Advances in Water Resources 33 (2010) 468–484

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Effect of simultaneous state–parameter estimation and forcing uncertainties on root-zone soil moisture for dynamic vegetation using EnKF

Alejandro Monsivais-Huertero [a,*], Wendy D. Graham [b], Jasmeet Judge [a], Divya Agrawal [a]

[a] Center for Remote Sensing, Agricultural and Biological Engineering Department, The University of Florida, P.O. Box 110570, Gainesville, FL 32611, USA
[b] Water Institute, The University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 7 July 2009
Received in revised form 19 January 2010
Accepted 29 January 2010
Available online 13 February 2010

Keywords:
Root-zone soil moisture
SVAT-vegetation models
Ensemble Kalman Filter

ABSTRACT

In this study, an EnKF-based assimilation algorithm was implemented to estimate root-zone soil moisture (RZSM) using the coupled LSP–DSSAT model during a growing season of corn. Experiments using both synthetic and field observations were conducted to understand effects of simultaneous state–parameter estimation, spatial and temporal update frequency, and forcing uncertainties on RZSM estimates. Estimating the state–parameters simultaneously with every 3-day assimilation of volumetric soil moisture (VSM) observations at 5 depths lowered the average standard deviation (ASD) and the root mean square error (RMSE) for RZSM by approximately 1.77% VSM (78%) and 2.18% VSM (93%), respectively, compared to the open-loop ASD where as estimating only states lowered the ASD by approximately 1.26% VSM (56%) and the RMSE by 1.66% VSM (71%). The synthetic case obtained RZSM estimates closer to the observations than the MicroWEX-2 case, particularly after precipitation/irrigation events. The differences in EnKF performance between MicroWEX-2 and synthetic observations may indicate other sources of errors in addition to those in parameters and forcings, such as errors in model biophysics.

1. Introduction

Accurate knowledge of root-zone soil moisture is crucial in hydrology, micrometeorology, and agriculture [25] for estimating energy and moisture fluxes at the land surface. Soil moisture plays a significant role in the partitioning of available energy at the ground surface into sensible and latent heat, as well as in partitioning of rainfall into infiltration and runoff [9,57]. Soil Vegetation Atmosphere Transfer (SVAT) models are used to simulate energy and moisture transport in soil and vegetation, and estimate these fluxes at the land surface and in the root zone. The interactions between the fluxes and vegetation become increasingly important as the vegetation grows. Most SVAT models rely on observations or empirical functions to simulate the effects of growing vegetation on land surface models. Coupling an SVAT model with a vegetation growth model allows inclusion of canopy effects without relying on these observations or empirical functions [21,39,42,7,41].

Even though the coupled SVAT-vegetation models capture the biophysics fairly well, their estimates of root-zone soil moisture (RZSM) diverge from reality due to errors in model conceptualization, computation, and numerical implementation, and due to uncertainties in model parameters, forcings, and initial conditions. The model estimates of moisture in the root zone can be significantly improved by assimilating remotely sensed [28,46,47,13], *in situ* [40] observations of soil moisture or assimilating remotely sensed surface temperature [18,49,12] into an SVAT model.

Early studies used for estimating time-varying states, assume that the only source of errors is uncertainties in model parameters [52,55]. Recent studies have incorporated errors in forcings, initial conditions, as well as those in model parameters [49,48,56,29,45]. Some studies that conduct sensitivity analyses assume a lumped error from all of the sources without incorporating errors from each, explicitly (e.g. [12]). All these sources of uncertainties can be explicitly incorporated using sequential data assimilation procedures, such as the Ensemble Kalman Filter (EnKF). The EnKF is a Monte-Carlo based Kalman Filter [19] that can be used for problems that are highly non-linear, such as those in hydrology. One of the main advantages of the EnKF compared to the Extended Kalman Filter is that it does not require linearization of the state equation during the propagation step. Reichle et al. [49] applied the EnKF to estimate soil moisture profile and found an error lower of 0.2% VSM to observations using the EnKF than the EKF. The EnKF obtained results closer to observations than the EKF with small number of ensemble members for applications with non-linear functions.

The EnKF is typically applied for dynamic state estimation, but model parameters can also be included in the framework through a

* Corresponding author.
E-mail address: monsivais@ufl.edu (A. Monsivais-Huertero).

0309-1708/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.advwatres.2010.01.011

state augmentation technique [3,42]. However, this increases the number of model states and parameters that are estimated simultaneously, resulting in increased computational demand and complexity. Moradkhani et al. [40] used two concatenated EnKF filters, in which the first estimates the parameters and the second filter estimates the states using the updated set of parameters. This implementation of the dual state–parameter estimation is computationally more efficient, however simultaneous state–parameter assimilation may be preferable due to tight coupling between states and relevant parameters in the system. Recent advances in high performance computing allow for simultaneous estimation of states and parameters. In this study, we implement the EnKF technique for simultaneous estimation of states and parameters in a coupled SVAT-vegetation model.

Even though tremendous progress has been made over the last few years to improve RZSM from SVAT models, some important issues still remain unresolved. First, most of the current studies update only the state estimates [45,29], not taking advantage of correlations between state and parameters during the update. As mentioned earlier, simultaneous state–parameter update is ideal.

Second, most of the studies involve either synthetic observations or empirically-derived near-surface soil moisture from remote sensing observations (e.g. [14,42]). Rigorous studies applying the EnKF to field observations are few, primarily due to lack of high temporal and spatial (vertical) density datasets such as in De Lannoy et al. [15]. Such studies can provide unique insights into the biophysics implemented in the model, particularly during changing land surface conditions.

Third, the temporal frequency of update that provides the most optimal estimates for RZSM is still debatable. Some synthetic studies such as Walker and Houser [56], found that daily near-surface soil moisture observations are required to achieve the best estimates of near-surface soil moisture. Longer times between observations offered only marginal impact in improving the RZSM estimates. For example, Hoeben and Troch [27] found that repeat observation times greater than 1-day (such as every 3, 6, or 12 h) did not improve the RMSE and it did not show any specific trend for more frequent update intervals due to saturation effects in the moisture profile. Most of the current studies using remotely sensed soil moisture observations assimilate every 2–3 days, primarily simulating expected temporal frequency of observations by near-future satellite-based microwave sensors [14,12].

Fourth, the spatial (vertical) frequency of the update in the root zone has received little attention, primarily due to lack of high density of measurements in the soil profile. For example, De Lannoy et al. [15] used field observations of SM at different depths to estimate the SM profile at temporal frequencies of 1 day, 1 week, and 8 weeks. They found that the assimilation of SM at different depths reduced the differences between the SM estimates and observations at deeper layers than the assimilation of near-surface SM. Recent research has focused on the combination of hydrological models and remotely sensed data to estimate soil moisture [14,41,45]. Most current studies use only near-surface soil moisture to improve RZSM estimates. Understanding the effects of soil moisture information throughout the soil profile on RZSM estimates is important for hydrologic applications such as partitioning infiltration into interflow and recharge. It accounts for the relationship between the observed soil moisture near land surface and soil moisture deeper in the soil profile.

Fifth, studies that lump the uncertainties due to forcings, parameters, and model physics (e.g. [14]) are not able to provide insights into the impact of each source of uncertainty on RZSM estimates. Very few studies include uncertainties in micrometeorological forcings (e.g. [56,15,45]). Walker and Houser [56] used a synthetic study comparing RZSM estimates with and without precipitation errors and found that precipitation errors significantly affect RZSM estimates. However, near-surface soil moisture estimates were not affected by these errors when near-surface soil moisture observations were assimilated.

The goal of this study is to address the five issues mentioned above using EnKF to merge observations with predictions from the coupled SVAT-vegetation model and estimate the RZSM during dynamic vegetation conditions. Here, we use observations from our second Microwave Water Energy Balance Experiment (MicroWEX-2) to compare with the results from EnKF using synthetic observations. Both the synthetic and the MicroWEX-2 observations were used to conduct four sets of comparisons: (i) estimation of state only with simultaneous estimation of state and parameters; (ii) assimilation of soil moisture at top 5 cm with assimilation of soil moisture throughout the root zone; (iii) assimilation conducted every 1, 3, and 10 days; and (iv) assimilation including uncertainties in precipitation/irrigation forcings with assimilation without forcing uncertainties.

In the next section, we briefly describe the MicroWEX-2 observations, the coupled SVAT-vegetation model, and the EnKF algorithm used in this study.

2. Experiment, model, and assimilation

2.1. MicroWEX-2

MicroWEX-2 was conducted from Day of Year (DoY) 78 (March 18) to DoY 153 (June 1) in 2004, to monitor micrometeorological, soil, and vegetation conditions as well as the microwave brightness temperatures during a growing season for sweet corn of variety Saturn SH2. Judge et al. [33] provides a detailed description of the experiment which is briefly summarized here. The experimental site was a 3.6 ha (9 acre) field located at the UF/IFAS Plant Science and Research Education Unit (PSREU), in North Central Florida (29.41N, 82.18W). The soils at the site are lake fine sand, with 89.4% sand, 7.1% clay, and a bulk density of 1.55 g/cm^3. Corn was planted at a row spacing of 76 cm, with a density of 8 plants/m^2. Irrigation and fertigation were conducted via a linear move system. Data used in this study include soil moisture, wind speed, upwelling and downwelling short- and long-wave radiation, precipitation, irrigation, relative humidity, and air temperature measured every 15 min. The soil moisture values were observed at six depths: 2, 4, 8, 32, 64, and 100 cm, using Campbell Scientific Water Content Reflectometers. An Eddy covariance system measured wind speed, and a net radiometer from Radiation and Energy Balance Systems (REBS) measured up- and down-welling short- and long-wave radiation. Four tipping-bucket rain gauges logged precipitation and irrigation at four locations in the field. Table 1 shows the different growth stages of corn and their associated vegetation characteristics observed during MicroWEX-2.

2.2. LSP–DSSAT model

The SVAT model used in this study is the Land Surface process (LSP) model [32,35]. It simulates 1-D coupled energy and moisture transport in soil and vegetation using a diffusion type equation, and estimates energy and moisture fluxes at the land surface and in the root zone. The model is forced with micrometeorological parameters such as air temperature, relative humidity, downwelling solar and long-wave radiation, irrigation/precipitation, and windspeed. The original version of the LSP model has been rigorously tested [32] and extended to wheat-stubble [34] and brome-grass [35] in the Great Plains, prairie wetlands in Florida [58], and to tundra in the Arctic [10]. In this study, we use a new version of the LSP model with a modified radiation flux parametrization at the land surface [7]. The LSP model includes 16 parameters, with

Table 1
The growing season of corn during MicroWEX-2 as described in [7].

Season	Number of days	Canopy height (cm)	LAI	Vegetation cover	Characteristics
Early	27 (DoY 78–105)	<17	<0.2	<0.22	Almost bare soil
Mid	20 (DoY 105–125)	17–73	0.2–1.82	0.22–1.00	Vegetation partially covering the terrain
Late	10 (DoY 125–135)	73–162	1.82–2.49	1.00	Maximum vegetative growth stage
Reproductive	19 (DoY 135–154)	162–200	2.49–2.75	1.00	Silking and ear formation, biomass increases primarily due to ear development

four parameters related to radiation balance, eight to latent and sensible heat fluxes and the remaining four being soil hydraulic properties, as shown in Table 2. The vegetation energy balance is calculated using equations from Verseghy et al. [54] for the water drainage from canopy, the bulk transfer approach from Trenberth [53] for the sensible heat flux and the latent heat flux following Peixoto and Oort [43]. The soil energy balance is calculated following the equations from Philip and de Vries [44] and de Vries [17].

The number of soil layers, with uniform constitutive properties, are user-defined. Two layers are used for this study: the first layer, top 1.7 m of soil is primarily sandy, with 89.4% sand and the second layer (1.7–2.7 m) constituted 40.5% sand. The soil has 35 computational blocks (nodes) in the two layers. The thickness of the blocks increases exponentially, with 4 blocks in the top 5 cm of the soil. A block-centered finite-differential scheme is employed to solve the coupled governing equations and estimate energy and moisture fluxes at land surface and in the root zone at an adaptive time step (seconds/minutes) in response to the forcings [7].

The LSP model was coupled to a vegetation growth model, viz. Decision Support System for Agrotechnology Transfer (DSSAT) model to provide the flux estimates during dynamic vegetation conditions [7]. The DSSAT is a crop model with modules for soil, soil–plant–atmosphere, weather, management, and crop development and growth [31]. The soil module simulates soil moisture using a bucket model [38] and soil temperature as an empirical function of air temperature and depth. The soil–plant–atmosphere module estimates evotranspiration (ET). In the weather module, meteorological forcings are read in, and in the management module, irrigation, fertilization, and pest control are read or generated. The crop module simulates the phenological development and growth, on a daily timestep, of a number of different crops, including soybeans, wheat, and cotton. The DSSAT model was tested and calibrated for its applicability to North-Central Florida [8] before it was coupled to the LSP model.

