OL experiments were obtained as described above using the stochastic weather generator.

[31] The OL ensemble is meant to characterize the "worst case" hillslope-scale soil moisture estimation scenario as no observational information is allowed to inform the state. In both the EnKF and OL cases, we assume only knowledge of the spatial distribution of categorical soil types as described in section 2.4.2. For both the EnKF and OL experiments, initial soil moisture conditions were obtained by performing an initial 27 day ensemble spin-up of the model with stochastically generated soil parameters and hydrometeorological forcings. The true soil moisture simulations, however, were initialized with hydrostatic conditions corresponding to 10% effective soil saturation. Although this impacts the first couple analysis cycles, it does not appreciably affect the overall conclusions.

[32] It is also important to note that the difference in size between the EnKF and OL ensembles may impact inferences from the respective simulations. At the relatively small spatial scale of the experiments and given the described methodology for generating hydrometeorological forcings, however, the primary driver of variance in soil moisture is uncertainty in soil hydraulic and thermal properties. Under these circumstances, *Flores et al.* [2010] demonstrate little change in the expected ensemble moments in ensembles ranging in size from 16 to 1024.

[33] Conclusions of any particular OSSE are specific to the synthetic true soil moisture used to produce the synthetic observations. In this case the true soil moisture is the result of a tRIBS-VEGGIE simulation subjected to particular combination of SHTPs and hydrometeorological forcings. The multiple-truth OSSE approach used here provides an experimental procedure to assess the robustness of the data assimilation system because it can be exercised with observations that correspond to noisy measurements of soil moisture conditions that are potentially far from the open loop ensemble mean. The robustness of the assimilation system is revealed by assessing the error characteristics averaged over all EnKF experiments. Because the true soil parameters and hydrometeorological forcings are generated with some degree of randomness, however, the results of any single EnKF experiment are less informative than the collection of experiments, considered together. To draw the broadest inference about the performance of the data assimilation system we constructed, the OSSEs presented here consider an average over four potential true realizations of the soil moisture state (i.e., a multiple-truth OSSE).

## 3. Results

[34] This study relies on synthetic experiments as a way to elucidate potential reductions in prediction errors of hillslope-scale predictions of soil moisture realized through assimilation of anticipated SMAP products. The synthetic nature of the experiments facilitates calculation of error diagnostics relative to the simulated true states. Especially useful is the comparison between error characteristics arising from the EnKF experiments and those from the OL experiments, in which no assimilation is performed.

### 3.1. Spatially Distributed Results

[35] At the watershed scale, the results of the assimilation experiments focus on (1) the spatial distribution of errors averaged over the nine analysis cycles and four OSSE experiments and (2) the temporal evolution of area-weighted average error characteristics at each analysis time. To facilitate deeper discussion of prediction errors in space and time, we present the results in terms of soil moisture vertically integrated in the near surface (i.e., top 10 cm) and soil profile (i.e., ~3000 cm). The absolute value of the bias between the ensemble mean and synthetic true soil moisture at each tRIBS-VEGGIE pixel, averaged over all four OSSEs and analysis times, is computed as

$$\mathrm{BIAS} = \frac{1}{N_A}\sum_{i=1}^{N_A}\left(\frac{1}{N_E}\sum_{j=1}^{N_E}\left|\bar{\theta}_{\mathrm{Ens},i,j} - \theta_{\mathrm{True},i,j}\right|\right), \qquad (23)$$

where $\bar{\theta}_{\mathrm{Ens},i,j}$ is the EnKF or OL ensemble mean near surface of profile-integrated soil moisture at the *i*th EnKF analysis time and for the *j*th OSSE, $\theta_{\mathrm{True},i,j}$ is the corresponding true near-surface or profile-integrated soil moisture, $N_A$ is the number of analysis cycles ($N_A$ = 9, over 27 days at 3 day intervals between observations), and $N_E$ is the number of true simulations ($N_E$ = 4). The root mean squared error (RMSE) averaged over all four synthetic true states and over the nine analysis cycles is computed as

$$\mathrm{RMSE} = \frac{1}{N_A}\sum_{i=1}^{N_A}\left(\frac{1}{N_E}\sum_{j=1}^{N_E}\sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(\theta_{\mathrm{Ens},i,j,k} - \theta_{\mathrm{True},i,j}\right)^2}\right), \qquad (24)$$

where $\theta_{\mathrm{Ens},i,j,k}$ is the vector of near-surface or profile-integrated soil moisture at the *i*th analysis time, *j*th EnKF or OL experiment, and *k*th EnKF or OL ensemble replicate. It should be noted that the bias and RMSE for the OL experiments will vary between OSSEs even though the OL soil moisture replicates and ensemble mean will not (because no assimilation has been performed). This definition of RMSE varies from that used in many other data assimilation studies, which focus on the RMSE of each ensemble replicate through time [e.g., *Leisenring and Moradkhani*, 2011; *Montzka et al.*, 2011]. This is largely because our applications of interest (military trafficability, landslide prediction, etc.) are more sensitive to instantaneous errors in soil moisture predictions. We temporally average instantaneous RMSEs to obtain an expectation of the instantaneous error over the time horizon of the experiment. Future studies

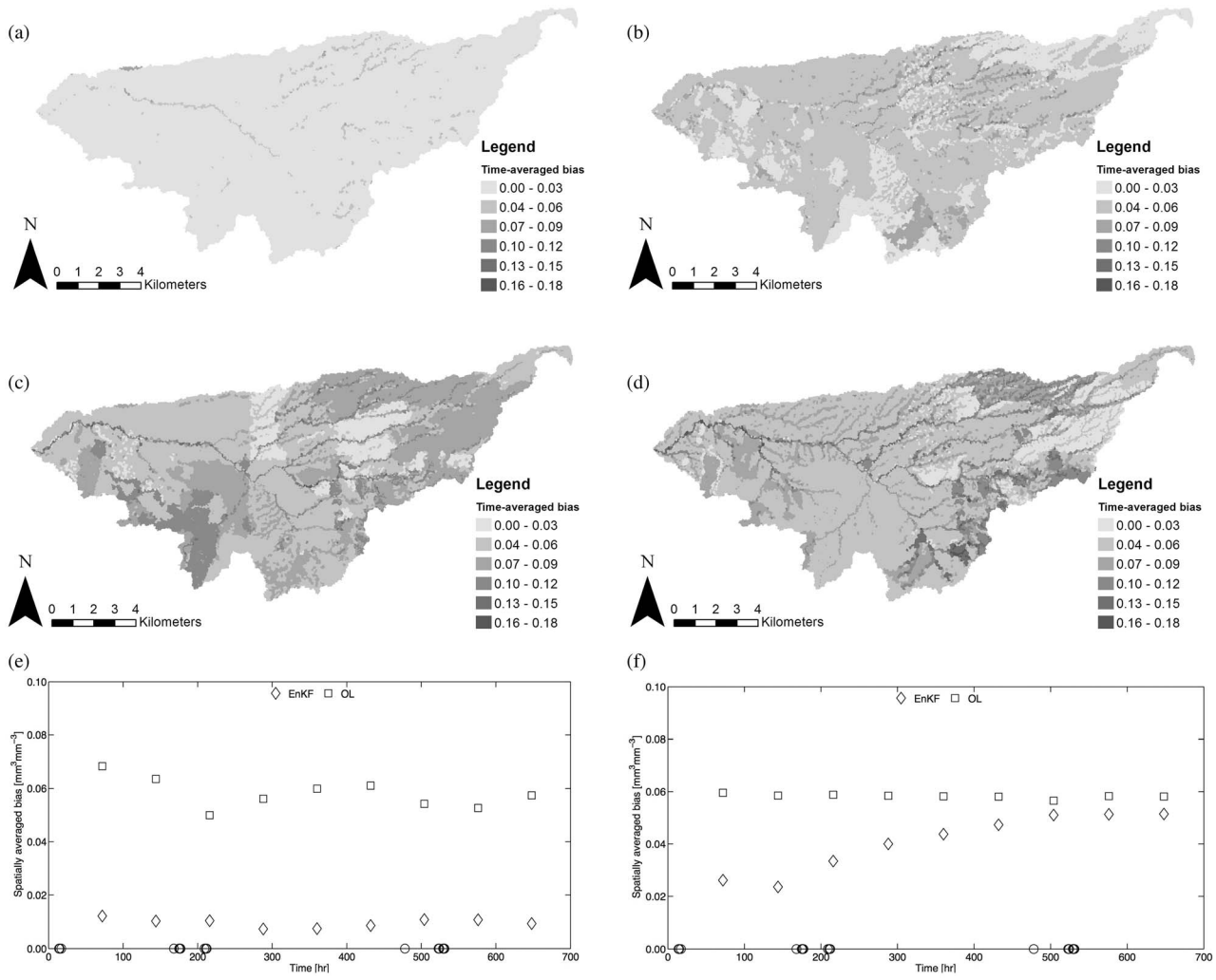


**Figure 4.** Absolute bias, averaged over the nine analysis times, for (a) EnKF-derived near-surface soil moisture, (b) EnKF-derived profile soil moisture, (c) OL-derived near-surface soil moisture, and (d) OL-derived profile soil moisture. Spatially averaged bias at the nine analyses for (e) near-surface and (f) profile moisture with open circles on the ordinate axes depicting temporal occurrence of rainfall.

geared toward discharge prediction and assimilation will utilize the more traditional definition of RMSE.

