Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Advances in Water Resources xxx (2005) xxx–xxx
Advances in Water Resources
www.elsevier.com/locate/advwatres

Numerical investigation of the impact of uncertainties in satellite rainfall estimation and land surface model parameters on simulation of soil moisture

Faisal Hossain [a], Emmanouil N. Anagnostou [b,*]

[a] Department of Civil and Environmental Engineering, Tennessee Technological University, Cookeville, TN 38505, United States
[b] Department of Civil and Environmental Engineering, University of Connecticut, 261 Glenbrook Road, U2037, Storrs, CT 06269, United States

Received 15 September 2004; received in revised form 20 February 2005; accepted 30 March 2005

Abstract

This study aims at evaluating the uncertainty in the prediction of soil moisture (1D, vertical column) from an offline land surface model (LSM) forced by hydro-meteorological and radiation data. We focus on two types of uncertainty: an input error due to satellite rainfall retrieval uncertainty, and, LSM soil-parametric error. The study is facilitated by in situ and remotely sensed data-driven (precipitation, radiation, soil moisture) simulation experiments comprising a LSM and stochastic models for error characterization. The parametric uncertainty is represented by the generalized likelihood uncertainty estimation (GLUE) technique, which models the parameter non-uniqueness against direct observations. Half-hourly infra-red (IR) sensor retrievals were used as satellite rainfall estimates. The IR rain retrieval uncertainty is characterized on the basis of a satellite rainfall error model (SREM). The combined uncertainty (i.e., SREM + GLUE) is compared with the partial assessment of uncertainty. It is found that precipitation (IR) error alone may explain moderate to low proportion of the soil moisture simulation uncertainty, depending on the level of model accuracy—50–60% for high model accuracy, and 20–30% for low model accuracy. Comparisons on the basis of two different sites also yielded an increase (50–100%) in soil moisture prediction uncertainty for the more vegetated site. This study exemplified the need for detailed investigations of the rainfall retrieval-modeling parameter error interaction within a comprehensive space-time stochastic framework for achieving optimal integration of satellite rain retrievals in land data assimilation systems.

Keywords: Uncertainty; Land surface model; Satellite rain retrievals; Parameter uncertainty; Soil moisture

1. Introduction

Soil moisture, defined as the water content in the upper layer of soil, is the hydrologic variable that controls the interactions (and feedbacks) between land surface and atmospheric processes. Over the last decade, a sizeable body of literature has accumulated on accurate characterization of spatio-temporal variability of soil moisture [35,1,32,28,23,16] (Hornberger et al., 2001); among others. Recent trends indicate that soil moisture estimates are increasingly derived from land data assimilation systems (LDAS) that comprise a physically based land surface model (LSM) forced by hydro-meteorological data and schemes for integrating remotely sensed observations (see for example [39,33,28]). Of particular importance therefore is the characterization of uncertainty in the prediction of soil moisture by LSMs. This facilitates the statistical (ensemble) pre-storm initialisation of distributed hydrologic models used in the prediction of floods, and the development of ensemble land surface boundary conditions for regional atmospheric

* Corresponding author. Tel.: +860 486 6806; fax: +860 486 2298.
E-mail address: manos@engr.uconn.edu (E.N. Anagnostou).

0309-1708/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.advwatres.2005.03.013

models issuing short to medium-range quantitative precipitation forecasts.

Knowledge of uncertainty is also a key to advance the efficiency of land data assimilation techniques [39,29]. One example of such a technique that requires characterization of model prediction (or forecast) error is the Kalman filtering approach used in LDAS, currently in operation over North America [33,27] and globally [34]. Accurate knowledge of soil moisture prediction error is essential for optimal performance of assimilation systems. In this study we seek to address the characterization of two types of uncertainty in soil moisture prediction on the basis of LSM: (1) errors in the rainfall estimation from satellite sensor observations; and, (2) the LSM parametric error (manifesting as non-uniqueness in soil hydraulic parameters). Parameter non-uniqueness is a phenomenon where no single parameter set uniquely defines the model state, but rather it is the ensemble that defines the universal set of equi-probable states. The non-uniqueness property is attributed to the simultaneous effect of various uncertainty sources such as, model formulation error, error due to initial/boundary conditions, input error etc. In this study we define parametric uncertainty as an attribute primarily identifiable with (but not limited to) LSM model formulation error. Hereafter, no distinction is made between the terms parametric uncertainty and modelling error (or model formulation error).