In the coupled LSP–DSSAT model, the LSP model provides the DSSAT model with estimates of soil moisture and temperature profiles and ET. The DSSAT model provides the LSP model with vegetation characteristics that influence heat, moisture, and radiation transfer at the land surface and in the vadose zone. The LSP–DSSAT model was rigorously tested and calibrated using observations [7]. The model estimates of surface fluxes, volumetric soil moisture (VSM) and soil temperature using the coupled LSP–DSSAT model were similar to those using the stand alone LSP model, indicating that the LSP–DSSAT model can be used to simulate fluxes for dynamic vegetation without the need of *in situ* vegetation observations.

Table 2
Parameters included in the LSP model [7]. The values for canopy parameters were from [23] and ranges for soil parameters were from [50].

Parameter	Description	Values
CANOPY		
z_{ob}	Bare soil roughness length (m)	0.004
x	Leaf angle distribution parameter	0.819
σ	Leaf reflectance	0.474
ϵ_c	Canopy emissivity	0.973
ϵ_s	Soil emissivity	0.953
c_d	Canopy drag coefficient	0.328
i_w	Canopy wind intensity factor	67.9
l_w	Leaf width (m)	0.0531
F_b	Base assimilation rate (kg CO_2/m^2 s)	-0.82×10^{-8}
ϵ_{photo}	Photosynthetic efficiency (kg CO_2/J)	0.897×10^{-6}
$soil_a$	Slope parameter for r_s (m^2 s/kg H_2O)	370
$soil_b$	Intercept parameter for r_s (m^2 s/kg H_2O)	−531
SOIL (0–1.7 m)		
λ	Pore-size index	0.1–0.9
ψ_0	Air entry pressure (m H_2O)	0.05–1.0
K_{sat}	Saturated hydraulic conductivity (m/s)	10^{-5}–10^{-3}
ϕ	Porosity (m^3/m^3)	0.2–0.55
SOIL (1.7–2.7 m)		
λ	Pore-size index	0.05
ψ_0	Air entry pressure (m H_2O)	0.019
K_{sat}	Saturated hydraulic conductivity (m/s)	8.93×10^{-5}
ϕ	Porosity (m^3/m^3)	0.41

2.3. Ensemble Kalman Filter

The assimilation algorithm used in this study is based up on the EnKF. The EnKF algorithm propagates an ensemble of state vectors in parallel such that each state vector represents one realization of generated model replicates. The state equation in the EnKF for each ensemble member is [22,40]:

$$x_t^{i-} = f\left(x_{t-1}^{i+}, u_{t-1}^{i}, \theta_{t-1}^{+}, t-1\right) + \omega_{t-1}^{i} \tag{1}$$

where $f(\cdot)$ is the non-linear model, x_t^{i-} is the state of the ith after update ensemble member prior to the update at time t, x_{t-1}^{i+} is the posterior state of the ith ensemble member at time $t-1$, u_{t-1}^{i} represents the meteorological forcings, θ_{t-1}^{+} is the parameters of the non-linear model, and ω_{t-1}^{i} is the model error. In our study, model physics is assume to be perfect $(\omega_{t-1}^{i} = 0)$.

The A matrix, holding the ensemble members, can be expressed by the form:

$$A = \{x^1, x^2, \ldots, x^N\} \tag{2}$$

where x_i represents each member of the ensemble, and N is the number of ensemble members.

If we collect all observations taken at time t into the observation vector d of dimension m, we can express the observation process as:

$$d^i = hx_t^{i-} + \epsilon_t^j \tag{3}$$

where $h(\cdot)$ is the operator relating the state variables to the observations, and ϵ is the error associated to the observations with zero mean.

The ensemble of pertubed observations can be written as:

$$D = \{d^1, d^2, \ldots, d^N\} \tag{4}$$

and the ensemble of perturbations as:

$$\gamma = \{\epsilon^1, \epsilon^2, \ldots, \epsilon^N\} \tag{5}$$

Mathematically, the EnKF can be represented by the standard equation

$$A^+ = A^- + K(D - HA^-) \tag{6}$$

where A^- denotes the ensemble of prior states and A^+, the posterior ensemble member, K is the Kalman gain, D is the ensemble of perturbed observation, and H the operator relating the ensemble of perturbated observations to the ensemble of states.

The Kalman gain can be calculated as [20]:

$$K = P_e H^T (H P_e H^T + R_e)^{-1} \tag{7}$$

where P_e represents the prior ensemble covariance matrix, and R_e is the observation error covariance matrix.

Then, the update equation can be written as:

$$A^+ = A^- + P_e H^T (H P_e H^T + R_e)^{-1} (D - HA^-) \tag{8}$$

Defining the ensemble perturbation matrix as $A' = A^- - \overline{A}$, with $\overline{A}$ being the mean matrix of A^-, and $P_e = A'(A')^T/(N-1)$ where N is the number of ensemble members, Eq. (8) can be written as:

$$A^+ = A^- + A'A'^T H^T (HA'A'^T H^T + \gamma\gamma^T)^{-1} (D - HA^-) \tag{9}$$

where γ is the ensemble of perturbations and is related to R_e by $R_e = \gamma\gamma^T/(N-1)$.

The common way of solving Eq. (9) involves the computation of the eigenvalues to decompose the term $HA'A'^T H^T + \gamma\gamma^T$. Using the Singular Value Decomposition [2], this term can be written as:

$$HA'A'^T H^T + \gamma\gamma^T = U \Lambda U^T \tag{10}$$

and its inverse matrix as:

$$(HA'A'^T H^T + \gamma\gamma^T)^{-1} = U \Lambda^{-1} U^T \tag{11}$$

Finally, Eq. (9) can be computed from:

$$A^+ = A^- + A'(HA')^T U \Lambda^{-1} U^T (D - HA^-) \tag{12}$$

3. Methodology

In this section, we describe the LSP–DSSAT simulations and implementation of the EnKF. We present the convergence and the sensitivity studies conducted, and the four sets of comparisons mentioned in Section 1.

3.1. LSP–DSSAT simulation

The coupled LSP–DSSAT model simulated the energy and moisture fluxes from planting on DoY 78, to harvest on DoY 153 during a corn-growing season in 2004. Micrometeorological forcings for the simulations were obtained from the observations during MicroWEX-2. Initial conditions were not known during MicroWEX-2 because sensor installation was completed 7 days after planting. The first values observed by the soil moisture and temperature sensors were used as the initial moisture and temperature values for the DoY 78.

3.2. Implementation and applicability of EnKF

In this study, the non-linear propagator $f(\cdot)$ in Eq. (1) represents the coupled LSP–DSSAT model; x is the state vector consisting of VSMs estimated by the LSP–DSSAT model; u_t is the meteorological forcings at time t, and θ is the time-invariant model parameters in the LSP model, shown in Table 2. The augmented state vector technique [22,30] is applied here to simultaneously assimilate states and parameters. In this study, we assume that the MicroWEX-2 observations are unbiased, following Lorenc and Hammom [37]. The augmented state vector technique is frequently used to account for bias in the model [16]. This technique has also been tested and applied by Baek et al. [4] and Reichle et al. [48].

The VSM observations during MicroWEX-2 were obtained from the top 1.7 m of layer of the soil, the extended state vector x_i (see Eq. (13)) includes the imperfectly unknown parameters describing the first layer. In this study, the second layer (1.7–2.7 m) of soil has constant constitutive properties (see Table 2) obtained from [7], which were assumed perfectly known and therefore not updated. The ith member of the ensemble is expressed as

$$x_i = \begin{bmatrix} VSM_{i1} \\ VSM_{i2} \\ \vdots \\ VSM_{ik} \\ \phi_i \\ \psi_{0_i} \\ K_{sat_i} \\ \lambda_i \end{bmatrix} \tag{13}$$

where k represents the number of nodes of the LSP model and ϕ_i, ψ_{0_i}, K_{sat_i}, and λ_i are soil parameters defined in Table 2.

Previous investigations have considered VSM observation errors between 3% and 5% [56,49]. In Walker and Houser [56], authors conclude that the error should be <5%, but preferably <3%. In this paper, the standard deviation of the error in VSM observations was 2% by volume [6] for both synthetic and MicroWEX-2 observations. In addition, error is considered to be Gaussian with zero mean.

The assimilation time was chosen to be 6 a.m. EST for this study, corresponding to the current and near-future availability of remotely sensed soil moisture [36,59,5].

3.3. Convergence of soil moisture estimates and sensitivity of RZSM to model parameters

To determine the optimal number of ensemble members, we analyzed the means and the standard deviations for an ensemble of 8000 members. The criterion for oscillations in mean and standard deviation of VSM at 0–5 cm, the depth exhibiting highest variability, in selecting the number of ensemble members for our study was $\leqslant 0.001$ m^3/m^3. Two representative days were chosen for each of the four growth stages of corn (Table 1) to understand the behaviour of mean and standard deviation of VSM for different numbers of ensemble members. One day was during a precipitation/irrigation event and the other was during the middle of a dry down period.

Randomly-generated, uniformly distributed parameters, with literature-based upper and lower bounds, were used in the EnKF. Table 2 shows the range considered for each parameter. The use of a uniform distribution avoids the generation of negative parameters. The time dependent correlations between the 16 parameters and the RZSM estimates were examined.

3.4. Synthetic and field observations

The synthetic truth was obtained from one of the realizations from an open-loop simulation of the LSP–DSSAT model with a randomly generated parameter set. The truth was perturbed with a Gaussian error with zero-mean and 0.02 m^3/m^3 standard deviation to generate synthetic observations. The truth was not included in the ensemble of 500 members during the assimilation by the EnKF. The field observations were obtained from the MicroWEX-2 experiment, described in Section 2.1. The Gaussian error with zero mean and 0.02 m^3/m^3 standard deviation was also added to the MicroWEX-2 observations.

3.5. Temporal and spatial frequency for assimilation

The VSM observations were assimilated at intervals of 1-, 3-, and 10-days. The observations obtained from five different depths, 0–5, 8, 32, 64, and 100 cm were assimilated in this study. The 0–5 cm VSM represented near-surface soil moisture, comparable to that derived from remote sensing measurements, and VSM at 0–100 cm represented the RZSM. The 0–5 cm VSM and RZSM estimates and observations were calculated by the equation:

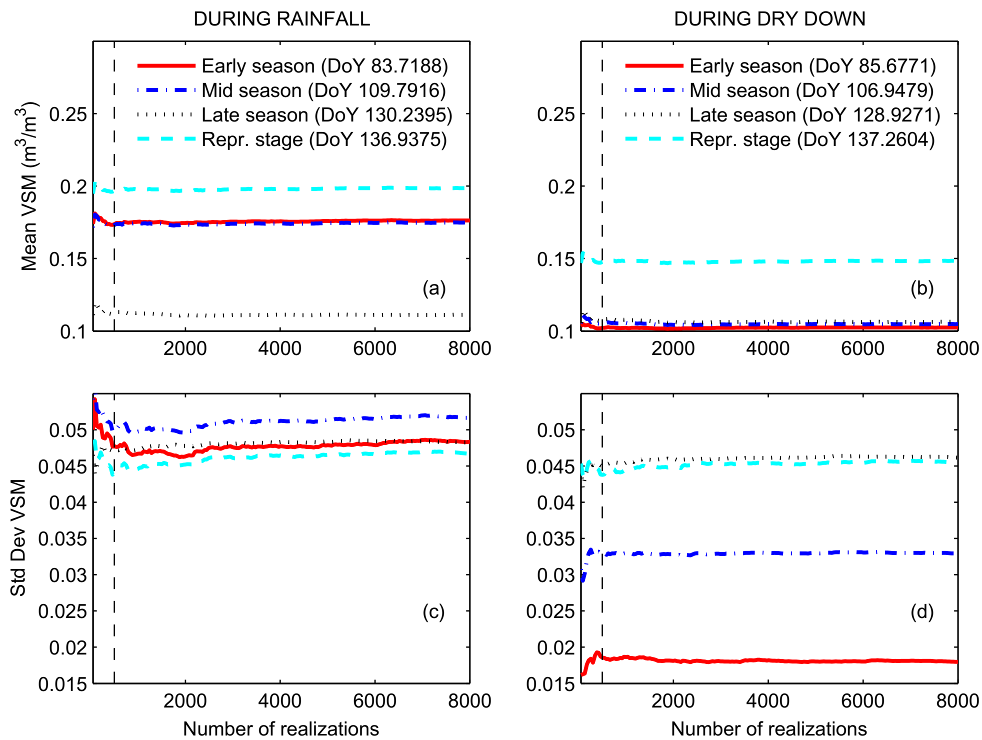

Fig. 1. Comparison of mean (a, b) and standard deviation (c, d) of VSM estimates at 0–5 cm by the LSP–DSSAT model for different number of realizations in the ensembles during different stages of growth on days with rainfall (a, b) and days during the dry down.