[36] At each analysis cycle, the pixel area-weighted bias can be computed as

$$\overline{\mathrm{BIAS}} = \sum_{p=1}^{N_p} \left( \frac{1}{N_E} \sum_{j=1}^{N_E} \left| \bar{\theta}_{\mathrm{Ens},p,j} - \theta_{\mathrm{True},p,j} \right| \right) \cdot a_p \Big/ \sum_{p=1}^{N_p} a_p, \quad (25)$$

where $a_p$ is the area of each tRIBS-VEGGIE pixel and $N_p$ is the number of pixels ($N_p$ = 19,447). Similarly, the pixel area-weighted RMSE can be computed at each analysis as

$$\overline{\mathrm{RMSE}} = \sum_{p=1}^{N_p} \left( \frac{1}{N_E} \sum_{j=1}^{N_E} \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left( \theta_{\mathrm{Ens},i,j,k} - \theta_{\mathrm{True},i,j} \right)^2} \right) \cdot a_p \Big/ \sum_{p=1}^{N_p} a_p. \quad (26)$$

The average bias in near-surface soil moisture after the nine EnKF analyses is less than 0.03 $m^3/m^3$ throughout virtually the entire watershed, while areas of the watershed where bias in near surface soil moisture is greater than 0.03 are limited primarily to valley bottoms (Figure 4a). The watershed-averaged bias in near-surface soil moisture, in fact, is less than 0.01 $m^3/m^3$ at all nine analyses, regardless of whether those analyses occur immediately after wetting or drying events (Figure 4e). By contrast, the corresponding bias in the OL estimate of near-surface soil moisture exhibits a spatial distribution exhibiting organization at a number of different scales (Figure 4c). The spatial arrangement of soil textural units is the primary driver of variation in near-surface soil moisture bias in the OL experiments, while valley bottoms seem also to be associated with higher bias in near-surface moisture. There is also a faint overprinting of the 4 × 4 $km^2$ rainfall forcing grid that can be seen in the central region of the watershed that explains some of the spatial variation in bias in OL-derived near-surface soil moisture. The watershed-averaged bias in OL-derived near-surface soil moisture exhibits some sensitivity to rainfall forcings, with rainfall events serving to decrease bias slightly (Figure 4e). Comparing the EnKF and OL bias in near-surface moisture, it is apparent that assimilation of the synthetic SMAP

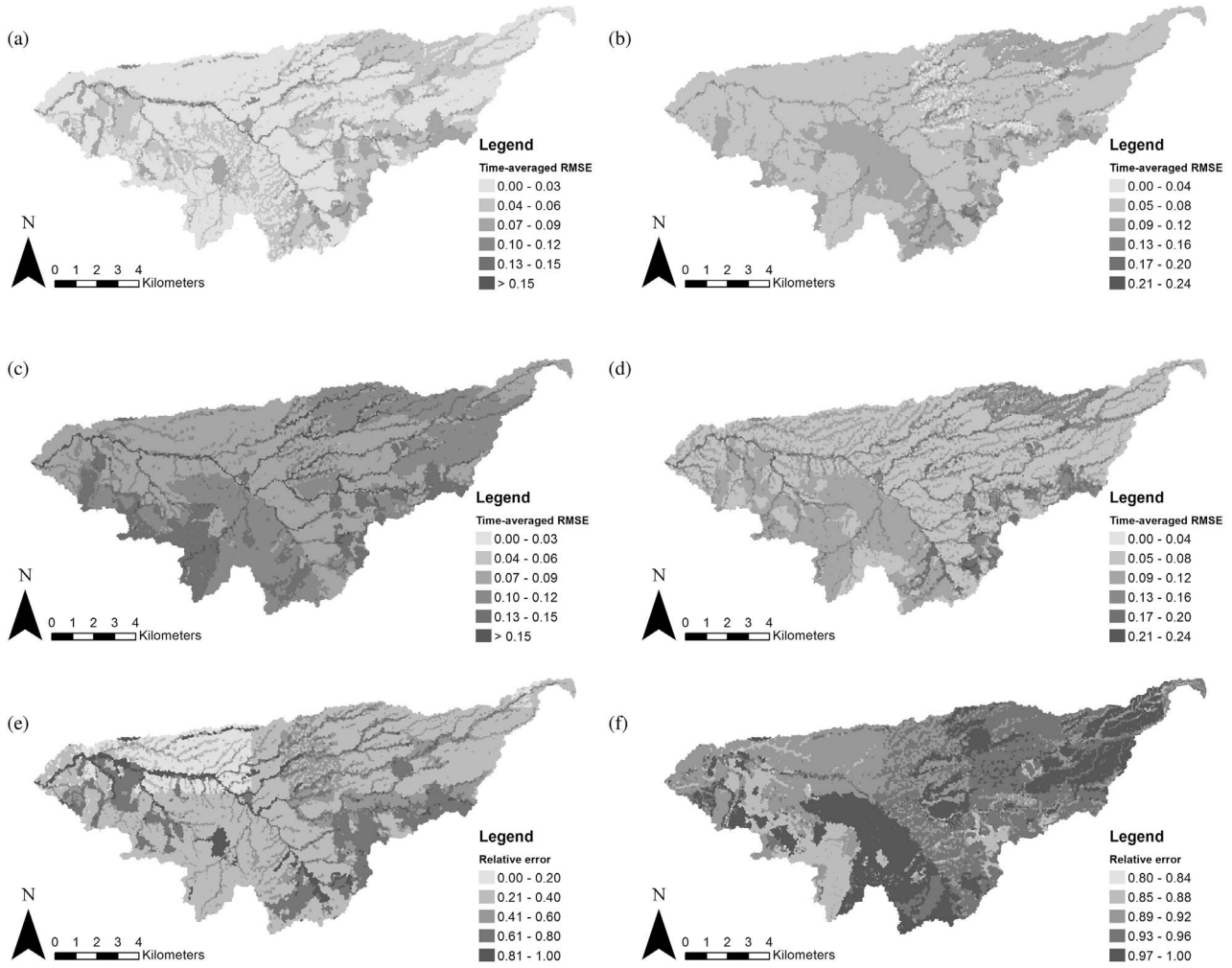


**Figure 5.** RMSE averaged over all four synthetic true experiments and temporally for the duration of the experiment for (a) EnKF-derived near-surface moisture, (b) EnKF-derived profile moisture, (c) OL-derived near-surface moisture, and (d) OL-derived profile moisture. Normalizing RMSE from EnKF experiments by that of OL experiments yields the relative RMSE for (e) near-surface moisture and (f) profile moisture.

observations leads to a significant reduction in bias across all soil texture units and almost all topographic settings present in the watershed.

[37] Bias in the EnKF estimate of profile-integrated soil moisture is significantly higher than the bias in the EnKF estimate of near-surface soil moisture, with a significant portion of the basin exhibiting bias in profile-integrated soil moisture between 0.03 and 0.06 $m^3/m^3$ (Figure 4b). While the spatial arrangement of soil textures seems to explain some of the organization in the bias, it also appears to exhibit variability at the hillslope scale, particularly in the northeast quadrant of the watershed. Interestingly, areas in the south central portion of the basin that are associated with the steepest slopes also seem to exhibit some of the lowest bias in profile-integrated moisture. The bias in the OL profile-integrated moisture throughout the watershed is generally higher than the corresponding EnKF profile-integrated moisture estimate (Figure 4d). Patterns in the spatial distribution of bias in OL profile-integrated moisture are primarily associated with the spatial arrangement of soil units, although valley bottoms as areas of enhanced bias are evident throughout almost the entire watershed. The spatially averaged bias in OL profile-integrated moisture is virtually constant throughout time, exhibiting virtually no sensitivity to rainfall forcings at the surface (Figure 4d). Importantly, the spatially averaged bias in EnKF profile-integrated moisture is initially reduced significantly through data assimilation, but asymptotically approaches the OL bias in profile-integrated moisture.