On one hand, understanding the error in rainfall measurement has implications for the global precipitation measurement (GPM), which is a mission to be launched by the international community by 2010 [36,6]. GPM is expected to provide rainfall measurements from space at scales finer than what is globally available today. Hence, satellite rainfall estimates would gradually become the dominant component of the atmospheric forcing data for LDAS. On the other hand, because current LDAS formulations have provisions for using multiple state-of-the-art LSMs in the assimilation technique [33], understanding the role played by model formulation error is as significant as error in satellite rainfall measurements.

A simultaneous assessment of uncertainty would allow comparisons with partial assessments (i.e., satellite rainfall or modeling error alone), which can potentially identify the implications of each error source (and hence a strategy for reducing the prediction uncertainty). In hydrologic remote sensing, the interactions between both sources of uncertainty imply that partial assessment may not be adequate to describe the full range of variability in soil moisture prediction. Studying the statistical characterization of soil moisture prediction error associated with both uncertainty sources is the key to advance the use of satellite rainfall remote sensing in LDAS. A similar paradigm of scientific inquiry for catchment-scale hydrology has proved useful in enhancing the usage of radar rainfall estimation in flood simulations of complex terrain basins [20]. Recently, collective research effort has emerged on this aspect called the Distributed Model Intercomparison Project (DMIP) [37]. DMIP provides an insightful treatment of the relative roles of uncertainties in input and model parameters for hydrologic models [37,10,9]. These studies investigated the uncertainty in stream flow prediction from a distributed rainfall-runoff model using ground radar estimates as rain input. However, the approach employed in those studies for error modeling of rainfall estimates omit satellite rainfall, which require a more complex error modeling strategy than pure random sampling [19]. Unlike radar rainfall estimation, where after careful quality control and error adjustments the residual error is left with primarily a random deviation component, the satellite rain retrieval uncertainty is associated with correlated rain, no-rain detection and false alarm error characteristics as well as systematic and random rain rate error components that have longer spatio-temporal correlation lengths (Hossain and Anagnostou, 2005). Furthermore, in terms of significance in large-scale water resources management, satellite data-land surface modeling system has a distinct role in hydrology given the global availability of satellite rainfall observations.

Thus, although work has been done of radar rainfall error propagation in stream flow simulation, much less is known about the hydrologic applications of satellite data. The important question about satellite rainfall data therefore concerns its error propagation through LSM in simulating soil moisture fields. A major difficulty of such investigation is caused by the complex error structure of satellite rainfall retrievals [19] and the non-linearity in error propagation. In that respect, our study seeks to address the relative impacts and interactions of uncertainty in satellite rainfall retrieval and LSM simulations (i.e., model parameter non-uniqueness). The numerical investigations presented in this paper are limited to single column (1D, vertical) simulations ignoring spatial error characteristics. The model parameter uncertainty is represented by the generalized likelihood uncertainty estimation (GLUE [4]) technique. Uncertainty in satellite rain retrieval is modeled on the basis of a satellite rainfall error model (SREM) developed by Hossain and Anagnostou [19]. The combined assessment of uncertainty—due to rainfall and modeling (SREM + GLUE)—is compared with partial assessments that accounted for either the modeling (GLUE only) or rain retrieval uncertainty (SREM only). Although it is recognized that there may be other important uncertainty criteria that are not examined here (such as, the detailed role of vegetation/climate and the effect of scale), we hope that this paper will provide a defendable proof-of-concept to trigger further studies involving a wide range of model structures, resolutions and objectives towards optimal integration of remotely sensed data in assimilation systems.