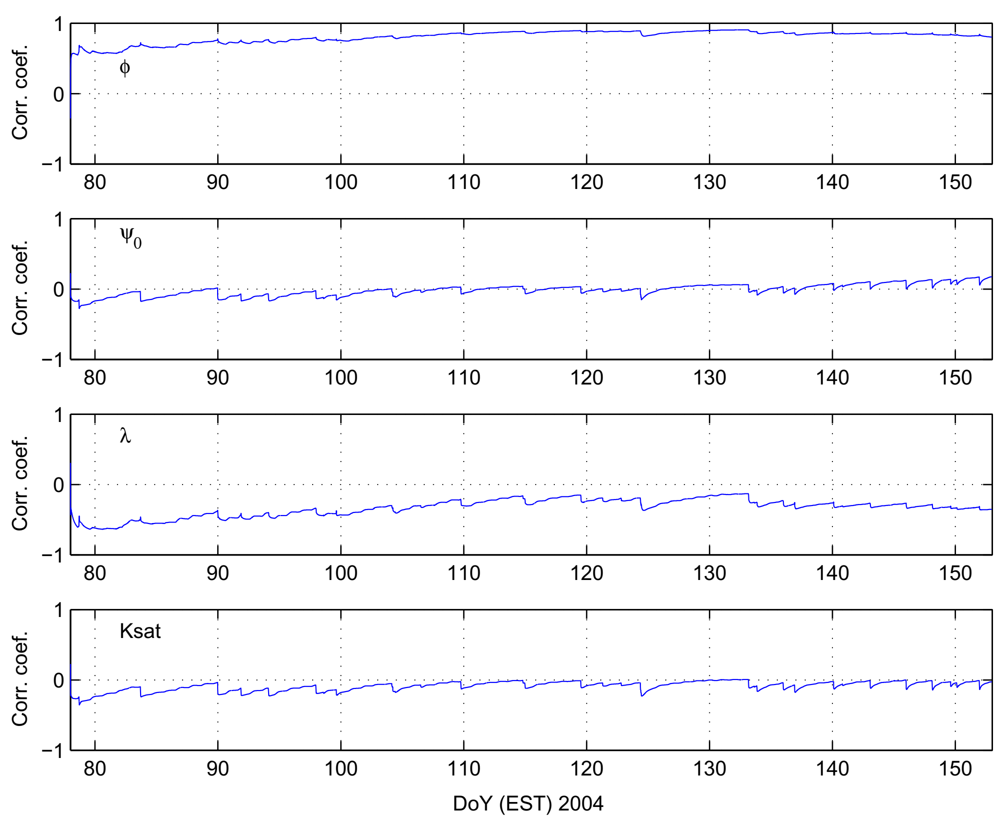

Fig. 2. Lag-zero (in time) correlation coefficient between the four sensitive parameters of soil, viz., ϕ, ψ_0, λ, and K_{sat}, and RZSM over the growing season.

$$VSM \text{ or } RZSM = \sum_{i=1}^{k} VSM_i \Delta z_i \qquad (14)$$

where k indicates the total number of nodes (blocks) within 0–5 cm or the root zone, Δz_i the thickness of the ith node, and VSM_i the volumetric soil moisture at ith node.

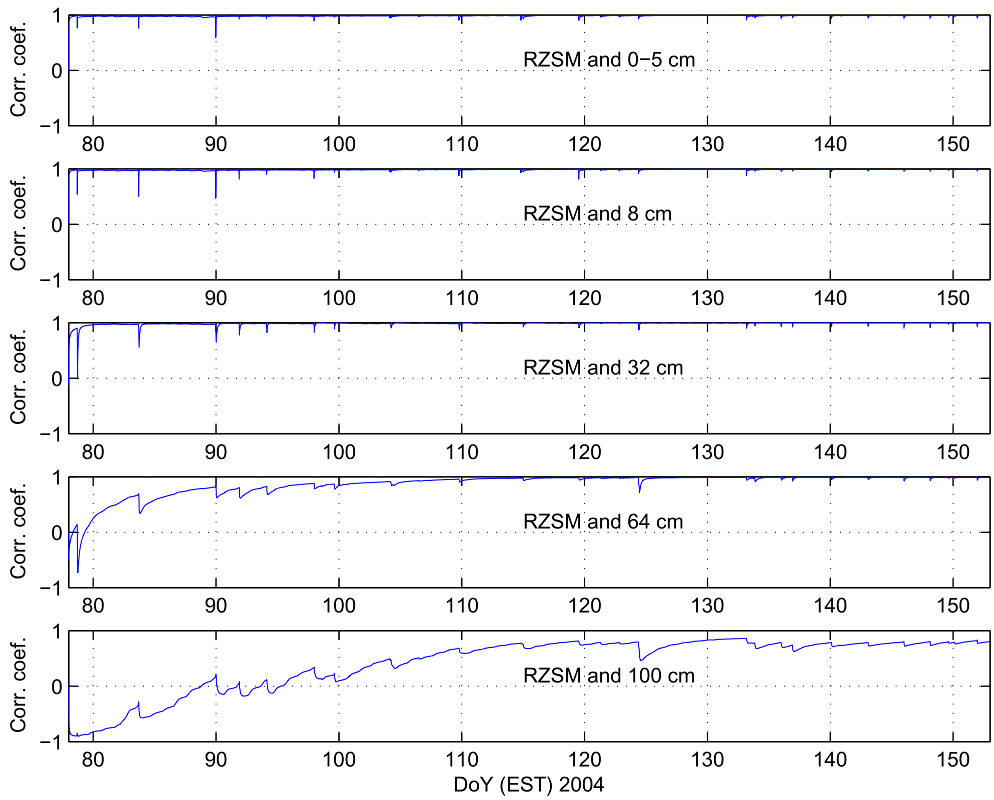

Fig. 3. Lag-zero (in time) correlation coefficient between RZSM and VSM at depths of 0–5, 8, 32, 64, and 100 cm during the growing season.

Table 3
Root mean square errors and differences (RMSE/D) and average standard deviations (ASD) of the RZSM (m^3/m^3) assimilating synthetic and field observations. S–P indicates simultaneous state–parameter estimation; S indicates state-only assimilation; 5-depths indicates scenarios when VSM observations at 0–5, 8, 32, 64, and 100 cm are assimilated.

Observations	Scenario	1-Day		3-Day		10-Day		Open loop	
		RMSE/D	ASD	RMSE/D	ASD	RMSE/D	ASD	RMSE/D	ASD
Synthetic	S–P, 0–5 cm	0.0013	0.0093	0.0024	0.0090	0.0019	0.0104	0.0233	0.0226
	S–P, 5-depths	0.0013	0.0055	0.0015	0.0049	0.0014	0.0081	0.0233	0.0226
	S, 0–5 cm	0.0050	0.0103	0.0087	0.0112	0.0106	0.0141	0.0233	0.0226
	S, 5-depths	0.0059	0.0096	0.0067	0.0100	0.0091	0.0123	0.0233	0.0226
MicroWEX-2	S–P, 0–5 cm	0.0114	0.0064	0.0189	0.0077	0.0274	0.0100	0.0250	0.0226
	S–P, 5-depths	0.0063	0.0044	0.0087	0.0056	0.0133	0.0077	0.0250	0.0226

Table 4
Root mean square errors and differences (RMSE/D) and average standard deviations (ASD) of VSM (m^3/m^3) at different depths in the soil profile for simultaneous state–parameters estimation and assimilating VSM observations at 0–5, 8, 32, 64, and 100 cm depths using synthetic and field observations.

Observations	Depth (cm)	1-Day		3-Day		10-Day		Open loop	
		RMSE/D	ASD	RMSE/D	ASD	RMSE/D	ASD	RMSE/D	ASD
Synthetic	0–5	0.0012	0.0065	0.0024	0.0058	0.0033	0.0098	0.0283	0.0270
	8	0.0009	0.0063	0.0019	0.0055	0.0028	0.0089	0.0281	0.0268
	32	0.0009	0.0056	0.0014	0.0050	0.0020	0.0082	0.0254	0.0242
	64	0.0012	0.0055	0.0015	0.0049	0.0011	0.0080	0.0200	0.0206
	100	0.0026	0.0083	0.0033	0.0067	0.0018	0.0105	0.0110	0.0198
MicroWEX-2	0–5	0.0104	0.0073	0.0123	0.0094	0.0137	0.0119	0.0205	0.0270
	8	0.0115	0.0083	0.0172	0.0083	0.0178	0.0103	0.0233	0.0268
	32	0.0071	0.0053	0.0085	0.0060	0.0083	0.0084	0.0167	0.0242
	64	0.0201	0.0040	0.0238	0.0050	0.0311	0.0073	0.0452	0.0206
	100	0.0090	0.0039	0.0144	0.0055	0.0204	0.0082	0.0405	0.0198

3.6. Uncertainty in forcings

Among all the inputs/forcings to the LSP–DSSAT model, precipitation/irrigation observations typically have the highest errors compared to other micrometeorological parameters. These errors can range between 2.9% and 12.3%, depending on the duration and the intensity of rainfall [24,11] and may also have a high impact on VSM estimation. To understand the impact of errors in

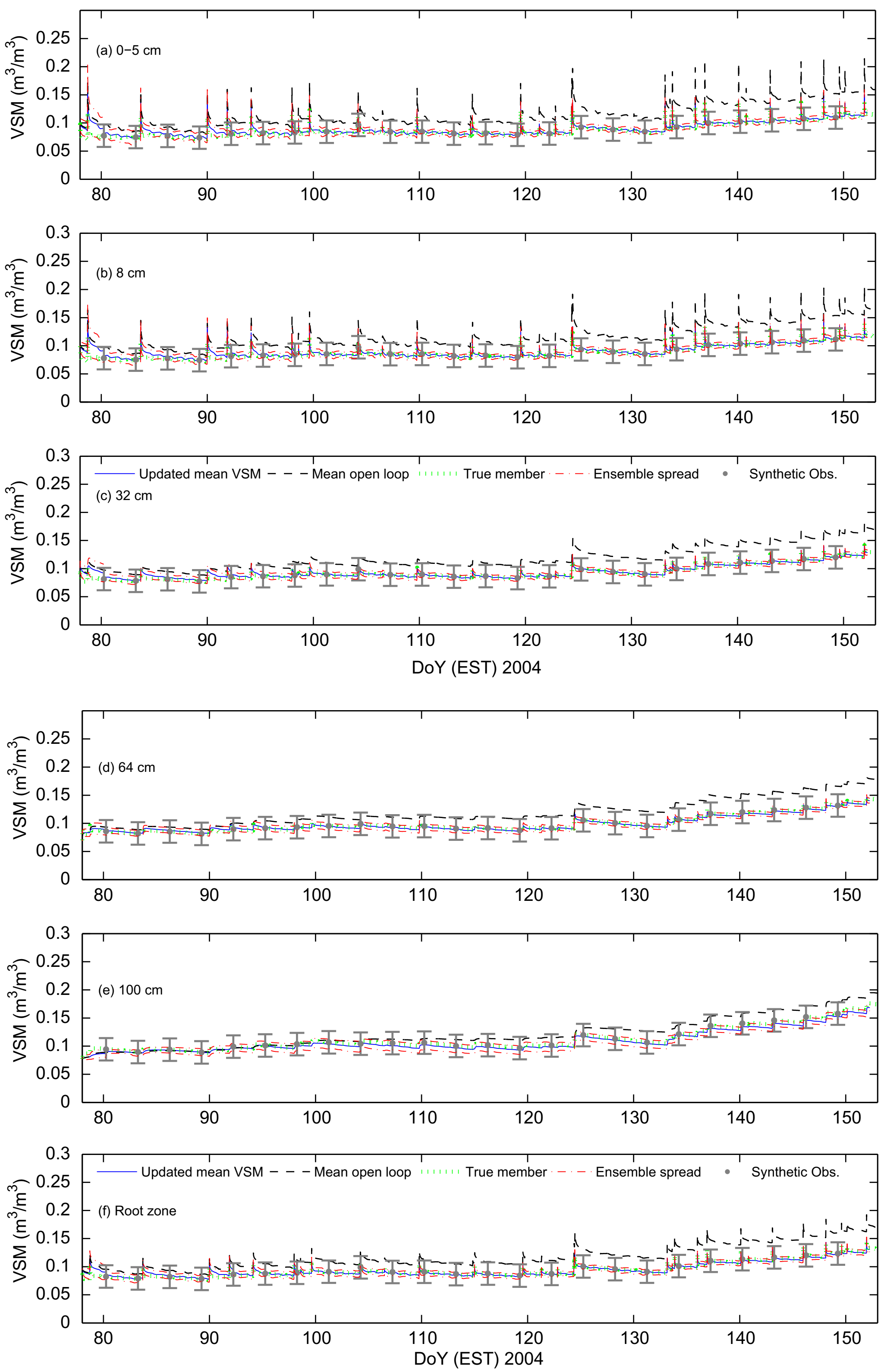

Fig. 4. Means and standard deviations at depths of (a) 0–5 cm, (b) 8 cm, (c) 32 cm, (d) 64 cm, (e) 100 cm, and (f) in the root zone for simultaneous state–parameter estimation when synthetic observations are assimilated every 3-days. These values are compared with true values and open-loop estimates.