[38] The spatial distribution of RMSE in EnKF near-surface soil moisture averaged over the nine analyses shows that most of the watershed exhibits an average RMSE of 0.03 $m^3/m^3$ or less, with much of the rest of the watershed exhibiting an RMSE of between 0.03 and 0.06 $m^3/m^3$ (Figure 5a). In comparison, RMSE in OL near-surface moisture is greater than 0.10 $m^3/m^3$ throughout nearly the entire watershed (Figure 5c). Notably, in both the OL and EnKF cases valley bottoms are areas of elevated RMSE. Normalizing the RMSE in near-surface moisture from each EnKF experiment by the RMSE in near-surface moisture

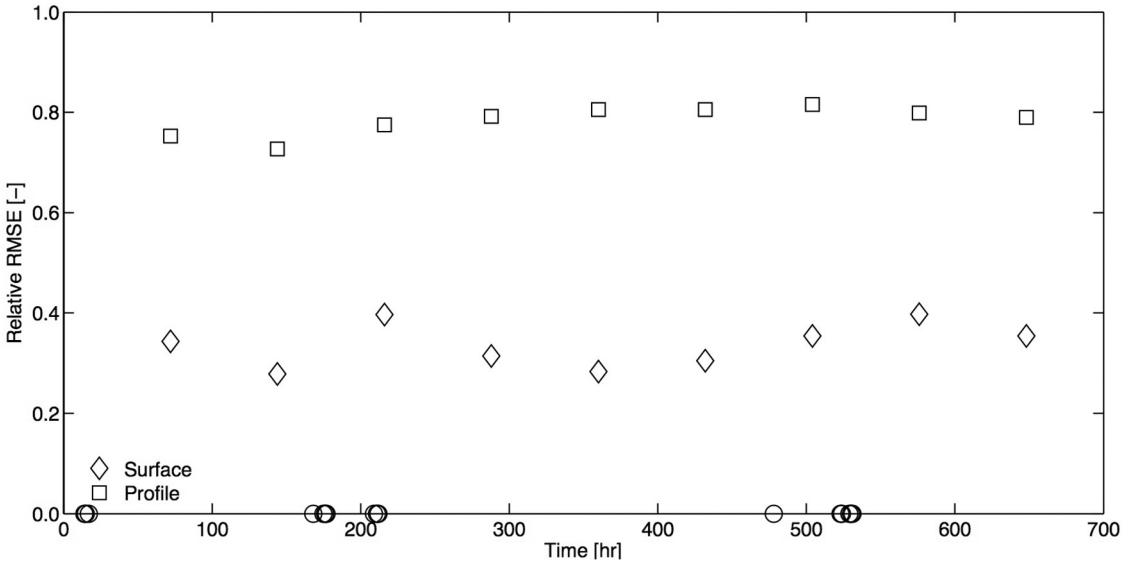


**Figure 6.** The temporal evolution of spatially averaged RMSE from EnKF experiments normalized by the corresponding RMSE from OL experiments. Open circles on the ordinate axis depict the temporal occurrence of rainfall.

from the corresponding OL experiment, and averaging over all four experiments and nine analyses, provides a measure of relative error representing the expected reduction in RMSE attributable to assimilation. Throughout most of the watershed, RMSE in EnKF near-surface moisture is at most 40% of the corresponding RMSE in OL near-surface moisture (Figure 5e). Areas where significant RMSE reductions in near-surface moisture are not realized by assimilating backscatter measurement are primarily associated with soil units that are limited in areal extent and valley bottoms. However, the average relative RMSE exhibits a faint overprinting of the rainfall forcing grid in the northwest region of the watershed in an area where the reduction in RMSE realized through data assimilation is most dramatic. Averaged over the entire watershed, the relative RMSE through time is fairly stable and EnKF RMSE in the near surface is approximately 35% of OL RMSE (Figure 6).

[39] In much of the watershed, RMSE in EnKF profile-integrated moisture averaged over the nine analyses is between 0.05 and 0.08 $m^3/m^3$ (Figure 5b). Spatial contrasts in EnKF profile-integrated moisture RMSE are largely associated with soil texture. The RMSE in OL profile-integrated moisture exhibits similarity with RMSE in EnKF profile-integrated moisture in both the range of variation and spatial pattern, although valley bottoms throughout the watershed tend to be associated with comparatively higher RMSE (Figure 5d). Normalizing the EnKF RMSE for profile-integrated moisture by the corresponding OL RMSE and averaging over OSSEs and analyses, we find that at best data assimilation leads to a 20% reduction in RMSE in profile-integrated moisture (Figure 5f). Moreover, in much of the watershed assimilating the synthetic backscatter observations produces no discernable improvement in predictions of profile-integrated soil moisture. Again, averaged over the entire watershed, the profile-integrated EnKF RMSE is approximately 80% of OL RMSE over the duration of the experiment (Figure 6).

### 3.2. Pixel-Scale Soil Moisture Dynamics

[40] Because the developed data assimilation system is targeted toward applications that require soil moisture information at intervals that may be finer than the satellite revisit, it is important to understand soil moisture dynamics during these intermediate times. To investigate the impacts of data assimilation at times intermediate to analyses, we identify two pixels in the computational domain for more detailed monitoring during the EnKF experiments (Figure 2b). Although these pixels correspond roughly to the location of two intensively monitored sites within the WGEW (the Lucky Hills site, pixel 3378; the Kendall site, pixel 4496) we emphasize the synthetic nature of the true soil properties and hydrometeorological forcings, which disallows any meaningful comparisons with actual data. It should also be underscored that the pixel-scale results presented here are diagnostic only and that no additional assimilation was performed at these sites. Important properties of each monitoring pixel are summarized in Table 4.

[41] For all four EnKF experiments the ensemble mean soil moisture as a function of depth and time varies significantly between each EnKF experiment at both locations (Figures 7c–7j). The ensemble mean soil moisture varies

**Table 4.** Properties of Kendall and Lucky Hills tRIBS-VEGGIE Observation Pixels

| Property | Pixel 3378 (Lucky Hills) | Pixel 4996 (Kendall) |
|---|---|---|
| Soil unit code | 22 (sandy loam) | 17 (loam) |
| Elevation (m) | 1378 | 1529 |
| Slope $\alpha_\nabla$ (m/m) | 0.114 | 0.099 |
| Aspect $\zeta_\nabla$ (deg) | 180° (south facing) | 315° (NW facing) |
| Pixel area $a_p$ ($m^2$) | 11,263 | 10,100 |
| Pixel length scale $\sqrt{a_p}$ (m) | 106.1 | 100.5 |
| Upstream contributing area ($km^2$) | 0.01 | 0.01 |

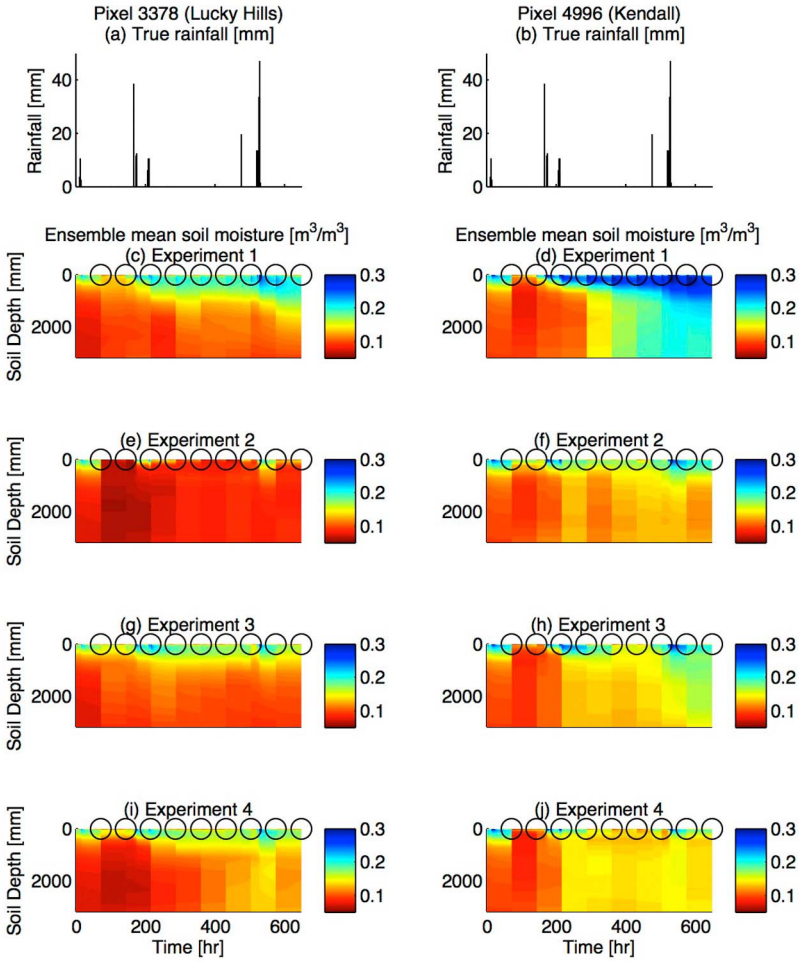


**Figure 7.** (left) Pixel 3378 (Lucky Hills) and (right) pixel 4996 (Kendall). (a, b) True rainfall versus time. Ensemble mean soil moisture is shown versus soil depth and time for (c, d) synthetic experiment 1, (e, f) synthetic experiment 2, (g, h) synthetic experiment 3, and (i, j) synthetic experiment 4. In Figures 7c–7j assimilation times are shown as black circles every 72 h at the surface.