The paper is organized as follows. In Section 2 we describe the study region and data. In Section 3 we describe the LSM used in this study. In Section 4, we present the stochastic models used to characterize the uncertainty in the two error sources (satellite rainfall and LSM). In Section 5 we describe the simulation framework and present the results. In Section 6, we present the major conclusions and suggested extensions of this study.

2. Study region and data

Two regions were chosen for the study: (1) Champaign in Illinois; and, (2) Perkins in Oklahoma (hereafter the two regions will be abbreviated as "CHAMPAIGN" and "PERKINS", respectively). CHAMPAIGN is a farmland located 40.01° North and 88.37° West. The site characteristics are typical of those found throughout Midwestern US with most of the land in agricultural production. The soil is silt loam with a bulk density of 1.5 gm/cm^3. The study period is 1 year (1998) when soybeans were planted on the farm. Atmospheric and radiation forcing data from a flux measuring system installed in the farm were recorded every 30 min for that year. The major atmospheric data comprised rainfall, temperature, humidity, surface pressure and wind. The radiation forcing data pertained to downward solar (short-wave) and downward long-wave radiation flux measurements. Soil moisture measurements at only the 5 cm depth are considered here, because at deeper depths the measurements are considered suspect (personal communication with Kenneth Mitchell of NOAA). This data is public domain and available as part of standardized testing protocols for simulation codes of the NOAH-LSM (discussed next). For more information on the study region and data measurement protocols the reader is referred to the User's Guide (ftp://ftp.emc.ncep.noaa.gov/mmb/gcp/ldas/noahlsm/ver_2.5).

PERKINS is located at 35.99° North and 97.05° West in Payne County, Oklahoma. It has perennial ground cover, intermittent farmland with broad-leaf deciduous trees. All requisite hydro-meteorological data except the downward long-wave radiation for NOAH-LSM operation are available at half-hourly intervals from the Oklahoma Mesonet network. The Oklahoma Mesonet [15,8] is a dense network of 114 automated measurement stations across Oklahoma (for more details on Mesonet see http://www.mesonet.ou.edu). The average composition of soil in a 75 cm column at PERKINS is about 45% sand, 35% silt and 20% clay and is classified as sandy-clay loam. Apart from vegetation, the most unique feature distinguishing PERKINS from CHAMPAIGN is that consistent soil moisture measurements are available at deeper depths: 5 cm, 25 cm and 60 cm from surface, thus allowing the soil vertical column to be modeled in the root zone. The year under study for PERKINS is 2001. Long-wave radiation was calculated from in situ measurements of air temperature (T_{air}, °C) and relative humidity (RLH, %) using methods outlined in Bras [7] as follows. First, the saturation vapor pressure was calculated from air temperature [7]. Using knowledge of relative humidity, the ambient vapor pressure was then computed. The long-wave emissivity was derived from Idso [25]. Finally, we used the Stefan–Boltzmann law to estimate the long-wave radiation (W/m^2). It is appropriate to mention that there is a potential limitation in the manner in which long-wave radiation is computed for PERKINS. It may be associated with errors during overcast conditions where cloud cover fraction would need to be factored in. We believe that such a potential limitation alone should not hamper our overall investigation, and particularly so when our intention is to primarily understand the role of uncertainties in satellite precipitation and model parameters. The weaknesses of this approach, if any, may be revealed in our results, and as a result, future studies may also employ more appropriate data sources for long-wave radiation (such as the atmospheric radiation measurement—ARM/CART network).

3. The land surface model

3.1. Model description

The LSM used in this study is the NOAH-LSM (also known as The Community NOAH-LSM—[11,31]). We chose NOAH-LSM as it is a popular operational model with a long heritage and more importantly, it is one of the four LDAS LSMs currently being evaluated over the United States [33]. NOAH-LSM is a stand-alone, uncoupled (offline), column (1-D) version used to execute single-site land surface simulations at 30 min intervals. NOAH-LSM is based on a typical one-dimensional soil-vegetation-atmosphere transfer (SVAT) approach that solves the coupled energy and water budgets at the land surface and within the unsaturated zone. In this traditional 1-D uncoupled mode, near surface atmospheric and radiation data are required as input forcing. NOAH-LSM simulates soil moisture (both liquid and frozen), soil temperature, snow pack, depth, snow pack water equivalent, canopy water content and the energy and water flux terms in terms of the surface energy balance and surface water balance. A four-layer soil configuration (comprising a total depth of 2 m) is adopted in the NOAH-LSM for capturing daily, weekly and seasonal evolution of soil moisture and mitigating possible truncation error in discretization [38]. The lower 1-m acts as gravity drainage at the bottom, and the upper