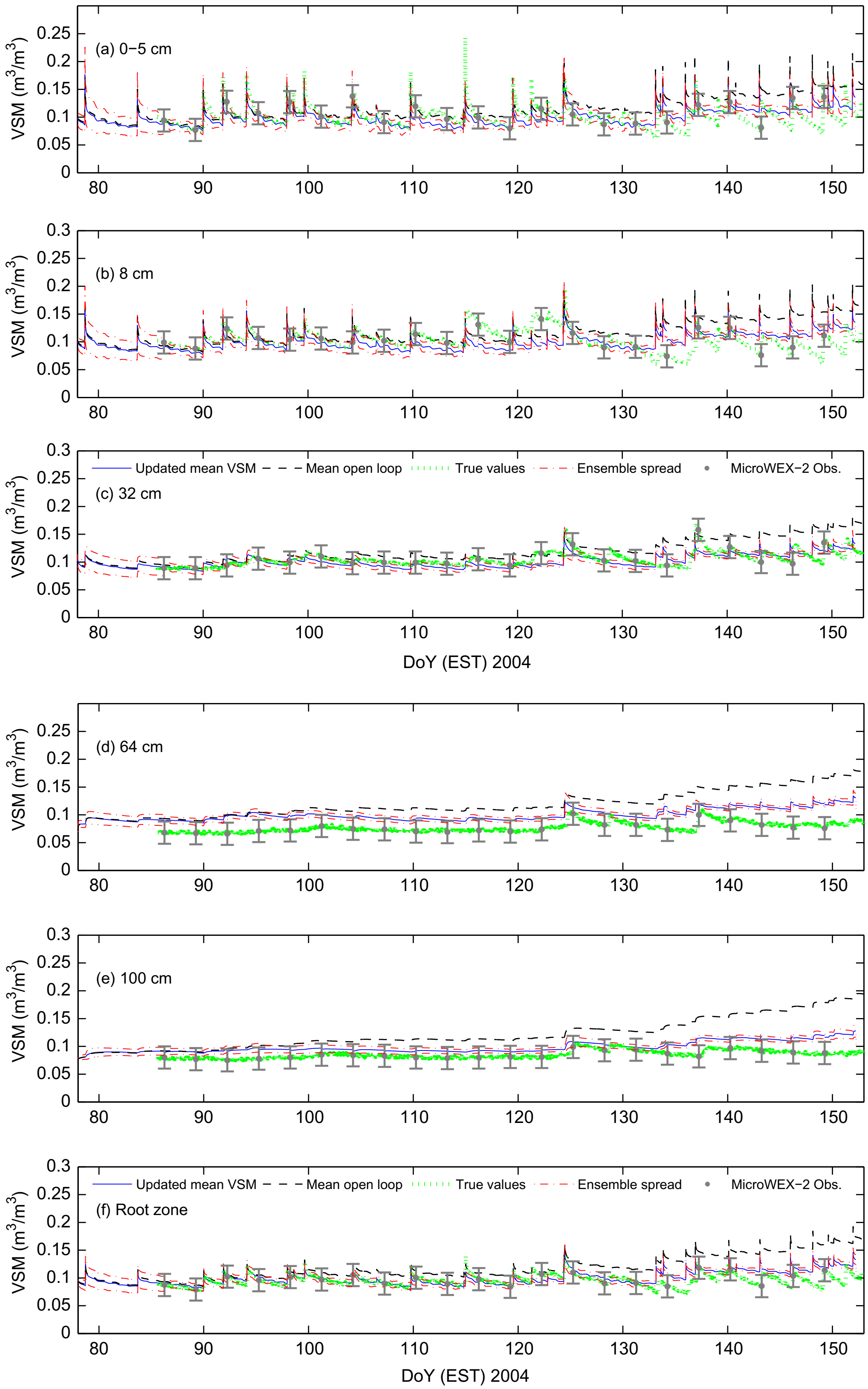

Fig. 5. Means and standard deviations at depths of (a) 0–5 cm, (b) 8 cm, (c) 32 cm, (d) 64 cm, (e) 100 cm, and (f) in the root zone for simultaneous state–parameter estimation when MicroWEX-2 observations are assimilated every 3-days. These values are compared with true values and open-loop estimates.

forcings in the estimation of RZSM, we introduced a Gaussian observation error with standard deviation equal to 12% of the observed value of precipitation/irrigation during events. No errors were introduced when there was no event.

4. Results and discussion

4.1. Convergence of soil moisture and sensitivity of RZSM to parameters

Fig. 1 shows the mean and standard deviations VSM at 0–5 cm obtained using different number of members in the ensemble, during the different growth stages for periods of precipitation and drydown. The mid-season shows the highest standard deviation during the precipitation/irrigation period, while the late and reproductive seasons show similar high standard deviations during the drydown (Fig. 1c and d). When the ensemble members increase beyond 500, minimal changes are observed in the mean and standard deviation. The maximum mean VSM oscillation is 0.0005 m^3/m^3 and the maximum standard deviation oscillation is 0.001 m^3/m^3 for 500 members. In this study all simulations were conducted using 500 members in the ensemble.

In this study, we vary the four soil parameters, viz., ϕ, ψ_0, λ, and K_{sat}, while the other 12 parameters, shown in Table 2, were set constant at their previously calibrated value from Casanova and Judge [7] because the correlation coefficients between RZSM and the other 12 parameters in the LSP model (see Table 2) were found to be very close zero (for e.g. 3×10^{-2}).

Fig. 2 shows the lag-zero correlation coefficients between the RZSM and the four soil parameters over the growing season. From the figure, ψ_0 and K_{sat} have similar correlations with RZSM, varying

Table 5
Means and standard deviations (Std. dev.) of the four soil parameters in the LSP–DSSAT model for simultaneous state–parameter estimation using synthetic and field observations. 5-Depths indicates scenarios when VSM observations at 0–5, 8, 32, 64, and 100 cm are assimilated.

Observations	Parameters	True value	Scenario	1-Day		3-Day		10-Day		Open loop	
				Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Synthetic	ϕ	0.276	0–5 cm	0.2897	0.0380	0.2826	0.0372	0.2970	0.0445	0.3698	0.0967
			5-depths	0.2785	0.0127	0.2788	0.0103	0.2828	0.0205	0.3698	0.0967
	ψ_0	0.6554	0–5 cm	0.5204	0.1859	0.4899	0.1956	0.4627	0.2056	0.5005	0.2744
			5-depths	0.5561	0.1706	0.5149	0.1694	0.4685	0.1857	0.5005	0.2744
	λ	0.7572	0–5 cm	0.6545	0.1514	0.5893	0.1675	0.5949	0.1647	0.4935	0.2372
			5-depths	0.6559	0.1143	0.6317	0.1028	0.5996	0.1178	0.4935	0.2372
	K_{sat}	5.85×10^{-4}	0–5 cm	6.06×10^{-4}	2.37×10^{-4}	5.90×10^{-4}	2.46×10^{-4}	5.47×10^{-4}	2.69×10^{-4}	5.13×10^{-4}	2.80×10^{-4}
			5-depths	6.15×10^{-4}	2.22×10^{-4}	6.02×10^{-4}	2.50×10^{-4}	5.55×10^{-4}	2.57×10^{-4}	5.13×10^{-4}	2.80×10^{-4}
MicroWEX-2	ϕ	–	0–5 cm	0.2078	0.0269	0.2984	0.0372	0.3715	0.0553	0.3698	0.0967
			5-depths	0.1811	0.0090	0.2241	0.0103	0.2147	0.0181	0.3698	0.0967
	ψ_0	–	0–5 cm	0.2235	0.1415	0.3032	0.1646	0.2128	0.1566	0.5005	0.2744
			5-depths	0.2616	0.1507	0.7210	0.2019	0.6354	0.2054	0.5005	0.2744
	λ	–	0–5 cm	0.2877	0.0838	0.4614	0.1490	0.6018	0.1579	0.4935	0.2372
			5-depths	0.0645	0.0090	0.0481	0.0574	0.0647	0.0655	0.4935	0.2372
	K_{sat}	–	0–5 cm	3.09×10^{-4}	1.85×10^{-4}	3.85×10^{-4}	2.18×10^{-4}	4.15×10^{-4}	2.35×10^{-4}	5.13×10^{-4}	2.80×10^{-4}
			5-depths	3.21×10^{-4}	1.47×10^{-4}	4.77×10^{-4}	1.95×10^{-4}	3.15×10^{-4}	2.03×10^{-4}	5.13×10^{-4}	2.80×10^{-4}

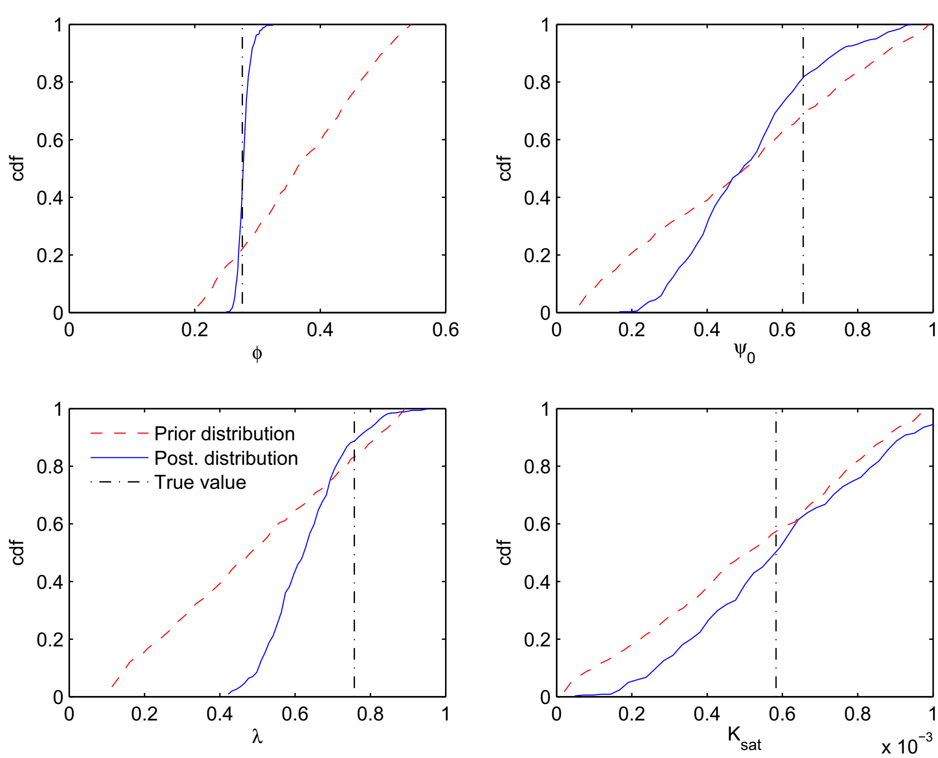

Fig. 6. Cumulative distribution function (CDF) of the estimates for porosity (ϕ), air entry pressure (ψ_0), pore-size index (λ), and saturated hydraulic conductivity (K_{sat}) at the beginning (first prior distribution) and at the end (last posterior distribution) of the growing season of sweet corn when synthetic observations were assimilated every 3-days.

from −0.25 to 0.1. The correlation coefficients for ψ_0 and K_{sat} increase (in negative magnitude) suddenly during precipitation/irrigation events. Fig. 3 shows the lag-zero (in-time) correlation between RZSM and VSM at depths of 0–5, 8, 32, 64, and 100 cm during the growing season. The RZSM is highly correlated with soil moisture in the top 32 cm throughout the growing season, with sudden decreases in correlation coinciding with precipitation/irrigation events. The correlation decreases with depth because the soil moisture values are higher at the near-surface and are lower in the root zone. The RZSM shows highly negative correlation values with depths below 42 cm (64 and 100 cm in Fig. 3) during early season. As the vegetation grows, the correlations decrease and then, become positive.

4.2. Comparison of simultaneous state–parameter estimation with state-only estimation

Table 3 shows the average standard deviations (ASD) and the root mean square errors (RMSE) of the RZSM for the synthetic observations and, the ASD and the root mean square difference (RMSD) for the MicroWEX-2 observations during the entire growing season for different assimilation scenarios. For both synthetic and MicroWEX-2 observations, the ASD for the assimilation cases were significantly lower, by approximately 0.0086–0.0177 m^3/m^3 (38–78%) than that for the open-loop simulation. It is also noted that the magnitude of the open-loop RMSE, of 0.0233 m^3/m^3, is similar to the observation error considered in this study. The RMSEs obtained for 10-day updates are similar to Pauwels et al. [42] and Muñoz-Sabatier et al. [41].