from extremely dry and exhibiting little dynamic variation beyond the near surface (Figure 7e) to being substantially wetter with significant dynamic variation throughout the entire profile (Figure 7d). It is again important to underscore that the uncertain hydrometeorological and soil parameter inputs to the model were the same across all EnKF experiments. Hence, between-experiment contrasts in the spatio-temporal dynamics of ensemble mean soil moisture reflect the influence of particular sets of observations through the successive updates.

[42] As expected, the data assimilation system leads to a progressive reduction in the ensemble standard deviation (visualized for simplicity as the ensemble standard deviation for all depths and times, averaged across all four experiments), $\sigma_{EnKF}$, during the duration of the experiment (Figures 8c and 8d). The first update, which occurs 72 h after the experiments begin, is particularly important in reducing $\sigma_{EnKF}$ (Figures 8c and 8d). Uncertainties in rainfall forcings supplied to the model tends to increase the $\sigma_{EnKF}$ between updates, but subsequent updates are effective in reducing $\sigma_{EnKF}$ to the levels that preceded the rainfall events (Figures 8c and 8d). Similarly, uncertainties in the nonrainfall meteorological forcings to the model lead to an increase in $\sigma_{EnKF}$ in the top ~100 mm, which grows as the time since the reinitialization occurs increases (Figures 8c and 8d). The areas with the lowest $\sigma_{EnKF}$ are below the evaporatively active horizons of the soil and tend to persist throughout the experiment. Importantly, when normalizing $\sigma_{EnKF}$ by the corresponding ensemble standard deviation from the OL experiment, $\sigma_{OL}$, it becomes apparent that the EnKF updates are leading to successive and significant reductions in the $\sigma_{EnKF}$ relative to $\sigma_{OL}$ (Figures 8e and 8f). During and immediately after rainfall events $\sigma_{EnKF}$ is on the order of $\sigma_{OL}$(Figures 8e and 8f).

[43] Because these experiments are synthetic we again find it useful to consider the soil moisture RMSE relative to the synthetic true values and averaged across all experiments. At pixel 3378 (Lucky Hills), RMSE in the top ~500 mm of the soil column decreases substantially from the beginning to the end of the experiment (Figure 9c). Rainfall events of uncertain magnitude tend to increase RMSE, but subsequent analyses are effective at reducing these estimation errors (Figure 9c). By the end of the experiment RMSE is typically less than 0.04 $m^3/m^3$ in the top ~500 mm of soil. Similar

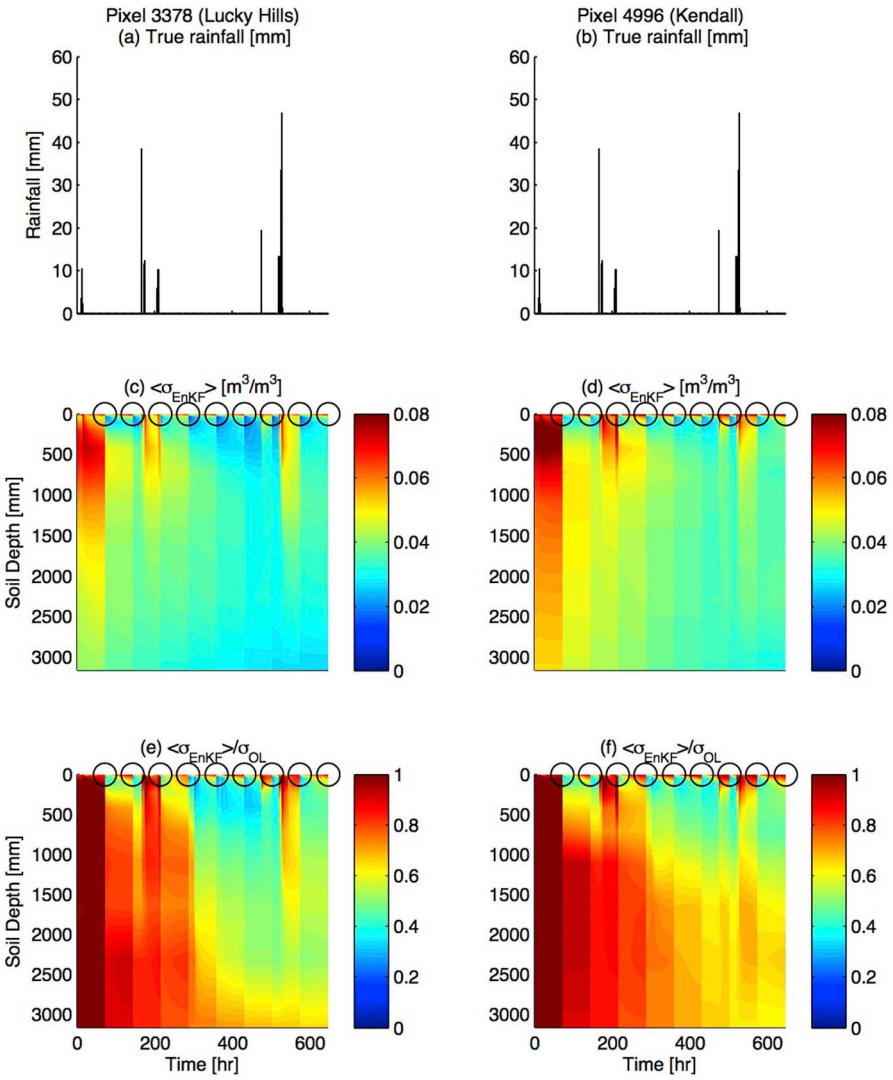


**Figure 8.** (left) Pixel 3378 (Lucky Hills) and (right) pixel 4996 (Kendall). (a, b) True rainfall versus time. (c, d) Ensemble standard deviation in soil moisture, averaged across all four synthetic experiments, versus soil depth and time and (e, f) ensemble standard deviation in soil moisture, averaged across all four synthetic experiments and normalized by the ensemble standard deviation in soil moisture for the open loop simulation. In Figures 8c–8f assimilation times are shown as black circles every 72 h at the surface.

trends are seen at pixel 4996 (Kendall), although the reductions in RMSE are not as persistent between updates when compared with pixel 3378 (Figure 9d). Normalizing RMSE for the EnKF experiments with RMSE for the OL experiment, data assimilation leads to a substantial reduction in RMSE relative to the OL approach in the top ~500–750 mm of the soil column by the end of the experiment (Figures 9e and 9f). However, at depths greater than 500–750 mm there is not a corresponding reduction in ⟨RMSE⟩ relative to the open loop approach (Figures 9e and 9f).

[44] The discrepancies between Figures 8c and 8d (ensemble standard deviation) and Figures 9c and 9d (ensemble RMSE) imply that there exists bias in the estimate of soil moisture, both at the pixel scale and, presumably, more broadly in the watershed throughout the experiment. The spatiotemporal dynamics of bias at the pixel scale are not shown for brevity, but absolute bias (computed as the absolute value of the difference between the ensemble mean soil moisture and synthetic true soil moisture) follows the RMSE normalized by the RMSE for the OL experiment (i.e., Figures 9e and 9f). That is, where the RMSE from the EnKF experiments is on the order of the RMSE from the OL experiments, the bias is correspondingly higher. Conversely, the absolute bias tends to be lower where EnKF experiment RMSE is substantially reduced relative to OL experiment RMSE, between the surface and approximately 750 mm. Potential reasons for this behavior are discussed below.