1-m of soil serves as root zone depth. A resistance approach is used to account for both aerodynamic and vegetation controls on energy fluxes. For more details on the physical description of the model, one may refer to Sridhar et al. [38], Margulis et al. [28] and Chen et al. [11]. In line with the minimum requirements for spin-up [14], a repeat run of NOAH-LSM was made with the year-long available data for each site to achieve equilibrium initial conditions. Since rainfall would have insignificant interaction with the frozen soil column during snow covered times (and transform mostly as surface runoff—an assumption we make), we truncated our effective study period for CHAMPAIGN and PERKINS to the 1 May–30 October, 1998 and 2001 periods, respectively.

3.2. Model fine-tuning

Our preliminary investigation with NOAH-LSM found it necessary to adjust NOAH-LSM vegetation parameter of 'fraction of green vegetation to make the model more representative of the point-scale soil moisture flux simulations at the two study regions. This procedure is essentially based on mild nudging (within physically acceptable limits) to force soil moisture simulations to mimic observations as closely as possible (see [21] for details). After the nudging procedure, the PERKINS fraction of green vegetation values were 15–20% higher than those for CHAMPAIGN, numerically manifesting the difference in vegetation between the two sites. Fig. 1a and b show the effect of fine-tuning during the study period for CHAMPAIGN and PERKINS, respectively. It is seen that NOAH-LSM is able to simulate the soil moisture variability at the 5 cm depth (for CHAMPAIGN) and at 5, 25 and 60 cm depths for PERKINS. The overall correlation of model predicted to measured soil moisture was calculated to be 0.8 (0.9) for PERKINS (CHAMPAIGN).

3.3. Model parameter uncertainty

NOAH-LSM parameter (model) uncertainty was accounted for the following five soil hydraulic parameters that we considered most sensitive to soil moisture simulation: (1) maximum volumetric soil moisture content (porosity) (SMCMAX, m^3/m^3); (2) saturated matric potential (PSISAT, m) (3) saturated hydraulic conductivity K (SATDK, $m\,s^{-1}$); (4) parameter 'B' of soil-water retention model of Clapp and Hornberger [12] (BB); and (5) soil moisture wilting point at which ET ceases (SMCWLT, m^3/m^3). The range (upper/lower) and optimal values for those parameters are shown in Table 1. These values were selected based on empirical studies by Clapp and Hornberger [12] and Cosby et al. [13], in situ land surface information, and considering the sampling requirements of GLUE [4]. We assume that the parameter uncertainty domain represented by the 5-D hyperspace characterizes adequately the parameter non-uniqueness, which is responsible for the modeling uncertainty in soil moisture simulation.

We have chosen GLUE as the framework to characterize the model parameter uncertainty in the NOAH-LSM formulation for the simulation of soil moisture. It is based on Monte Carlo (MC) simulation: a large number of model runs are performed, each with random parameter values each sampled from uniform probability distribution (e.g., Table 1). The acceptability of each run is assessed through comparison of the predicted versus observed hydrologic variables on the basis of a selected likelihood measure. Simulations with likelihood values below a certain threshold are rejected as non-behavioural. The likelihoods of these non-behavioural parameters are set to zero and are thereby removed from subsequent analysis. Following the rejection of non-behavioral runs, the likelihood weights of the retained (i.e., behavioral) runs are rescaled so that their cumulative total is one [17]. In this study the GLUE method was applied to uncertainty estimation of soil moisture simulation by NOAH-LSM at the 5 cm depth (for CHAMPAIGN) and at 5 cm, 25 cm and 60 cm depth (for PERKINS). Thus at each time step (at 30 min intervals), the predicted soil moisture from the behavioral runs are likelihood weighted and ranked to form a cumulative distribution of soil moisture simulation from which quantiles can be used to represent modelling uncertainty. While GLUE is based on a Bayesian conditioning approach, the likelihood measure is achieved through a goodness of fit criterion as a substitute for a more traditional likelihood function. We have considered the classical index of efficiency, E_{NS} [30] as the measure of likelihood,