For all time-intervals of assimilation, the estimation of state–parameters simultaneously results in lower ASD, by 0.0086–0.0177 m^3/m^3 (38–78%) compared to open-loop ASD, than those obtained for cases when only states are estimated (0.0130–0.0086 m^3/m^3; 38–58%). For the simultaneous state–parameter case, assimilation of VSM observations at five different depths (0–5, 8, 32, 64, and 100 cm) in the root zone produces lower ASD, by 0.003 m^3/m^3 (25%) compared to the cases that assimilate only near-surface soil moisture VSM (0–5 cm) observations. For the state-only estimation in the synthetic case, the ASD reduced by 0.0012 m^3/m^3 (10%) and the RMSE by 0.001 m^3/m^3 (18%), when VSM at the five depths are assimilated, compared to the case when only 0–5 cm VSM is assimilated. This indicates that the assimilation of additional VSM observations in the profile improves the VSM estimates, similar to [15].

The RMSEs obtained for all temporal update frequencies are significantly lower, by 0.005 m^3/m^3 (51%) for the simultaneous estimation case than those for the state-only estimation case. When the forcings were assumed to be known perfectly, the simultaneous state–parameter estimates with daily or 3-day assimilation of VSM observations at the five depths resulted in the lowest uncertainty in RZSM estimation.

4.3. Comparison between synthetic and MicroWEX-2 observations

4.3.1. Estimates of states (VSM)

Table 4 shows the RMSE/RMSD, and ASD for VSM estimation at different depths using synthetic and MicroWEX-2 observations. Estimates of RZSM using synthetic observations show similar

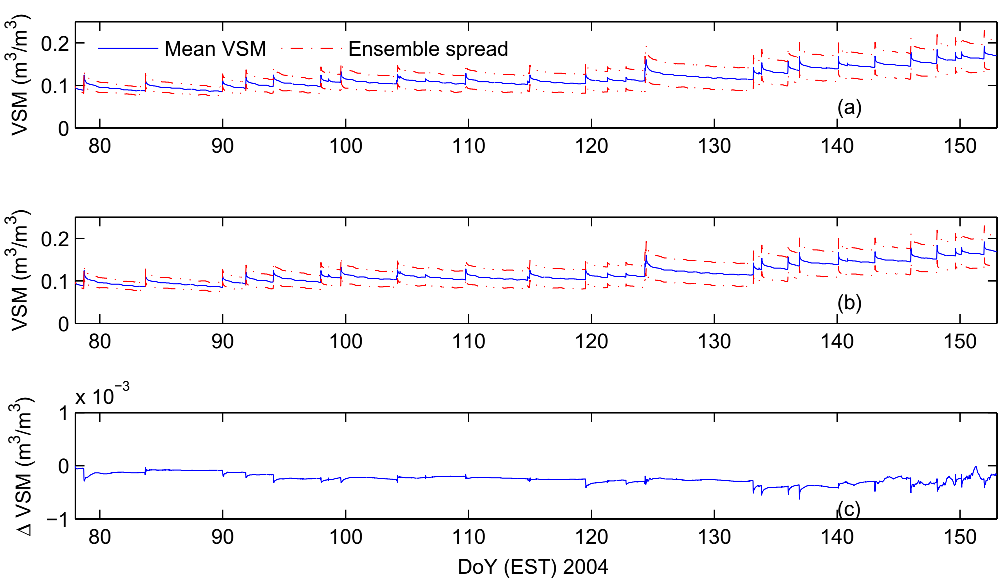

Fig. 7. Mean and standard deviation of RZSM estimates from open-loop simulation (a) without precipitation/irrigation uncertainties, (b) with precipitation/irrigation uncertainties, and (c) difference between open-loop simulations with and without precipitation/irrigation uncertainties.

Table 6
RMSE (for synthetic observations)/RMSD (for MicroWEX-2 observations) and average of the standard deviation (ASD) for the RZSM estimates for simultaneous state–parameter estimates using synthetic and MicroWEX-2 observations that were assimilated daily and every 3 days for the two scenarios: (A) without and (B) with uncertainties in precipitation/irrigation.

Observations	Scenario	1-Day		3-Day		Open loop	
		RMSE/RMSD	ASD	RMSE/RMSD	ASD	RMSE/RMSD	ASD
Synthetic	(A) Without uncertainty	0.0013	0.0055	0.0015	0.0049	0.0233	0.0226
	(B) With uncertainty	0.0016	0.0055	0.0015	0.0068	0.0230	0.0228
MicroWEX-2	(A) Without uncertainty	0.0063	0.0044	0.0087	0.0056	0.0250	0.0226
	(B) With uncertainty	0.0062	0.0046	0.0083	0.0057	0.0247	0.0228

ASD as those obtained using MicroWEX-2 observations. However, because the "perfect" model physics in the LSP–DSSAT do not follow the MicroWEX-2 observations, unlike the synthetic case, the RMSDs for MicroWEX-2 are higher than the RMSEs for synthetic observations (see Figs. 4 and 5). VSM estimates at 0–5, 8 cm and RZSM show the biggest differences after DoY 130 during the dry-down periods for MicroWEX-2 observations (Fig. 5). Although the posterior VSM estimates are close to the observations, the estimates do not follow the same trend as observations during the propagation phase. For the synthetic case, the RMSE values for RZSM shown in Tables 3 and 4 are of the order of 2% VSM or 0.02 m^3/m^3 and are comparable to those obtained by Reichle et al. [49] and Heathman et al. [26] for their assimilation of synthetic near-surface soil moisture using EnKF and direct-insertion filtering techniques, respectively. The mean VSM is close to the true value throughout the soil profile and the prior standard deviation of the estimates is lower than the error in observations for all time intervals of assimilation (see Table 4). Similar to Reichle et al. [49], the RMSE and ASD of the VSM at deeper layers are lower, at 0.001 (40%) and 0.001 m^3/m^3 (15%), respectively, compared to those of near-surface VSM estimates (see Table 4) for synthetic observations. This is because near-surface layers are more sensitive to the precipitation/irrigation events than the deeper layers.

The RMSD of 0–5 cm estimates are higher, at 0.012 m^3/m^3 (12% of the seasonal mean VSM), when MicroWEX-2 observations are assimilated at the five depths throughout the soil profile, compared to RMSE, at 0.0009 m^3/m^3 (7% of the seasonal mean VSM), using synthetic observations. The mean VSM at deeper layers of 64 and 100 cm are biased high compared to the MicroWEX-2 observations, with an RMSD of 0.0238 and 0.0144 m^3/m^3, respectively, for the entire growing season (see Fig. 5 and Table 4). However, the VSM at deeper layers matched well with synthetic observations (see Fig. 4 and Table 4), with RMSEs of 0.0015 and 0.0033 m^3/m^3, respectively. These differences between the estimates of MicroWEX-2 and synthetic scenarios are likely due to the assumption of homogeneous soil in the LSP–DSSAT model and also to a very low model estimate uncertainty of about 0.0050 m^3/m^3 compared to observation uncertainties of 0.02 m^3/m^3) (see Section 3.2). However, the RZSM estimates are not impacted by the assumption of homogeneous soil and have similar low values of RMSDs and RMSEs for both MicroWEX-2 and synthetic scenarios. For the MicroWEX-2 scenario, RZSM estimates do not follow the observations closely during dry-down periods after DoY 130 (see Fig. 5), with the average differences between the model estimates and observations at about 0.0161 m^3/m^3. A maximum difference of 0.053 m^3/m^3 is observed after DoY 136. The mean VSM at depths of 0–5 and 8 cm exhibit significant differences between model estimates and observations, of 0.013 and 0.0201 m^3/m^3, respectively, after DoY 130. In addition, a difference of 0.0341 m^3/m^3 is observed at 8 cm for DoY 115–125. Such differences are not observed in the synthetic scenarios, and are likely due to imperfect implementation of soil and plant characterization in the model compared to actual field observations. Typically, land is ploughed and disced up to 30 cm at the beginning of the growing season. This may result in heterogeneity of soil constitutive properties. Assumptions of homogeneous soil in the LSP–DSSAT model does not allow parameters to vary in the soil profile, as necessary to match reality. Also, the DSSAT model may need to

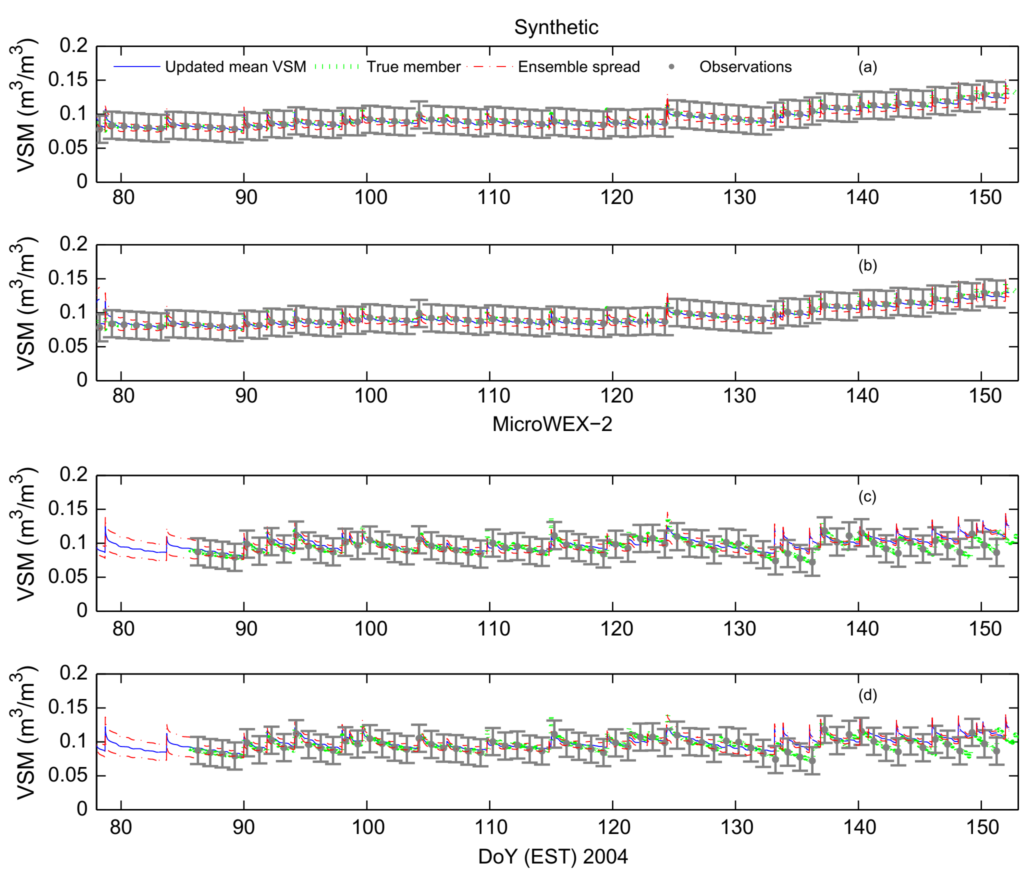

Fig. 8. Mean and standard deviation of RZSM estimates for simultaneous state–parameter estimation using VSM observations at the five depths of 0–5, 8, 32, 64, and 100 cm when assimilation was conducted every day for: (a) synthetic observations without precipitation/irrigation uncertainties, (b) synthetic observations with precipitation/irrigation uncertainties, (c) MicroWEX-2 observations without precipitation/irrigation uncertainties, and (d) MicroWEX-2 observations with precipitation/irrigation uncertainties.

be updated for canopy and root distribution parameters along with the soil parameters in the LSP model. Such insights into errors associated with model physics during the season cannot be obtained from synthetic studies.

4.3.2. Estimates of parameters

Table 5 gives the means and standard deviations of ϕ, ψ_0, λ, and K_{sat} at the end of the growth season, when synthetic and MicroWEX-2 observations are assimilated every 1-, 3-, and 10 days. In the synthetic case, estimates for all the parameters are close to their true values, for all assimilation scenarios, with a difference between estimates and true values at <20%. Fig. 6 shows the CDF of the estimates for the four parameters at the end of the season for synthetic observations. Parameter ϕ estimates has a Gaussian distribution, whereas K_{sat} keeps a shape similar to a uniform distribution.

In the MicroWEX-2 case, the parameter estimates are significantly different for different assimilation scenarios; with biggest differences for ψ_0 and K_{sat} parameters. The mean values are given in Table 5. ψ_0 has the higher standard deviation, and ϕ the lowest standard deviation. Even though, these mean values seem realistic for sandy soils, further analysis of these parameters could not be done due to lack of knowledge of their true values.

4.4. Temporal frequency of VSM observations

We discuss the uncertainties in state and parameter estimates for updates every 1-, 3-, and 10-days, using the synthetic case only, because these uncertainties are similar in both synthetic and MicroWEX-2 cases.