## 4. Discussion and Conclusions

[45] This narrowly focused proof-of-concept study is meant to develop an algorithmic framework by which coarse-scale observations—in this case 3 km synthetic L band radar data—can be used to inform a hydrologic model capable of resolving hillslope scales through data

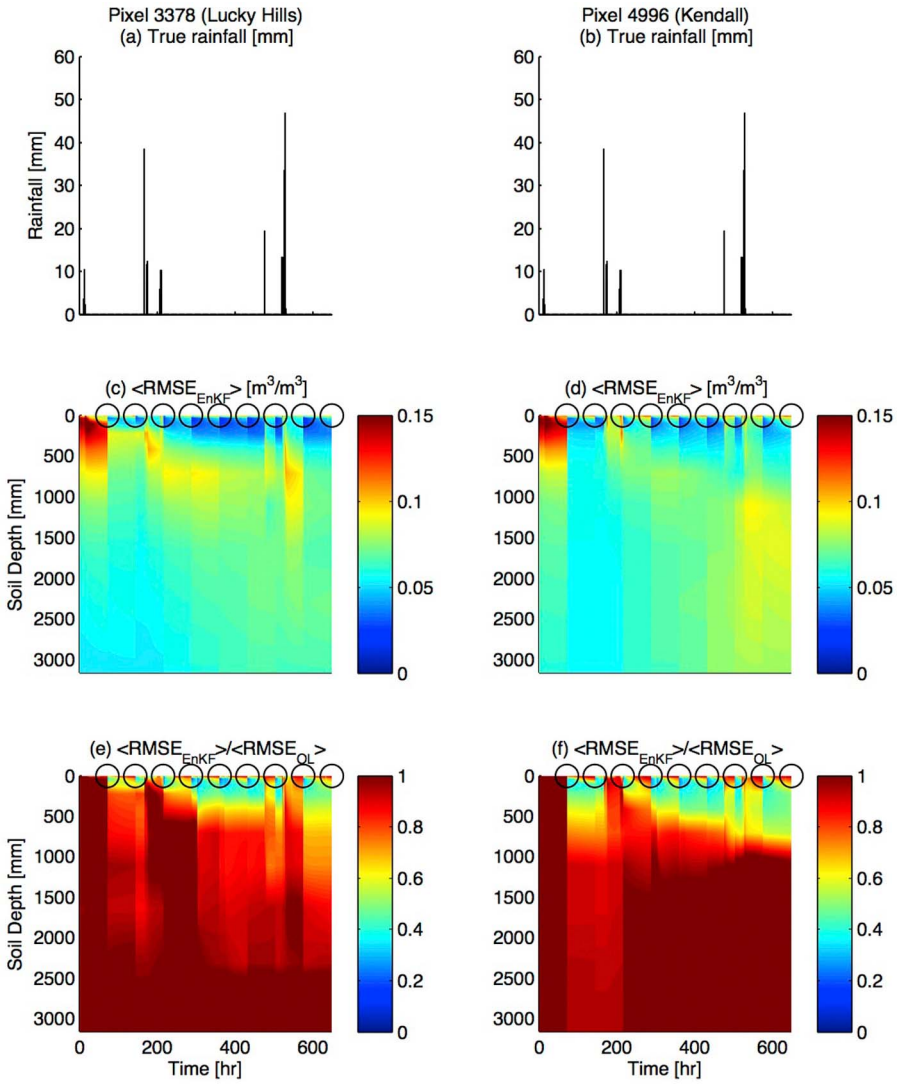


**Figure 9.** (left) Pixel 3378 (Lucky Hills) and (right) pixel 4996 (Kendall). (a, b) True rainfall versus time. (c, d) Ensemble RMSE in soil moisture, averaged across all four synthetic experiments, versus soil depth and time. (e, f) Ensemble RMSE in soil moisture, averaged across all four synthetic experiments and normalized by the ensemble RMSE in soil moisture for the open loop simulation. In Figures 9b–9f assimilation times are shown as black circles every 72 h at the surface.

assimilation. The novelty of the outlined study is threefold: (1) backscatter observations are assimilated directly, (2) downscaling of coarse backscatter observations from 3 km to hillslope-scale soil moisture estimates is achieved via the filter rather than as a preprocessing step, and (3) the filter allows information to propagate from the near surface to all layers of the soil.

[46] Despite the relatively ideal conditions of the experimental setup, there is little improvement in the estimate of total moisture storage in the entire soil profile. Inspection of the pixel-scale moisture dynamics reveals that, relative to the open loop ensemble, significant improvement in estimation of soil moisture from the surface through soil depths of approximately 500–750 mm is seen, followed by a significant decay in predictive skill at greater depth (e.g., Figures 9e and 9f). We presume the results at these pixels to be broadly indicative of behavior throughout the rest of the watershed and arise for a number of important and interrelated reasons. The total thickness of the simulated soil column (~3 m) may, in part, explain the lack of improvement in profile-integrated soil moisture relative to improvement in near-surface moisture, particularly if the ensemble spread is larger in deeper layers. The assumed depth is likely unreasonably thick for WGEW and many other desert settings, and a more realistic depth (and distribution) of soil thicknesses will be used in future experiments. The significant improvement seen at depths up to 500–750 mm suggests that perhaps profile-integrated moisture estimation with L band backscatter is a reasonable objective in thinly soil-mantled landscapes. The lack of improvement in estimation of profile-integrated soil moisture under the idealized experimental setup, however, may also arise from the role of process dynamics in vertical moisture redistribution. Soil moisture in this arid hydroclimatic setting varies most dynamically in the near surface, which is correspondingly the region to which radar backscatter is most sensitive. The

dynamic decoupling between soil moisture in the near surface and in deeper layers may induce relatively small and perhaps spurious correlations between soil moisture at depth and the predicted observations. Analysis increments computed in the EnKF update may not appreciably, therefore, improve soil moisture profile predictions at depth. It may also be that the short experiment length was not sufficient to capture events that are sufficiently extreme to establish hydrologic connectivity throughout the entire soil column or to allow the development of spatial patterns and correlations in soil moisture at substantial depths in the soil column. Future experiments will test this hypothesis and, if confirmed, it may underscore the importance of continuity in data assimilation systems and in the observations themselves. In heavily managed or grazed lands, moreover, vertical decoupling of soil moisture may arise from heterogeneity in soil properties. In such settings it is unlikely that even long experiments will allow for dramatic improvements in estimation of profile-integrated moisture without additional observational constraint.

[47] Despite the lack of improvement in the prediction of profile-integrated soil moisture, the inspection of the pixel-scale dynamics indicates some improvement in predictability of soil moisture at depths greater than the penetration depth of the radar. Future study will be devoted to understanding the factors (e.g., vegetation, soils, hydrometeorology) that control the depth of soil where predictability can be significantly improved through assimilation of radar (or radiometer) measurements with relatively shallow penetration depths. This would necessitate defining what constitutes a significant improvement, which may depend on the application of interest. Such a study would probably benefit from a more realistic treatment of the soils, vegetation, and hydrometeorological conditions in the study area, for example, using measured quantities to parameterize the hydrologic model. Moreover, actual radar observations from instruments like the Passive and Active L- and S-band (PALS) instrument used in a targeted field campaign or reanalysis of existing field campaign data (e.g., the North American Monsoon Experiment [*Higgins and Gochis*, 2007]) would benefit any such study.

[48] There are attributes of the OSSEs presented here that potentially complicate interpretation of the results presented here and should be addressed in future studies. Specifically, the way in which the true soil moisture and synthetic observations are produced in this study could be causing biased observations to be assimilated, a fundamental violation of the Kalman Filter assumptions. And because there was only one large OL simulation, there may be considerable bias between the soil moisture statistics of the EnKF and OL ensembles. A more robust approach and only a slight modification of our experimental setup would produce the soil parameter and hydrometeorological forcings using the same methods, but use the ensemble mean soil moisture from an OL simulation as the synthetic true from which synthetic observations would be produced. A more robust methodology for the multiple-truth OSSE would then extend this same procedure multiple times for different soil parameter and hydrometeorological ensembles. A subsequent or alternative experimental step would use a bias estimation (e.g., *de Lannoy et al.* [2006, 2007]) or CDF-matching approach to remove bias in the observations prior to the Kalman update.

[49] In this study the soil parameters are not simultaneously updated with the model state. As a result the developed data assimilation system is inherently suboptimal and RMSE and bias in soil moisture predictions at soil depths greater than 500–750 mm remain high. There are several avenues through which this may be addressed in future studies. State augmentation is a historically popular way to jointly estimate the states and parameters of hydrologic models with data assimilation strategies [e.g., *Bras and Restrepo-Posada*, 1980]. *Vrugt et al.* [2005] incorporated the strengths of both global optimization techniques and data assimilation into a single framework to simultaneously estimate the states and parameters of a lumped hydrological model. *Moradkhani et al.* [2005b] employed an EnKF to first update the parameters of the HyMOD model, propagated the model state using the updated parameters, and updated the model state using another EnKF. Also using the HyMOD model, *Moradkhani et al.* [2005a] used a particle filtering approach to retrieve both states and parameters of the HyMOD conceptual model and point out some important limitations to Kalman update-based data assimilation procedures. More recently, *Montzka et al.* [2011] demonstrated the utility of a particle filtering approach to estimate hydraulic parameters of the HYDRUS 1-D model using synthetic microwave soil moisture retrievals.