$$E_{\mathrm{NS}} = \left[1 - \frac{\sigma_{\mathrm{e}}^2}{\sigma_{\mathrm{obs}}^2}\right] \tag{1}$$

where σ_{e} is the variance of errors and σ_{obs} the variance of soil moisture observations, computed over the entire study period. A point to note is that for PERKINS, the index of efficiency was computed as a depth-weighted average (weighted by the thickness of each soil layer). This yielded an aggregate measure of model accuracy that could be used to select parameters representative of the vertical soil column of the root zone.

To implement the GLUE methodology, each parameter of NOAH-LSM was specified the range of possible values shown in Table 1. Constant (calibrated) values for all other NOAH-LSM parameters were used. Model predictions of soil moisture were carried out, and the model likelihood measure was calculated using the efficiency index of Eq. (1). From the specified parameter ranges, MC simulations were conducted that allowed the selection of a large number of behavioral parameter sets characterized by a simulation efficiency index value

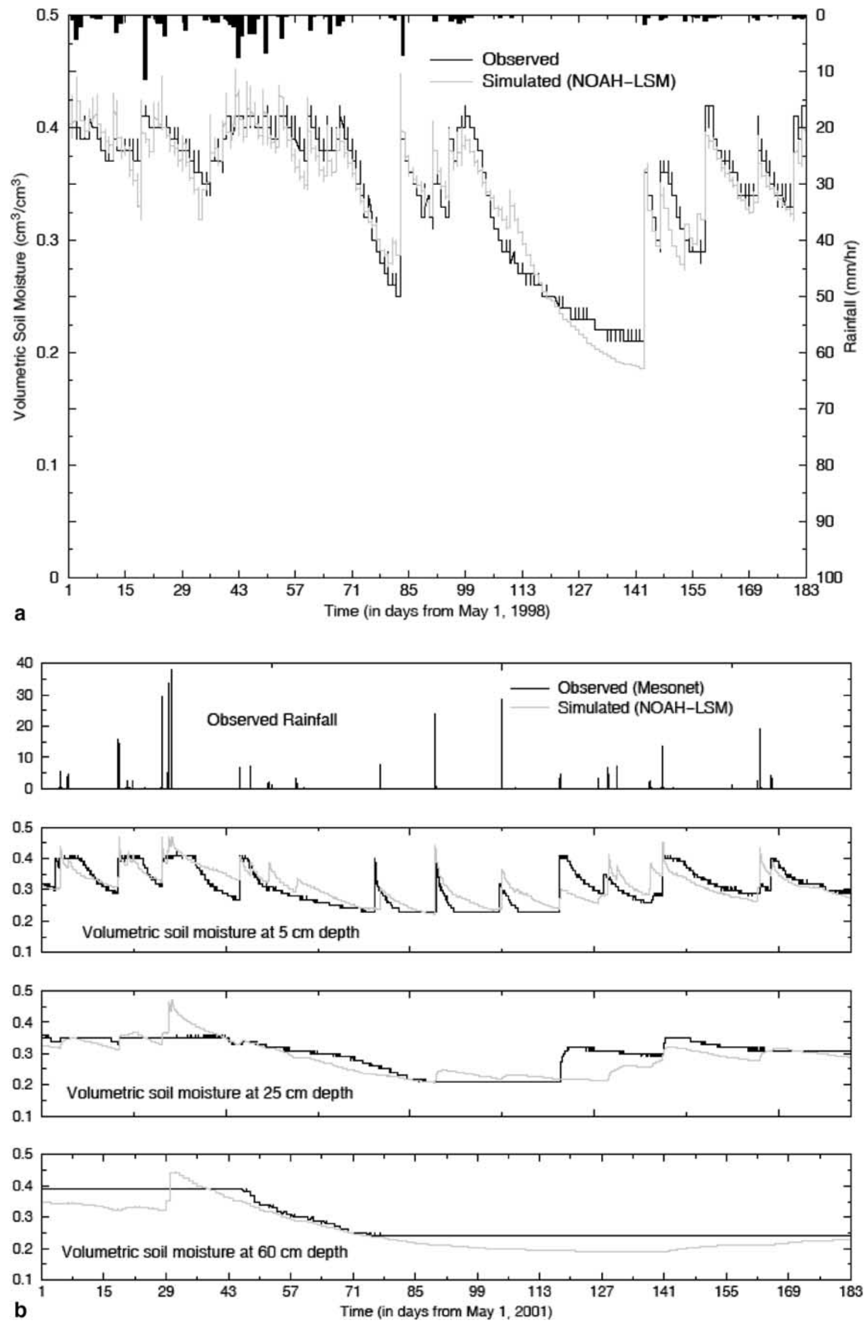


Fig. 1. (a) NOAH-LSM simulation of soil moisture (with adjustment for vegetation parameters) at 5 cm depth for CHAMPAIGN. Rainfall is shown on the opposite *x*-axis. (b) NOAH-LSM simulation of soil moisture (with adjustment for vegetation parameters) for PERKINS. Uppermost panel—observed rainfall from Mesonet; From bottom panel and up—soil moisture simulation at 60 cm, 25 cm and 5 cm depth, respectively.

Table 1
Uncertainty ranges and optimal values for soil hydraulic parameters of NOAH-LSM

Parameter	Minimum value	Maximum value	Optimal value		Sampling strategy
			CHAMPAIGN	PERKINS	
SMCMAX (m^3/m^3)	0.05	0.50	0.41	0.47	Uniform
PSISAT (m)	0.01	0.65	0.140	0.36	Uniform
SATDK (m/s)	1.00×10^{-6}	1.77×10^{-4}	3.39×10^{-6}	7.00×10^{-5}	Log (uniform)
BB	2.00	15.00	14.4	7.70	Uniform
SMCWLT (m^3/m^3)	0.01	0.20	0.100	0.119	Uniform

greater than an assigned minimum threshold value. For further details on GLUE implementation the reader is referred to Beven and Binley [4], Freer et al. [17] and Beven and Freer [5].