4.4.1. Estimate of states (VSM)

Table 4 compares the ASDs and RMSEs of VSM in the soil profile for dual state–parameter estimates when observations at the top 5 cm and at depths throughout the profile are assimilated every 1-, 3-, and 10-days. The ASDs are similar for both 1- and 3-day updates, both being lower than that every 10-day update. The highest ASDs are observed at the deepest layer (100 cm) and the near-surface layer (0–5 cm). The near-surface VSM estimate shows a high sensitivity to precipitation/irrigation events (see Fig. 4a).

From Fig. 4, assimilation results in a mean that is closer to the true value in comparison to the mean VSMs for the open-loop case. In all cases, the daily assimilation produces similar ASDs in the VSM as the 3-day assimilation, both being lower than those obtained from the 10-day assimilation. This result is in agreement with Walker and Houser [56] where the optimal temporal observation frequency was from 1 to 5 days. Most current studies use an assimilation interval of every 3-days, matching the interval of satellite revisit (for example, [14,12]). The lowest RMSE of 0.0013 m^3/m^3 (94% reduction of the open-loop RMSE) is obtained for the case when state–parameters are simultaneously estimated and VSM observations at different depths in the profile are assimilated everyday. However, the RZSM ASD for cases when only states are estimated, reduced marginally by 0.001 m^3/m^3 (10%) in comparison with simultaneous state–parameters estimation when near-surface observations are assimilated and the frequency of updates increased from every 10 days to daily. This demonstrated

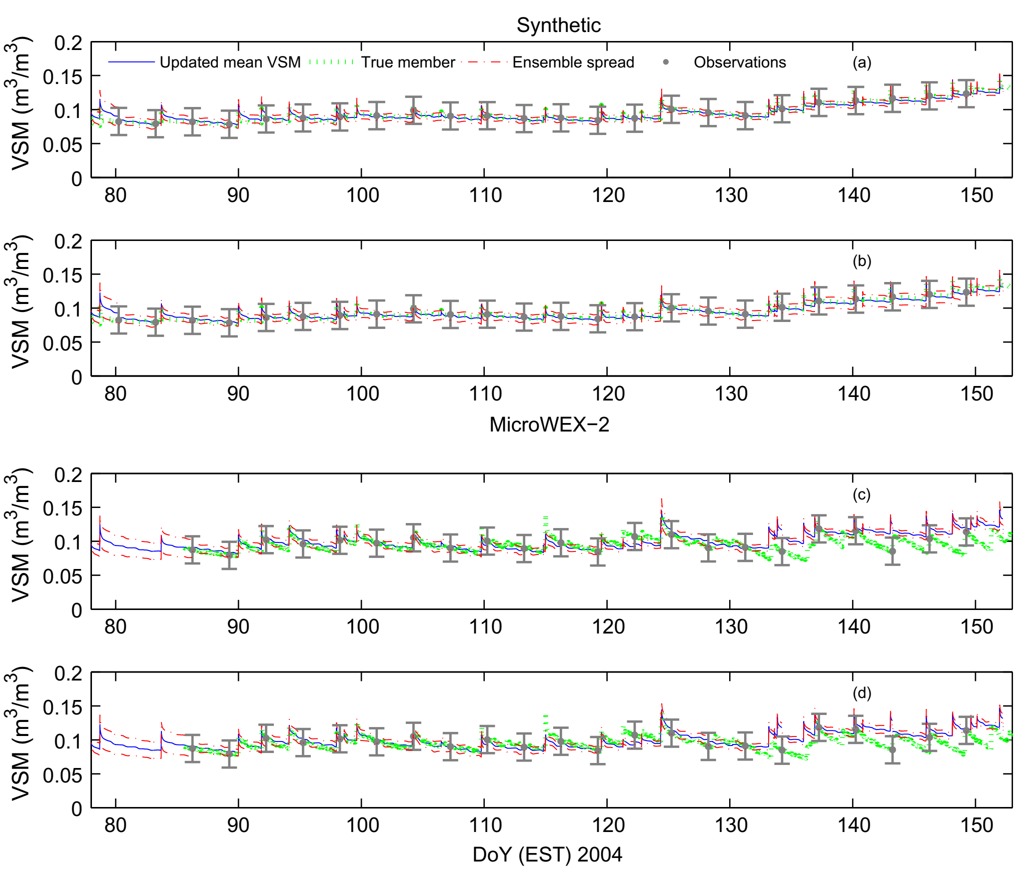

Fig. 9. Mean and standard deviation of RZSM estimates for simultaneous state–parameter estimation using VSM observations at the five depths of 0–5, 8, 32, 64, and 100 cm when assimilation was conducted every 3-days for: (a) synthetic observations without precipitation/irrigation uncertainties, (b) synthetic observations with precipitation/irrigation uncertainties, (c) MicroWEX-2 observations without precipitation/irrigation uncertainties, and (d) MicroWEX-2 observations with precipitation/irrigation uncertainties.

that the RZSM uncertainty in the state-only estimation may not improve substantially even if near-surface observations, as obtained from remotely sensed measurements, are assimilated more frequently. The ASDs and the RMSE/D of RZSM estimates are similar for both daily and every 3-day updates with reductions from the open loop of 77% and 95% in the ASD and RMSE, respectively (see Table 3).

4.4.2. Estimate of parameters

For the synthetic case, 1- and 3-day updates obtain similar parameter values, ASDs, and RMSEs, when assimilating VSM observations at the five depths within the root zone compared to those observed when VSM observations at top 5 cm are assimilated (Table 5). When the parameters are estimated using MicroWEX-2 observations, 1- and 3-day updates estimate different values (see Table 5). Porosity is the only parameter whose estimates are similar for both 1- and 3-day update, which is also the most sensitive parameter.

4.5. Forcing uncertainties

The results discussed so far have assumed perfectly known precipitation/irrigation forcings to the LSP–DSSAT model, with the prediction uncertainty resulting only from unknown parameters. To understand the effect of precipitation/irrigation measurement errors in estimates of VSM and parameters, a Gaussian error, with standard deviation equal to 12% of precipitation/irrigation value observed, was introduced on rainy days. Fig. 7 shows the comparison of the open-loop simulation with perfect forcings and with the case when errors in precipitation/irrigation events were intro-

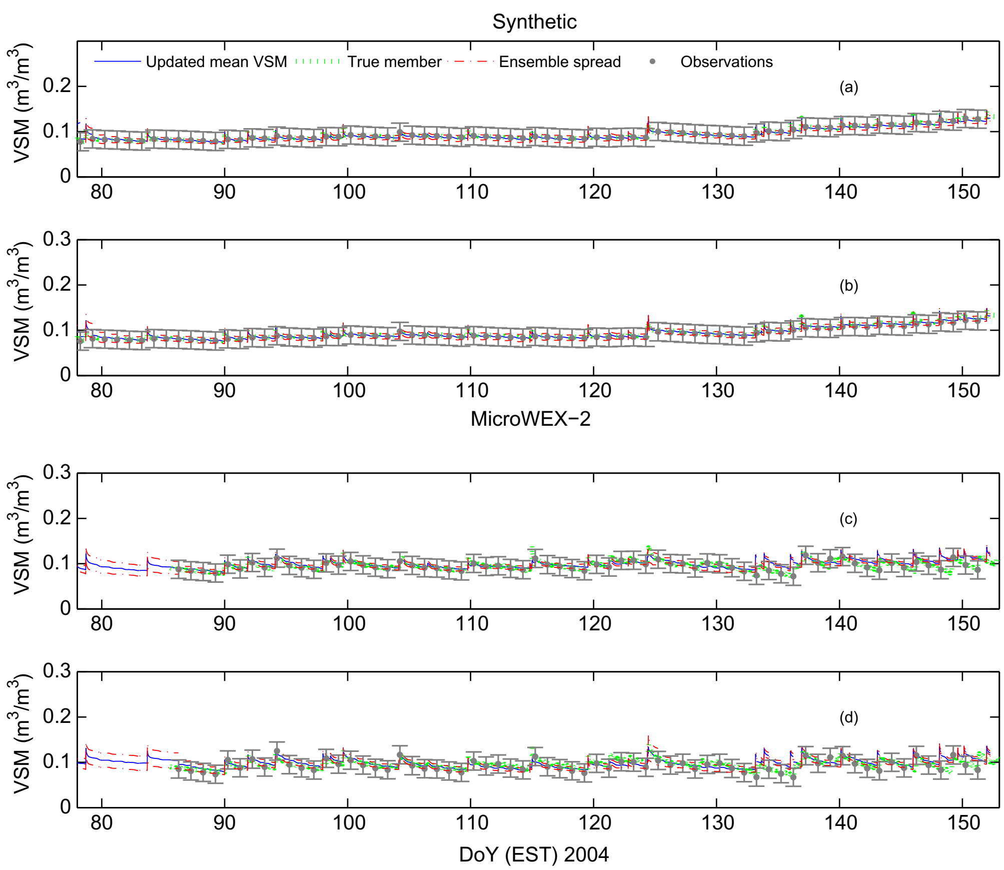

Fig. 10. Mean and standard deviation of RZSM estimates for simultaneous state–parameter estimation using VSM observations with precipitation/irrigation uncertainties when assimilation was conducted every day for: (a) synthetic observations at the five depths of 0–5, 8, 32, 64, and 100 cm, (b) synthetic observations at the two depths of 0–5 and 64 cm, (c) MicroWEX-2 observations at the five depths of 0–5, 8, 32, 64, and 100 cm, and (d) MicroWEX-2 observations at the two depths of 0–5 and 64 cm.

Table 7
RMSE (for synthetic)/RMSD (for MicroWEX-2) and average of the standard deviation (ASD) for the RZSM and VSM at different depths in the soil profile using synthetic and MicroWEX-2 observations that were assimilated daily using two depth combinations. 5-Depths indicates VSM at 0–5, 8, 32, 64, and 100 cm and 2-depths indicates VSM at 0–5 and 64 cm. All simulations presented here include precipitation/irrigation uncertainties.

Depth (cm)	Synthetic				MicroWEX-2			
	5-Depths		2-Depths		5-Depths		2-Depths	
	RMSE	ASD	RMSE	ASD	RMSD	ASD	RMSD	ASD
0–5	0.0011	0.0065	0.0018	0.0069	0.0099	0.0078	0.0094	0.0093
8	0.0009	0.0063	0.0015	0.0066	0.0109	0.0085	0.0137	0.0100
32	0.0011	0.0056	0.0011	0.0061	0.0067	0.0054	0.0091	0.0073
64	0.0015	0.0055	0.0012	0.0065	0.0196	0.0042	0.0172	0.0062
100	0.0028	0.0079	0.0022	0.0112	0.0082	0.0042	0.0071	0.0086
Root zone	0.0016	0.0055	0.0013	0.0060	0.0062	0.0046	0.0067	0.0061

duced. The introduction of the errors did not change the mean RZSM or the ensemble spread significantly.

Table 6 compares the RMSE, the RMSD, and the ASD of RZSM for simultaneous state–parameter estimates when observations at five depths are assimilated every 1- and 3-days, considering errors in precipitation/irrigation for synthetic and MicroWEX-2 observations. Figs. 8 and 9 show the means and standard deviation of RZSM estimates for simultaneous state–parameter estimation