[50] Unique to this study are the implications of the data assimilation process for estimation of soil moisture with hydrologic models capable of resolving hillslopes and coarse L band backscatter observations. This particular set of experiments in this hydroclimatic setting seems to suggest that assimilation of SMAP data is most useful in combating uncertainties in the hydrometeorological forcings and soil parameters input to the model in the near-surface environment. It is plausible that assimilation of L band radar data could improve predictions of near-surface moisture in steeper regions where the spatial distribution of near-surface moisture is characterized by hillslope-scale variation to a greater degree than observed in this study. However, topographic effects on radar aperture synthesis and satellite geolocation errors may necessitate substantial preprocessing and orthorectification of the raw radar data prior to assimilation [*Shimada*, 2010]. We are preparing to apply the data assimilation system in the Reynolds Creek Experimental Watershed and Dry Creek Experimental Watershed in southwest Idaho to investigate this further.

[51] The OSSE approach necessitated simplifying assumptions to reduce the complexity of the problem. One particularly important assumption is that the surface roughness (equation (1)) and the length scale of the roughness correlation function (equations (8)) are known. Characteristics of surface roughness are well constrained in relatively few locations because, in part, labor intensive technologies like laser scanners or pin meters (see *Bryant et al.* [2007] for a discussion of surface roughness measurement) are required to measure them directly. Although some recent work suggests that surface roughness can be retrieved from satellite or airborne microwave imagery, these estimates are uncertain and associated with spatial scales larger than hillslopes [*Shi et al.*, 1997; *Rahman et al.*, 2007, 2008]. For the

foreseeable future, therefore, it will be necessary to obtain surface roughness characteristics through calibration or by extrapolating statistical relationships between measured roughness and spatially extensive landscape properties such as soil texture. Future work will also investigate the potentially important effects of scattering due to water stored in the canopy through the use of forward models that explicitly account for vegetation water content [*Bindlish and Barros*, 2001].

[52] **Acknowledgments.** This work was supported by U.S. Army RDECOM ARL Army Research Office under grants W911NF-04-1-0119, W911NF-09-1-0534, and W911NF-11-1-0310, NASA grants NNG05GA17G, NNX10AG84G, and NNX11AQ33G, and the MIT Martin Family Society of Fellows for Sustainability. The first author would like to thank Shawn Benner and Jim McNamara at Boise State University for their comments on the introduction and the associate editor and three anonymous reviewers for their helpful comments.

## References

Altese, E., O. Bolognani, M. Mancini, and P. A. Troch (1996), Retrieving soil moisture over bare soil from ERS 1 synthetic aperture radar data: Sensitivity analysis based on a theoretical surface scattering model and field data, *Water Resour. Res.*, *32*(3), 653–661, doi:10.1029/95WR03638.

Bindlish, R., and A. P. Barros (2001), Parameterization of vegetation backscatter in radar-based, soil moisture estimation, *Remote Sens. Environ.*, *76*(1), 130–137, doi:10.1016/S0034-4257(00)00200-5.

Bovolo, C. I., S. J. Abele, J. C. Bathurst, D. Caballero, M. Ciglan, G. Eftichidis, and B. Simo (2009), A distributed framework for multi-risk assessment of natural hazards used to model the effects of forest fire on hydrology and sediment yield, *Comput. Geosci.*, *35*, 924–945, doi:10.1016/j.cageo.2007.10.010.

Bras, R. L., and P. Restrepo-Posada (1980), Real time automatic parameter calibration in conceptual runoff forecasting models, in *Proceedings of the Third International Symposium on Stochastic Hydraulics*, pp. 61–70, Organizing Comm. of the Third Int. Symp. on Stochastic Hydraulics, Tokyo, Japan.

Brunet, P., R. Clément, and C. Bouvier (2010), Monitoring soil water content and deficit using electrical resistivity tomography (ERT)—A case study in the Cevennes area, France, *J. Hydrol.*, *380*(1–2), 146–153, doi:10.1016/j.jhydrol.2009.10.032.

Bryant, R., et al. (2007), Measuring surface roughness height to parameterize radar backscatter models for retrieval of surface soil moisture, *IEEE Geosci. Remote Sens. Lett.*, *4*(1), 137–141, doi:10.1109/LGRS.2006.887146.

Camporese, M., C. Paniconi, M. Putti, and S. Orlandini (2010), Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data, *Water Resour. Res.*, *46*, W02512, doi:10.1029/2008WR007536.

Chen, F., et al. (1996), Modeling of land-surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res.*, *101*(D3), 7251–7268, doi:10.1029/95JD02165.

Chen, K. S., S. K. Yen, and W. P. Huang (1995), A simple model for retrieving bare soil moisture from radar-scattering coefficients, *Remote Sens. Environ.*, *54*, 121–126, doi:10.1016/0034-4257(95)00129-O.

Crow, W. T., and D. Ryu (2009), A new data assimilation approach for improving runoff prediction using remotely sensed soil moisture retrievals, *Hydrol. Earth Syst. Sci.*, *13*(1), 1–16, doi:10.5194/hess-13-1-2009.

Crow, W. T., and E. F. Wood (2003), 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. Water Resour.*, *26*(2), 137–149, doi:10.1016/S0309-1708(02)00088-X.

Crow, W. T., M. Drusch, and E. F. Wood (2001), An observation system simulation experiment for the impact of land surface heterogeneity on AMSR-E soil moisture retrieval, *IEEE Trans. Geosci. Remote Sens.*, *39*(8), 1622–1631, doi:10.1109/36.942540.

De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest (2006), State and bias estimation for soil moisture profiles by an ensemble Kalman filter: effect of assimilation depth and frequency, *Water Resour. Res.*, *43*, W06401, doi:10.1029/2006WR005100.

De Lannoy, G. J. M., R. H. Reichle, P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest (2007), Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, *Water Resour. Res.*, *43*, W09410, doi:10.1029/2006WR005449.

Dobson, M. C., and F. T. Ulaby (1986), Preliminary evaluation of the SIR-B response to soil moisture, surface roughness, and crop canopy, *IEEE Trans. Geosci. Remote Sens.*, *24*(4), 510–516.

Dubois, P. C., J. van Zyl, and T. Engman (1995), Measuring soil moisture with imaging radars, *IEEE Trans. Geosci. Remote Sens.*, *33*(4), 915–926, doi:10.1109/36.406677.

Dunne, S., and D. Entekhabi (2005), An ensemble-based reanalysis approach to land data assimilation, *Water Resour. Res.*, *41*, W02013, doi:10.1029/2004WR003449.

Dunne, S., and D. Entekhabi (2006), Land surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment, *Water Resour. Res.*, *42*, W01407, doi:10.1029/2005WR004334.

Engman, E. T. (1991), Applications of microwave remote sensing of soil moisture for water resources and agriculture, *Remote Sens. Environ.*, *35*, 213–226, doi:10.1016/0034-4257(91)90013-V.

Entekhabi, D., et al. (2010), The Soil Moisture Active Passive (SMAP) Mission, *Proc. IEEE*, *98*(5), 704–716, doi:10.1109/JPROC.2010.2043918.

Eom, H. J., and W.-M. Boener (1986), Scattering from a layered medium connected with rough interfaces: Matrix doubling method, *IEEE Trans. Geosci. Remote Sens.*, *24*(6), 937–939, doi:10.1109/TGRS.1986.289709.

Evans, D. L., T. G. Farr, and J. J. van Zyl (1992), Estimates of surface roughness derived from synthetic aperture radar (SAR) data, *IEEE Trans. Geosci. Remote Sens.*, *30*(2), 382–389, doi:10.1109/36.134087.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10,143–10,162, doi:10.1029/94JC00572.

Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, *53*, 343–367, doi:10.1007/s10236-003-0036-9.

Evensen, G. (2004), Sampling strategies and square root analysis schemes for the EnKF, *Ocean Dyn.*, *54*, 539–560, doi:10.1007/s10236-004-0099-2.

Ferraris, L., S. Gabellani, N. Rebora, and A. Provenzale (2003), A comparison of stochastic models for spatial rainfall downscaling, *Water Resour. Res.*, *39*(12), 1368, doi:10.1029/2003WR002504.