The GLUE method has a drawback that limits its application for computationally demanding models. It requires analysis of multiple simulation scenarios based on uniform random sampling of the model parameter hyperspace. This requirement can be prohibitive for models that are slow-running [3,4]. Hossain et al. [21] and Hossain and Anagnostou (2004b) provide an extensive review about this limitation, and propose an efficient sampling technique as an addendum to GLUE. In this technique, the uncertainty in soil moisture simulation (model output) is approximated through a Hermite polynomial chaos expansion of normal random variables that represent the model's parameter (model input) uncertainty. The unknown coefficients of the polynomial are calculated using limited number of model simulation runs. The calibrated polynomial is then used as a fast-running proxy to the slower-running LSM to predict the degree of representativeness of a randomly sampled model parameter set. The herein study has employed this efficient sampling scheme formulated by Hossain et al. [21] to substantially reduce the computational burden of the analyses. It should be noted that the Hermite polynomial scheme is used only to accelerate parameter sampling by avoiding unnecessary model runs due to non-behavioral parameter sets, and that selected parameter set's degree of representativeness is always verified on the basis of actual model runs.

4. Satellite rainfall error model

The one-dimensional (1-D) satellite rainfall error model (hereafter referred to as SREM-1D) developed by Hossain and Anagnostou [19] was used to characterize the satellite rainfall retrieval error. The approach is to stochastically simulate spatially independent (1-D), temporally correlated, realizations of satellite rainfall retrievals by corrupting a more accurate measurement of rainfall process. The more accurate source was derived from half-hourly rain gauge measurements (hereafter labeled as 'reference rainfall'). The three most pertinent aspects of the SREM-1D uncertainty framework, are: (1) conversion of reference rainfall rates to reference instantaneous rainfall rates; (2) modeling of the sensor's probability of detection for rain and no-rain events; and, (3) modeling of retrieval error based on a multiplicative error model with temporal correlation. For details on the algorithmic structure of SREM-1D the reader is referred to Hossain and Anagnostou [19].