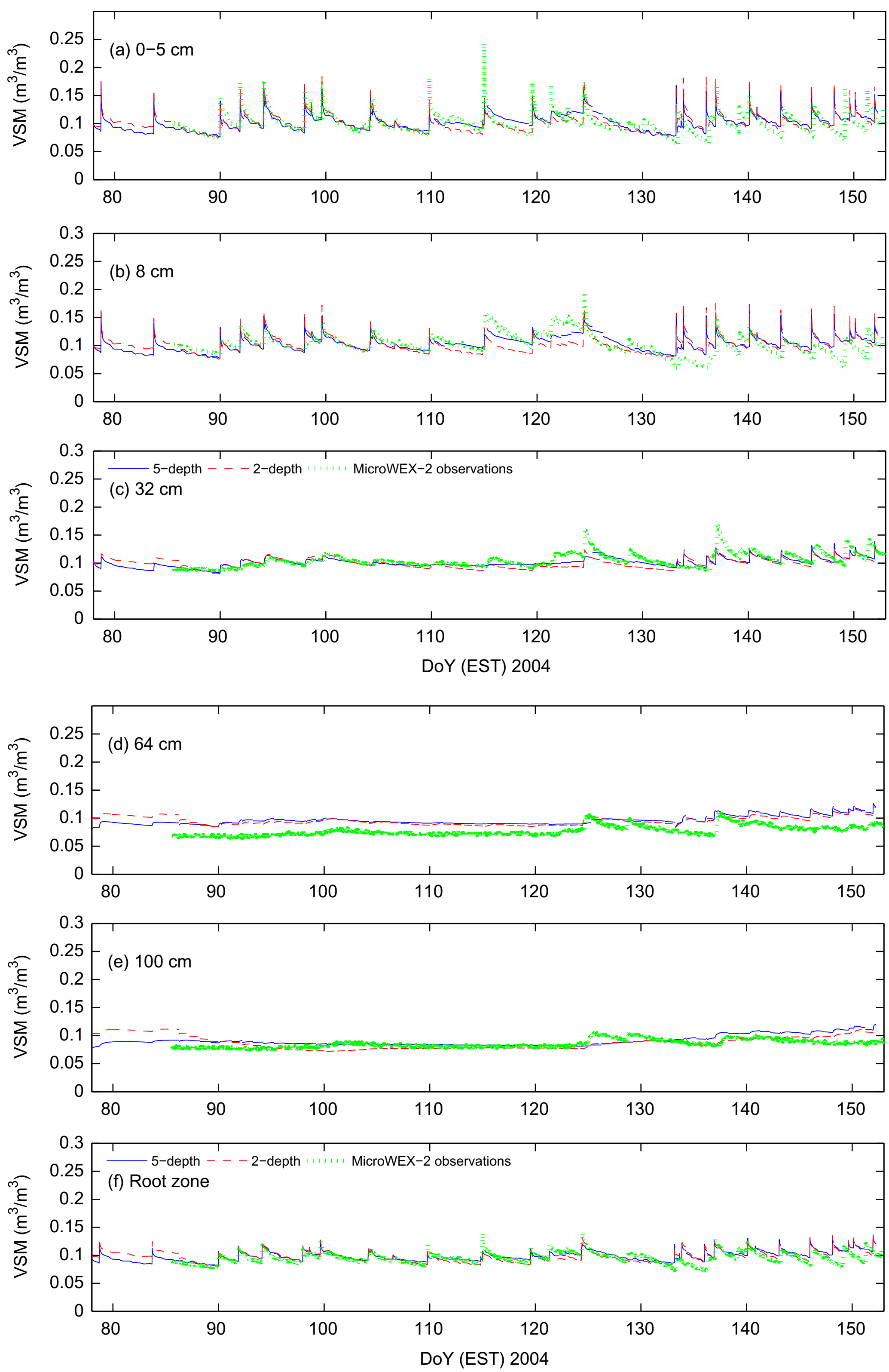

Fig. 11. VSM estimates by the LSP–DSSAT model at depths of (a) 0–5 cm, (b) 8 cm, (c) 32 cm, (d) 64 cm, (e) 100 cm, and (f) in the root zone when MicroWEX-2 observations are assimilated everyday, including uncertainties in precipitation/irrigation for observations at the five depths of 0–5, 8, 32, 64, and 100 cm, and for observations at the two depths of 0–5 and 64 cm.

using VSM observations at five depths when assimilating daily and every 3-days, respectively, synthetic and MicroWEX-2 observations. The ASD of the RZSM estimates do not change significantly for any of the cases, similar to the open-loop simulations. In contrast, introducing uncertainties in precipitation/irrigation does not change substantially the RZSM estimates when either synthetic or MicroWEX-2 observations were assimilated every 3 days (see Fig. 9). The uncertainties introduced by such errors is still low at 0.0056 m^3/m^3 compared to errors in the observations (0.02 m^3/m^3). This is particularly evident in the late and reproductive stages when frequent irrigation was applied (Fig. 9).

4.6. Spatial (vertical) frequency of observations

As mentioned in Section 1, most of the current studies assimilate remotely sensed VSM observations at 0–5 cm. As discussed in Sections 4.3 and 4.4, this study found that assimilation of VSM at different depths in the soil profile results in significant improvement in RZSM and parameter estimates. We investigated this further and compared the results from assimilating VSM observations at 0–5, 8, 32, 64, and 100 cm, and assimilating 0–5 and 64 cm for both synthetic and MicroWEX-2 cases everyday. We selected the depth combination of 0–5 and 64 cm because VSM at 0–5 cm is minimally correlated to soil layers deeper than 64 cm in the LSP–DSSAT model [1]. For both synthetic and MicroWEX-2 observations, the depth combinations obtain similar results for RZSM estimates (see Fig. 10). The RZSM estimates show a higher RMSD for the MicroWEX-2 observations than the RMSE obtained for the synthetic observations.

Table 7 shows the RMSE/RMSD and the ASD for the RZSM estimates and VSM estimates at different depths when considering precipitation/irrigation errors and the five and two depth combinations for synthetic and MicroWEX-2 observations. For both the synthetic and MicroWEX-2 scenarios, the RMSE/RMSD and ASD are not significantly different in the soil moisture profile at <0.010 m^3/m^3 (78% reduction of the open-loop RMSE) and <0.008 m^3/m^3 (69% reduction of the open-loop ASD), respectively, for the two depth combinations. However for MicroWEX-2 scenario, for VSMs at 8 and 32 cm the RMSDs are 0.0137 (41% reduction of the open-loop RMSE) and 0.0091 m^3/m^3 (61% reduction of the open-loop RMSE), respectively, when the VSM observations at 0–5 and 64 cm are assimilated compared to 0.0109 (53% reduction of the open-loop RMSE) and 0.0067 m^3/m^3 (60% reduction of the open-loop RMSE), respectively, when those at the five depths are assimilated. In the synthetic case, the RMSE and ASD of the VSM estimates at different depths in the soil profile are about 0.0010 m^3/m^3 (96% reduction of the open-loop RMSE) and 0.0055 m^3/m^3 (80% reduction of the open-loop RMSE), respectively. In contrast, in the MicroWEX-2 case, the RMSD and ASD range between 0.0062–0.0196 m^3/m^3 (30–80% reduction of the open-loop RMSE) and 0.0042–0.0100 m^3/m^3 (66–85% reduction of the open-loop ASD), respectively. The MicroWEX-2 case has higher RMSDs and ASDs for both the five depths and the two depth assimilation combinations compared to those for the synthetic case. This difference is more evident for VSM estimates of 0–5 and 8 cm, where RMSD is 0.0100 m^3/m^3 (64% reduction of the open-loop RMSE) for MicroWEX-2 case and is 0.0011 m^3/m^3 (95% reduction of the open-loop RMSE) for synthetic case. Fig. 11 pre-

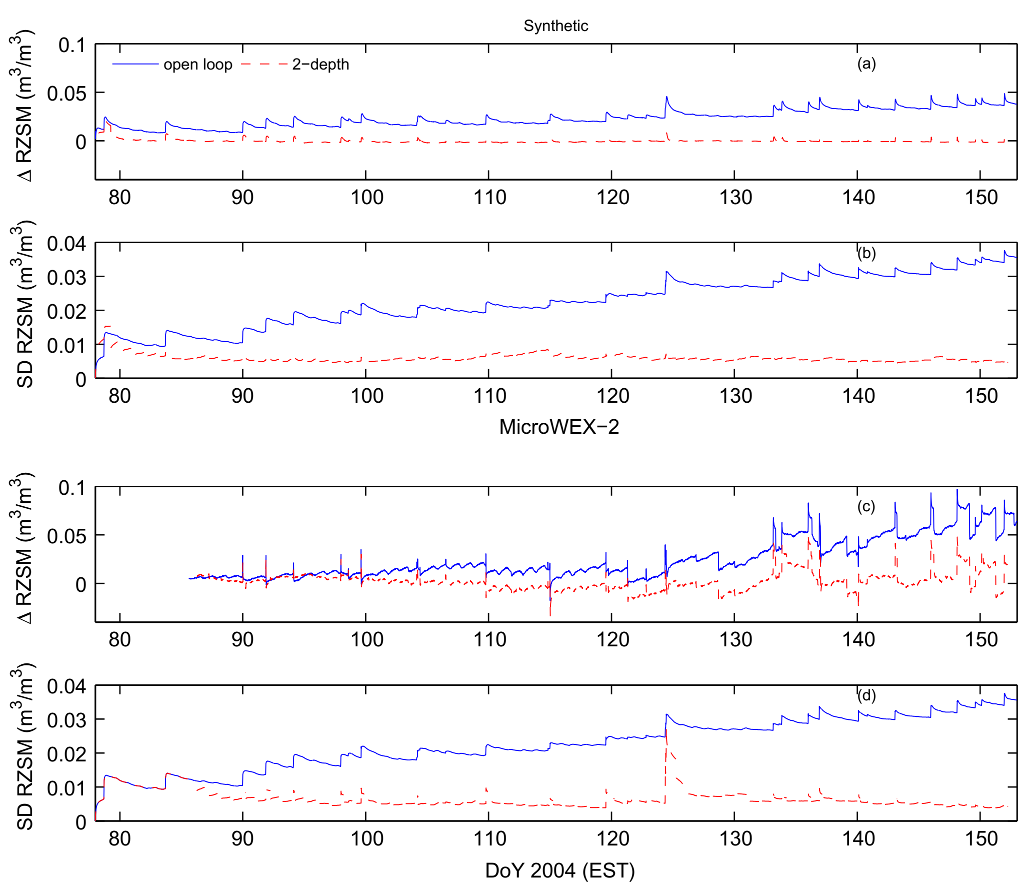

Fig. 12. Differences ($\Delta RZSM$) and standard deviations (SD) of RZSM estimates between the true values and those from the open-loop simulations and from the simultaneous state–parameter estimates assimilating VSM observations at 0–5 and 64 cm every day. (a) $\Delta RZSM$ for synthetic case, (b) SD for synthetic case, (c) $\Delta RZSM$ for MicroWEX-2, and (d) SD for MicroWEX-2.

sents the VSM estimates at different depths in the soil profile for the synthetic and MicroWEX-2 cases when precipitation/irrigation uncertainties are considered and two depth combination. For MicroWEX-2 case, the VSM estimates at the depths of 0–5 and 8 cm are different for the two depth combination scenarios at 0.0157 (30% reduction of the open-loop RMSE) and 0.0123 m^3/m^3 (50% reduction of the open-loop RMSE), respectively. During DoY 105–130, at depths of 0–5 and 8 cm, five-depth assimilation produces VSM estimates closer to the MicroWEX-2 observations. At lower depths, both depth combinations predict similar VSM estimates during the growing season. This demonstrates that detailed profile measurements are not necessary and two combination of depths, near surface and deeper in the soil, provide as accurate estimates of RZSM.

Fig. 12 shows the differences in the RZSM ($\Delta RZSM$) between the truth and the open loop; and between the truth and the case when VSM at 0–5 and 64 cm were assimilated every day. For the MicroWEX-2 case, the field observations were assumed to be the true value. As expected, the updated RZSM estimates are closer to the true values compared to the open loop estimates. At the end of the season, the $\Delta RZSM$ for the open loop were 0.038 and 0.062 m^3/m^3 for the synthetic and MicroWEX-2 cases, respectively (Fig. 12a and c). Daily updates decreased the $\Delta RZSM$ to 0.002 and 0.019 for synthetic and MicroWEX-2 cases, respectively (Fig. 12a and c). The uncertainties, ASD also decreased upon assimilation (Fig. 12b and d). The $\Delta RZSM$s and the ASDs are higher immediately after precipitation/irrigation events, primarily after DoY 130, when MicroWEX-2 observations are assimilated compared to the synthetic case. These differences could further indicate the need to update canopy and root distribution parameters in the DSSAT model, as mentioned in Section 4.3.1.

5. Summary and conclusions

In this study, an EnKF-based assimilation algorithm was implemented and tested to improve RZSM estimates from the coupled LSP–DSSAT model. The objective of this study was to use field observations during a growing season of corn to understand the effects of simultaneous state–parameter estimation, spatial and temporal frequency of updates, and forcing uncertainties on RZSM estimates from the LSP–DSSAT model. A comparison of EnKF performance using high spatio-temporal density observations from the MicroWEX-2 experiment and synthetic observations was conducted to differentiate model errors in biophysics from errors in parameters and forcings.

The RZSM was found to be the most sensitive to four of the 12 parameters in the LSP–DSSAT model, viz., porosity (ϕ), pore-size index (λ), air-entry pressure (ψ_0), and saturated hydraulic conductivity (K_{sat}). Assimilating synthetic observations produced estimates of VSM close to the true values throughout the soil profile with an average error of about 0.2% VSM (90% reduction of the open-loop RMSE) for the entire growing season, whereas assimilating MicroWEX-2 observations produced estimates that were biased high from the observations in the deeper layers with an average error of about 1.5% VSM (40% reduction of the open-loop RMSE) in the top 32 cm. In the deeper layers, the VSM were biased high with an average error of 2.3% VSM (50% reduction of the open-loop RMSE) during the growing season for the MicroWEX-2 observations. This difference between VSM error for the field experiment and synthetic case indicated that the field observations did not follow the model assumptions (i.e. homogeneous soil) and/or physics perfectly, unlike the synthetic observations.