Flores, A. N., V. Y. Ivanov, D. Entekhabi, and R. L. Bras (2009), Impacts of hillslope-scale organization in topography, soil moisture, soil temperature, and vegetation on modeling surface microwave radiation emission, *IEEE Trans. Geosci. Remote Sens.*, *47*(8), 2557–2571, doi:10.1109/TGRS.2009.2014743.

Flores, A. N., D. Entekhabi, and R. L. Bras (2010), Reproducibility of soil moisture ensembles when representing soil parameter uncertainty and correlation using a Latin hypercube–based approach, *Water Resour. Res.*, *46*, W04506, doi:10.1029/2009WR008155.

Fung, A. K. (1994), *Microwave Scattering and Emission Models and Their Applications*, Artech House, Boston, Mass.

Fung, A. K., Z. Li, and K. S. Chen (1992), Backscattering from a randomly rough dielectric surface, *IEEE Trans. Geosci. Remote Sens.*, *30*(2), 356–369, doi:10.1109/36.134085.

Gessler, P. E., O. A. Chamran, F. Althouse, and L. Holmes (2000), Modeling soil-landscape and ecosystem properties using terrain attributes, *Soil Sci. Soc. Am. J.*, *64*(6), 2046–2056, doi:10.2136/sssaj2000.6462046x.

Grayson, R. B., A. W. Western, F. H. S. Chiew, and G. Bloschl (1997), Preferred states in spatial soil moisture patterns: Local and nonlocal controls, *Water Resour. Res.*, *33*(12), 2897–2908, doi:10.1029/97WR02174.

Hawk, K. (1992), Climatology of station storm rainfall in the continental United States: Parameters of the Bartlett-Lewis and Poisson rectangular pulses models, MS thesis, Mass. Inst. of Technol., Cambridge.

Higgins, W., and D. Gochis (2007), Synthesis of results from the North American Monsoon Experiment (NAME) process study, *J. Clim.*, *20*(9), 1601–1607, doi:10.1175/JCLI4081.1.

Hoeben, R., and P. A. Troch (2000), Assimilation of active microwave observation data for soil moisture profile estimation, *Water Resour. Res.*, *36*, 2805–2819, doi:10.1029/2000WR900100.

Islam, S., D. Entekhabi, R. L. Bras, and I. Rodriguez-Iturbe (1990), Parameter estimation and sensitivity analysis for the modified Bartlett-Lewis rectangular pulses model of rainfall, *J. Geophys. Res.*, *95*(D3), 2093–2100, doi:10.1029/JD095iD03p02093.

Ivanov, V. Y., E. R. Vivoni, R. L. Bras, and D. Entekhabi (2004a), Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: A fully distributed physically based approach, *J. Hydrol.*, *298*(1–4), 80–111, doi:10.1016/j.jhydrol.2004.03.041.

Ivanov, V. Y., E. R. Vivoni, R. L. Bras, and D. Entekhabi (2004b), Catchment hydrologic response with a fully distributed triangulated irregular network model, *Water Resour. Res.*, *40*, W11102, doi:10.1029/2004WR003218.

Ivanov, V. Y., R. L. Bras, and D. C. Curtis (2007), A weather generator for hydrological, ecological, and agricultural applications, *Water Resour. Res.*, *43*, W10406, doi:10.1029/2006WR005364.

Ivanov, V. Y., R. L. Bras, and E. R. Vivoni (2008a), Vegetation-hydrology dynamics in complex terrain of semi-arid areas: 1. A mechanistic approach to modeling dynamic feedbacks, *Water Resour. Res.*, *44*, W03429, doi:10.1029/2006WR005588.

Ivanov, V. Y., R. L. Bras, and E. R. Vivoni (2008b), Vegetation-hydrology dynamics in complex terrain of semi-arid areas: 2. Energy-water controls of vegetation spatiotemporal dynamics and topographic niches of favorability, *Water Resour. Res.*, *44*, W03430, doi:10.1029/2006WR005595.

Jencso, K. G., B. L. McGlynn, M. N. Gooseff, S. M. Wondzell, K. E. Bencala, and L. A. Marshall (2009), Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale, *Water Resour. Res.*, *45*, W04428, doi:10.1029/2008WR007225.

Kashif Gill, M., M. W. Kemblowski, and M. McKee (2007), Soil moisture data assimilation using support vector machines and ensemble Kalman filter, *J. Am. Water Resour. Assoc.*, *43*(4), 1004–1015, doi:10.1111/j.1752-1688.2007.00082.x.

Kerr, Y. H., P. Waldteufel, J.-P. Wigneron, J.-M. Martinuzzi, J. Font, and M. Berger (2001), Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission, *IEEE Trans. Geosci. Remote Sens.*, *39*(8), 1729–1735, doi:10.1109/36.942551.

Kerr, Y., F. Secherre, J. Lastenet, and J.-P. Wigneron (2003), SMOS: Analysis of perturbing effects over land surfaces, in *IGARSS '03: Proceedings of the 2003 IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, pp. 908–910, IEEE Press, Piscataway, N. J.

Kim, S. (2009), Characterization of soil moisture responses on a hillslope to sequential rainfall events during late autumn and spring, *Water Resour. Res.*, *45*, W09425, doi:10.1029/2008WR007239.

Kollet, S. J., R. M. Maxwell, C. S. Woodward, S. Smith, J. Vanderborght, H. Vereecken, and C. Simmer (2010), Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources, *Water Resour. Res.*, *46*, W04201, doi:10.1029/2009WR008730.

Koster, R. D., and M. J. Suarez (1996), Energy and water balance calculations in the MOSAIC LSM, *NASA Tech. Memo., TM-104606*, *9*, 76 pp.

Kumar, M., C. J. Duffy, K. M. Salvage, and M. Karen (2009), A second-order accurate, finite volume-based, integrated hydrologic modeling (FIHM) framework for simulation of surface and subsurface flow, *Vadose Zone J.*, *8*, 873–890, doi:10.2136/vzj2009.0014.

Larson, K. M., J. J. Braun, E. E. Small, V. U. Zavorotny, E. D. Gutmann, and A. L. Bilich (2010), GPS multipath and its relation to near-surface soil moisture content, *IEEE J. Sel. Top. Appl. Remote Sens.*, *3*(1), 91–99, doi:10.1109/JSTARS.2009.2033612.

Leisenring, M., and H. Moradkhani (2011), Snow water equivalent prediction using Bayesian data assimilation methods, *Stochastic Environ. Res. Risk Assess.*, *25*, 253–270.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for GCMs, *J. Geophys. Res.*, *99*(D7), 14,415–14,428, doi:10.1029/94JD00483.

Livings, D. M., S. L. Dance, and N. K. Nichols (2008), Unbiased ensemble square root filters, *Physica D*, *237*(8), 1021–1028, doi:10.1016/j.physd.2008.01.005.

Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne (2002), Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, *Water Resour. Res.*, *38*(12), 1299, doi:10.1029/2001WR001114.

Margulis, S. A., D. Entekhabi, and D. McLaughlin (2006), Spatiotemporal disaggregation of remotely sensed precipitation for ensemble hydrologic modeling and data assimilation, *J. Hydrometeorol.*, *7*(3), 511–533, doi:10.1175/JHM492.1.

Mätzler, C., and A. Standley (2000), Technical note: Relief effects for passive microwave remote sensing, *Int. J. Remote Sens.*, *21*(12), 2403–2412, doi:10.1080/01431160050030538.

Maxwell, R. M., and N. L. Miller (2005), Development of a coupled land surface and groundwater model, *J. Hydrometeorol.*, *6*(3), 233–247, doi:10.1175/JHM422.1.

Merlin, O., A. Chehbouni, G. Boulet, and Y. Kerr (2006), Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data, *J. Hydrometeorol.*, *7*(6), 1308–1322, doi:10.1175/JHM552.1.

Mialon, A., L. Coret, Y. H. Kerr, F. Secherre, and J.-P. Wigneron (2008), Flagging the topographic impact on the SMOS signal, *IEEE Trans. Geosci. Remote Sens.*, *46*(3), 689–694, doi:10.1109/TGRS.2007.914788.

Montzka, C., H. Moradkhani, L. Weihermuller, H. J. H. Franssen, M. Canty, and H. Vereecken (2011), Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, *J. Hydrol.*, *399*(1–2), 410–421, doi:10.1016/j.jhydrol.2011.01.020.

Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser (2005a), Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, *28*(2), 135–147, doi:10.1016/j.advwatres.2004.09.002.

Moradkhani, H., K.-L. Hsu, H. Gupta, and S. Sorooshian (2005b), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, *41*, W05012, doi:10.1029/2004WR003604.