The rain retrieval considered in this study is from satellite IR, as at global scale, these observations offer the finest temporal sampling characteristics (1/2-hourly) necessary to resolve the dynamic variability of soil moisture in the root zone. We considered here hourly averaged IR rainfall fields produced by NASA's Multi-satellite Precipitation Analysis (MPA) algorithm [24] as representative of the current level of IR rainfall estimation characteristics. This community release product is known as 3B41RT. Hossain and Anagnostou [19] had calibrated SREM-1D parameters for 3B41RT over the US on the basis of coincident rain profile estimates from TRMM Precipitation Radar [26]. Fig. 2a shows the cumulative hyetographs of actual IR (3B41RT) rainfall products and the corresponding Mesonet rainfall data over PERKINS for the year 2002 (1 January–30 October) when MPA became operational on a best effort basis. The 3B41RT rainfall is compared against the quantile envelop associated with 5–95% percentiles, predicted by SREM-1D using as input the Mesonet rain rates (Fig. 2a). The 3B41RT rainfall hyetographs, which is considered an observed realization, is enveloped by the SREM-1D quantiles. In Fig. 2b, we show a similar quantile envelop of SREM-1D simulations for CHAMPAIGN. There are currently no 3B41RT products available for the retrospective period of 1998 over CHAMPAIGN.

5. Simulation framework and results

Our study is essentially a sensitivity investigation addressing the 'relative' impact and non-linear interaction of uncertainties in modeling and satellite rainfall estimation. The words 'relative' and 'satellite' are stressed herein because this study does not focus on the model structure or rainfall estimation deficiencies per se. Rather the purpose of this study is to quantify the response of a given model structure (i.e., one that is used in the scientific community) to remotely sensed rainfall measurements by a space-borne passive sensor relative to the scenario of non-existence of uncertainty in rainfall (e.g., gauge measurements) and model parameters. We therefore argue that 1-year simulation period (with 6 months for comparison of sensitivities) is adequate to study these relative impacts. Furthermore, this study also does not address the spatial or lateral variability of soil moisture and the surface groundwater interaction. It is also highlighted that this study examines in a stand-alone fashion the sensitivity of soil moisture prediction (1D), given the near impossibility of completely defining the interdependency between all possible combinations of hydrologic and energy variables. We assume that the reference rainfall (from surface measurements) and the optimal parameters for NOAH-LSM yield accurate predictions of soil moisture with low uncertainty (see Fig. 1a and b). If it is further assumed that SREM-1D and GLUE sample adequately the error structure in satellite rainfall and NOAH-LSM