For both synthetic and MicroWEX-2 observations, lower average ASD and RMSE were obtained during the growing season when states and parameters were updated simultaneously than when only states were updated. For the synthetic case, simultaneous state–parameter updates produced consistent and accurate parameter estimates, with updated parameter estimates within one standard deviation of the true value, for all combinations of temporal and spatial observation frequencies. However, the simultaneous state–parameter estimation using MicroWEX-2 observations produced substantially different parameter estimates when using VSM observations at five depths (0–5, 8, 32, 64, and 100 cm) than when using only near-surface (0–5 cm) observations. This again pointed out the possibility of unaccounted for model error due to heterogeneous soil and vegetation conditions in the actual field experiment. However, for state-only estimates, the RZSM ASD reduced marginally by 0.4% VSM (30% in comparison with simultaneous state–parameter estimates) when near-surface observations were incorporated and the frequency of updates increased from every 10 days to daily. This demonstrated that the RZSM uncertainty in the state-only estimation may not improve substantially even if near-surface observations, as obtained from remotely sensed measurements, are assimilated more frequently.

When forcings were assumed to be known perfectly, the simultaneous state–parameter estimates with daily or 3-day updates resulted in the lowest uncertainty in RZSM estimation when VSM observations at the five depths were assimilated. Introducing errors in precipitation/irrigation forcings marginally increased the RZSM ASD by 0.02% VSM (5%) and the RZSM RMSE/RMSD by 0.02% VSM (2%) in comparison to the perfectly-known-forcing case using either synthetic or MicroWEX-2 observations for every 3 days assimilation, and the RZSM ASD and the RZSM RMSE/RMSD also marginally increased to 0.02% VSM (4%) and 0.02% VSM (3%), respectively, when observations were assimilated every day. Comparing different depth combinations to estimate RZSM, it was found that VSM observations at all depths (0–5, 8, 32, 64, and 100 cm) produced similar results to using only 0–5 and 64 cm observations. This demonstrated that detailed profile measurements were not necessary in this study and a two-depth combination, one at near-surface and one deeper in the soil, provided as similar estimates of RZSM. Remotely-sensed near-surface observations could be combined with *in situ* VSM observations from sites such as the Soil Climate Analysis Network (S.C.A.N.) [51] to produce similar results.

Despite realistic uncertainties attributed to soil parameters and precipitation/irrigation forcing, the uncertainties in VSM predictions by the model were much lower than the assumed observation errors. This resulted in relatively little change in either the mean or standard deviation of VSM estimates when spatial and temporal observation frequency were varied. In this study, only the errors in soil and precipitation/irrigation forcings were considered. Thus, the observed differences in EnKF performance between synthetic and field observations may indicate errors in model biophysics that were not considered here, such as those in soil conditions or predictions of LAI and root distribution.

Acknowledgements

This research was supported by the NSF Earth Science Directorate (EAR-0337277) and the NASA New Investigator Program (NASA-NIP-00050655). The authors acknowledge computational resources and support provided by the University of Florida High-Performance Computing Center for the simulations conducted in this study.

References

[1] Agrawal D, Monsivais-Huertero A, Graham WD, Judge J. Parameter sensitivity analysis for root zone soil moisture in SVAT models. In: 2008 Joint assembly. Fort Lauderdale, FL, USA: Amer. Geophys. Union; 2008.

[2] Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J, et al. LAPACK user's guide, 3rd ed. Philadelphia, PA: Soc. for Ind. and Applied Mathem.; 1999.
[3] Anderson JL. An ensemble adjustment Kalman filter for data assimilation. Mon Weath Rev 2001;129(12):2884–903.
[4] Baek S, Hunt BR, Kalnay E, Ott E, Szunyogh I. Local ensemble Kalman filtering in the presence of model bias. Tellus A 2006;58(3):293–306.
[5] Barre MJP, Duesmann B, Kerr YH. SMOS: the mission and the system. IEEE Trans Geosci Remote Sens 2008;46(3):587–93.
[6] Campbell Scientific. Cs616 and cs625 water content reflectometers. Tech. report, Logan, UT; 2006. %3chttp://www.campbellsci.com/documents/manuals/cs616.pdf%3e.
[7] Casanova JJ, Judge J. Estimation of energy and moisture fluxes for dynamic vegetation using coupled SVAT and crop-growth models. Water Resour Res 2008;44:W07415. doi:10.1029/2007WR006503.
[8] Casanova JJ, Judge J, Jones JW. Calibration of the ceres-maize model for linkage with a microwave remote sensing model. Trans Am Soc Agric Eng 2006;49(3):783–92.
[9] Chaubey I, Hann CT, Grunwald S, Salisbury JM. Uncertainty in the model parameters due to spatial variability of rainfall. J Hydrol 1999;220:48–61.
[10] Chung Y. A snow–soil–vegetation–atmosphere-transfer/radiobrightness model for wet snow. PhD thesis, The University of Michigan; 2007.
[11] Ciach GJ. Local random errors in tipping-bucket rain gauge measurements. J Atmos Ocean Tech 2003;20(1–2):752–9.
[12] Crow WT, Chan STZ, Entekhabi D, Houser PR, Hsu AY, Jackson TJ, et al. An observing system simulation experiment for hydros radiometer-only soil moisture products. IEEE Trans Geosci Remote Sens 2005;43(6):1289–303.
[13] Crow WT, Kustas WP, Prueger JH. Monitoring root-zone soil moisture through the assimilation of a thermal remote sensing-based soil moisture proxy into a water balance model. Remote Sens Environ 2008;112(4):1268–81.
[14] Crow WT, Wood EF. 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 2003;26(2):137–49.
[15] De Lannoy GJM, Houser PR, Pauwels VRN, Verhoest NEC. State and bias estimation for soil moisture profiles by an ensemble Kalman filter: effect of assimilation depth and frequency. Water Resour Res 2007;43:W06401. doi:10.1029/2006WR005100.
[16] De Lannoy GJM, Reichle RH, Houser PR, Pauwels VRN, Verhoest NEC. Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. Water Resour Res 2007;43:W09410. doi:10.1029/2006WR005449.
[17] de Vries D. Simultaneous transfer of heat and moisture in porous media. Trans Am Geophys Union 1958;39(5):909–16.
[18] Entekhabi D, Nakamura H, Njoku EG. Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. IEEE Trans Geosci Remote Sens 1994;32(2):438–48.
[19] Evensen G. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res 1994;99(C5):10143–62.
[20] Evensen G. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynam 2003;53(4):343–67.
[21] Garcia-Quijano JF, Barros AP. Incorporating canopy physiology into a hydrological model: photosynthesis, dynamic respiration, and stomatal sensitivity. Ecol Model 2005;185(1):29–49.
[22] Gelb A. Applied optimal estimation. The MIT Press; 1974. p. 102–55 [Ch. 4].
[23] Goudriaan J. Crop micrometeorology: a simulation study. 1st ed. Wageningen, The Netherlands: Centre for Agricultural Publishing and Documentation; 1977.
[24] Habib E, Krajewski WF, Kruger A. Sampling errors of tipping-bucket rain gauge measurements. J Hydrol Eng 2001;6(2):159–66.
[25] Hanson JD, Rojas KW, Schaffer MJ. Calibrating the root zone water quality model. Agron J 1999;91:171–7.
[26] Heathman GC, Starks PJ, Ahuja LR, Jackson TJ. Assimilation of surface soil moisture to estimate profile soil water content. J Hydrol 2003;279(1–4):1–17.
[27] Hoeben R, Troch PA. Assimilation of active microwave observation data for soil moisture profile estimation. Water Resour Res 2000;36(10):2805–19.
[28] Houser PR, Shuttleworth WJ, Famiglietti JS, Gupta HV, Syed KH, Goodrich DC. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. Water Resour Res 1998;34(12):3405–20.
[29] Huang C, Li X, Gu J. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. Remote Sens Environ 2008;112(3):888–900.
[30] Jazwinski AH. Stochastic processes and filtering theory. Mathematics in science and engineering, vol. 64. Academic Press; 1970.
[31] Jones JW, Hoogenboom G, Poter C, Boote K, Batchelor WBL, Hunt LA, et al. The DSSAT cropping system model. Eur J Agron 2003;18(3–4):235–65.
[32] Judge J, Abriola L, England A. Numerical validation of the land surface process component of an LSP/R model. Adv Water Resour 2003;26(7):733–46.
[33] Judge J, Casanova JJ, Lin T, Tien KC, Jang M, Lanni O, et al. Field observations during the second microwave, water, and energy balance experiment (MicroWEX-2): from March 17 through June 3, 2004. Circular no. 1480, tech. rep., Center for Remote Sensing, University of Florida; 2005. %3chttp://www.edis.ifas.ufl.edu/AE360%3e.
[34] Judge J, England A, Laymon CA, Crosson C, Hornbuckle B, Boprie WL, et al. A growing season land surface process/radiobrightness model for wheat-stubble in the southern great plains. IEEE Trans Geosci Remote Sens 1999;37(5):2152–8.
[35] Judge J, England AW, Metcalfe JR, McNichol D, Goodison BE. Calibration of an integrated land surface process and radiobrightness (LSP/R) model during summertime. Adv Water Resour 2008;31(1):189–202.
[36] Kerr Y, Waldteufel P, Wigneron JP, Martinuzzi JM, Font J, Berger M. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Trans Geosci Remote Sens 2001;39(8):1729–35.
[37] Lorenc A, Hammon O. Objective quality control of observations using Bayesian methods: theory, and a practical implementation. Q J R Meteorol Soc 1988;114(480):515–43.
[38] Manabe S. Climate and the ocean circulation. 1. The atmospheric circulation and the hydrology of the earth's surface. Mon Weath Rev 1969;97(11):739–74.
[39] Mo X, Liu S, Xu Z, Xiang Y, McVicar T. Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data from the North China Plain. Ecol Model 2005;183(2–3):301–22.
[40] Moradkhani H, Sorooshian S, Gupta HV, Houser PR. Dual state–parameter estimation of hydrological models using Ensemble Kalman filter. Adv Water Resour 2005;28(2):135–47.
[41] Muñoz-Sabater J, Rüdiger C, Calvet JC, Fritz N, Jarlan L, Kerr Y. Joint assimilation of surface soil moisture and LAI observations into a land surface model. Agric Forest Meteorol 2008;148(8–9):1362–73.
[42] Pauwels VRN, Verhoest NEC, De Lannoy GJM, Guissard V, Lucau C, Defourny P. Optimization of a coupled hydrology-crop growth model through the assimilation of observed soil moisture and leaf area index using an ensemble Kalman filter. Water Resour Res 2007;43:W04421. doi:10.1029/2006WR004942.
[43] Peixoto J, Oort A. Physics of climate. New York: American Institute of Physics; 1992.
[44] Philip J, de Vries D. Moisture movement in porous materials under temperature gradients. Trans Am Geophys Union 1957;38(2):222–32.
[45] Reichle RH, Crow WT, Kepenne CL. An adaptive ensemble Kalman filter for soil moisture data assimilation. Water Resour Res 2008;44:W03423. doi:10.1029/2007WR006357.
[46] Reichle RH, Koster RD. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. Geophys Res Lett 2005;32:L02404. doi:10.1029/2004GL021700.
[47] Reichle RH, Koster RD, Liu P, Mahanama SPP, Njoku EG, Owe M. Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SSMR). J Geophys Res 2007;112:D09108. doi:10.1029/2006JD008033.
[48] Reichle RH, McLaughlin DB, Entekhabi D. Hydrologic data assimilation with the Ensemble Kalman filter. Mon Weath Rev 2002;130(1):103–14.
[49] Reichle RH, Walker JP, Koster RD, Houser PR. Extended versus ensemble Kalman filter for land data assimilation. J Hydrometeorol 2002;3(6):728–40.
[50] Rossi C, Nimmo JR. Modeling of soil water retention from saturation to oven dryness. Water Resour Res 1994;30(3):701–8.
[51] S.C.A.N. Soil Climate Analysis Network (S.C.A.N.); 2009. %3chttp://www.wcc.nrcs.usda.gov/scan/%3e.
[52] Thiemann M, Trosset M, Grupta H, Sorooshian S. Bayesian recursive parameter estimation for hydrologic models. Water Resour Res 2001;37(10):2521–35.
[53] Trenberth KE. Climate system modeling. New York: Cambridge University Press; 1995.
[54] Verseghy S, McFarlane N, Lazare M. CLASS-A Canadian land surface scheme for GCMs, II. Vegetation model and coupled runs. Int J Climatol 1993;13(4):347–70.
[55] Vrugt JA, Gupta HV, Bouten W, Sorooshian SA. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res 2003;39(8). doi:10.1029/2002WR001642.
[56] Walker JP, Houser PR. Requirements of a global near-surface soil moisture satellite mission: accuracy, repeat time, and spatial resolution. Adv Water Resour 2004;27(8):785–801.
[57] Western AW, Bloschl G. On the spatial scaling of soil moisture. J Hydrol 1999;217:203–24.
[58] Whitfield B, Jacobs J, Judge J. Intercomparison study of the land surface process model and the common land model for a prairie wetland in Florida. J Hydrometeorol 2006;7(6):1247–58.
[59] Zhan X, Houser PR, Walker JP, Crow WT. A Method for retrieving high-resolution surface soil moisture from Hydros L-band radiometer and radar observations. IEEE Trans Geosci Remote Sens 2006;44(6):1534–44.