Moran, M. S., et al. (2008), Preface to the special section on Fifty Years of Research and Data Collection: U.S. Department of Agriculture Walnut Gulch Experimental Watershed, *Water Resour. Res.*, *44*, W05S01, doi:10.1029/2007WR006083.

Njoku, E., and D. Entekhabi (1996), Passive microwave remote sensing of soil moisture, *J. Hydrol.*, *184*, 101–129, doi:10.1016/0022-1694(95)02970-2.

Njoku, E. G., and J.-A. Kong (1977), Theory for passive microwave remote sensing of near-surface soil moisture, *J. Geophys. Res.*, *82*(20), 3108–3118, doi:10.1029/JB082i020p03108.

Njoku, E. G., W. J. Wilson, S. H. Yueh, S. J. DiNardo, F. K. Li, T. J. Jackson, V. Lakshmi, and J. Bolten (2002), Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99, *IEEE Trans. Geosci. Remote Sens.*, *40*(12), 2659–2673, doi:10.1109/TGRS.2002.807008.

Oleson, K., et al. (2004), Technical description of the Community Land Model (CLM), *Tech. Note NCAR/TN-461+STR*, Natl. Cent. for Atmos. Res., Boulder, Colo.

Paloscia, S., G. Macelloni, and E. Santi (2006), Soil moisture estimates from AMSR-E brightness temperatures by using a dual-frequency algorithm, *IEEE Trans. Geosci. Remote Sens.*, *44*(11), 3135–3144, doi:10.1109/TGRS.2006.881714.

Pan, M., and E. F. Wood (2010), Impact of accuracy, spatial availability, and revisit time of satellite-derived surface soil moisture in a multiscale ensemble data assimilation system, *IEEE J. Sel. Top. Appl. Remote Sens.*, *3*(1), 49–56, doi:10.1109/JSTARS.2010.2040585.

Parada, L. M., and X. Liang (2008), Impacts of spatial resolutions and data quality on soil moisture data assimilation, *J. Geophys. Res.*, *113*(D10), D10101, doi:10.1029/2007JD009037.

Pellarin, T., J.-P. Wigneron, J.-C. Calvet, and P. Waldteufel (2003), Global soil moisture retrieval from a synthetic L-band brightness-temperature data set, *J. Geophys. Res.*, *108*(D12), 4364, doi:10.1029/2002JD003086.

Peters-Lidard, C. D., M. S. Zion, and E. F. Wood (1997), A soil-vegetation-atmosphere transfer scheme for modeling spatially variable water and energy balance processes, *J. Geophys. Res.*, *102*, 4303–4324, doi:10.1029/96JD02948.

Qu, Y., and C. J. Duffy (2007), A semidiscrete finite volume formulation for multiprocess watershed simulation, *Water Resour. Res.*, *43*, W08419, doi:10.1029/2006WR005752.

Rahman, M. M., et al. (2007), A derivation of roughness correlation length for parameterizing radar backscatter models, *Int. J. Remote Sens.*, *28*(18), 3995–4012, doi:10.1080/01431160601075533.

Rahman, M. M., et al. (2008), Mapping surface roughness and soil moisture using multi-angle radar imagery without ancillary data, *Remote Sens. Environ.*, *112*(2), 391–402, doi:10.1016/j.rse.2006.10.026.

Reichle, R. H., D. Entekhabi, and D. B. McLaughlin (2001), Downscaling of radiobrightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach, *Water Resour. Res.*, *37*, 2353–2364, doi:10.1029/2001WR000475.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi (2002), Hydrologic data assimilation with the ensemble Kalman filter, *Mon. Weather Rev.*, *130*, 103–114, doi:10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2.

Reichle, R. H., W. T. Crow, and C. L. Keppenne (2008), An adaptive ensemble Kalman filter for soil moisture data assimilation, *Water Resour. Res.*, *44*, W03423, doi:10.1029/2007WR006357.

Sakov, P., and P. R. Oke (2008), Implications of the form of the ensemble transformation in the ensemble square root filters, *Monthly Weather Rev.*, *136*(3), 1042–1053.

Sandells, M. J., I. J. Davenport, and R. J. Gurney (2008), Passive L-band microwave soil moisture retrieval error arising from topography in otherwise uniform scenes, *Adv. Water Resour.*, *31*, 1433–1443, doi:10.1016/j.advwatres.2008.01.012.

Schaap, M. G., and F. J. Leij (1998), Database related accuracy and uncertainty of pedotransfer functions, *Soil Sci.*, *163*, 765–779, doi:10.1097/00010694-199810000-00001.

Selker, J. S., L. Thévenaz, H. Huwald, A. Mallet, W. Luxemburg, N. van de Giesen, M. Stejskal, J. Zeman, M. Westhoff, and M. B. Parlange (2006), Distributed fiber-optic temperature sensing for hydrologic systems, *Water Resour. Res.*, *42*, W12202, doi:10.1029/2006WR005326.

Seyfried, M. S., and B. P. Wilcox (1995), Scale and the nature of spatial variability: Field examples having implications for hydrologic modeling, *Water Resour. Res.*, *31*(1), 173–184, doi:10.1029/94WR02025.

Shi, J., J. Wang, A. Y. Hsu, P. E. O'Neill, and E. T. Engman (1997), Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data, *IEEE Trans. Geosci. Remote Sens.*, *35*(5), 1254–1266, doi:10.1109/36.628792.

Shimada, M. (2010), Ortho-rectification and slope correction of SAR data using DEM and its accuracy evaluation, *IEEE J. Sel. Top. Appl. Remote Sens.*, *3*(4), 657–671, doi:10.1109/JSTARS.2010.2072984.

Tabatabaeenejad, A., and M. Moghaddam (2006), Bistatic scattering from three-dimensional layered rough surfaces, *IEEE Trans. Geosci. Remote Sens.*, *44*(8), 2102–2114, doi:10.1109/TGRS.2006.872140.

Topp, G. C., J. L. Davis, and A. P. Annan (1980), Electromagnetic determination of soil water content: Measurements in coaxial transmission lines, *Water Resour. Res.*, *16*, 574–582, doi:10.1029/WR016i003p00574.

Tyler, S. W., J. S. Selker, M. B. Hausner, C. E. Hatch, T. Torgersen, C. E. Thodal, and S. G. Schladow (2009), Environmental temperature sensing using Raman spectra DTS fiber-optic methods, *Water Resour. Res.*, *45*, W00D23, doi:10.1029/2008WR007052.

Ulaby, F. T., R. K. Moore, and A. K. Fung (1986), *Microwave Remote Sensing*, vols. *1–3*, 2162 pp., Artech House, Norwood, Mass.

van Loon, E. E., and P. A. Troch (2002), Tikhonov regularization as a tool for assimilating soil moisture data in distributed hydrological models, *Hydrol. Processes*, *16*(2), 531–556, doi:10.1002/hyp.352.

Vivoni, E. R., V. Y. Ivanov, R. L. Bras, and D. Entekhabi (2004), Generation of triangulated irregular networks based on hydrological similarity, *J. Hydrol. Eng.*, *27*, 959–973.

Vrugt, J. A., C. H. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*, W01017, doi:10.1029/2004WR003059.

Wang, J. R., E. T. Engman, J. C. Shiue, M. Rusek, and C. Steinmeier (1986), The SIR-B observations of microwave dependence on soil moisture, surface roughness, and vegetation covers, *IEEE Trans. Geosci. Remote Sens.*, *24*(4), 510–516, doi:10.1109/TGRS.1986.289665.

Wang, X., C. H. Bishop, and S. J. Julier (2004), Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble?, *Mon. Weather Rev.*, *132*, 1590–1605, doi:10.1175/1520-0493(2004)132<1590:WIBAEO>2.0.CO;2.

Wigneron, J.-P., P. Waldteufel, A. Chanzy, J.-C. Calvet, and Y. Kerr (2000), Two-dimensional microwave interferometer retrieval capabilities over land surfaces (SMOS mission), *Remote Sens. Environ.*, *73*(3), 270–282, doi:10.1016/S0034-4257(00)00103-6.

Williams, C. J., J. P. McNamara, and D. G. Chandler (2008), Controls on the temporal and spatial variability of soil moisture in a mountainous landscape: The signatures of snow and complex terrain, *Hydrol. Earth Syst. Sci. Discuss.*, *5*, 1927–1966, doi:10.5194/hessd-5-1927-2008.

Zreda, M., D. Desilets, T. P. A. Ferré, and R. L. Scott (2008), Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons, *Geophys. Res. Lett.*, *35*, L21402, doi:10.1029/2008GL035655.