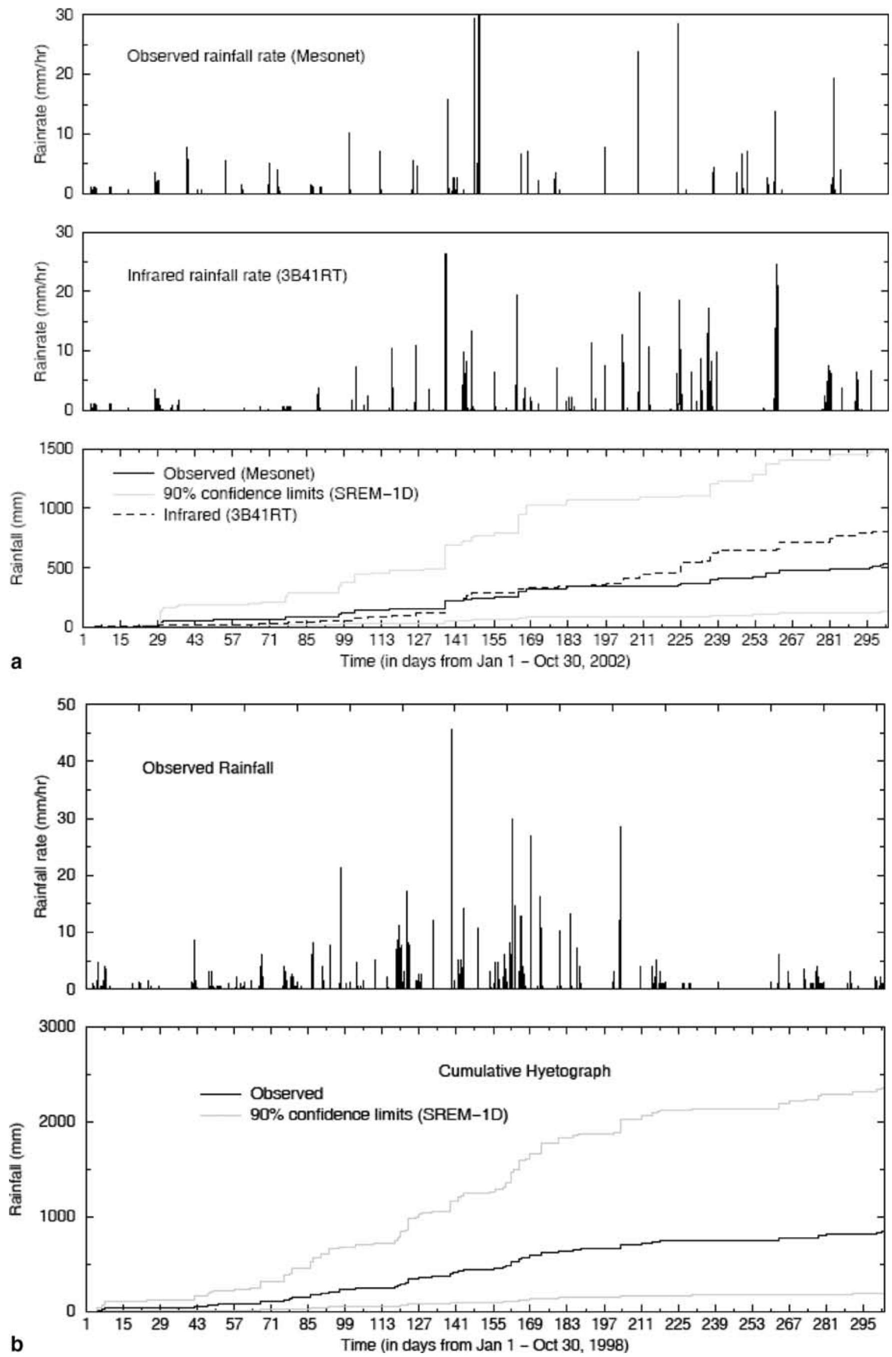


Fig. 2. (a) Rainfall estimation by infra-red technique (3B41RT) with simulated ranges of uncertainty over PERKINS for 2002. (b) Rainfall estimation by the assumed infra-red technique (SREM-1D) over CHAMPAIGN for 1998.

parameters, respectively, then, based on these two assumptions we can construct the following logical inferences: (1) Propagation of multiple realizations of SREM-1D rainfall processes via NOAH-LSM at optimal parameters will reflect the partial uncertainty in soil moisture prediction due to satellite rainfall estimation error (uppermost panel—Fig. 3a); (2) Propagation of reference rainfall to NOAH-LSM via multiple GLUE model parameter realizations will reflect the *partial* uncertainty in soil moisture prediction due to modeling uncertainty (middle panel—Fig. 3b); and (3) Combining SREM-1D and GLUE on NOAH-LSM will reflect the *total* uncertainty in soil moisture prediction due to *both* sources of uncertainty (lowermost panel—Fig. 3c).

5.1. Relative impact of uncertainties

As a demonstration of the relative impacts and interactions of uncertainties, multiple (500) realizations were conducted from SREM-1D and using GLUE. For

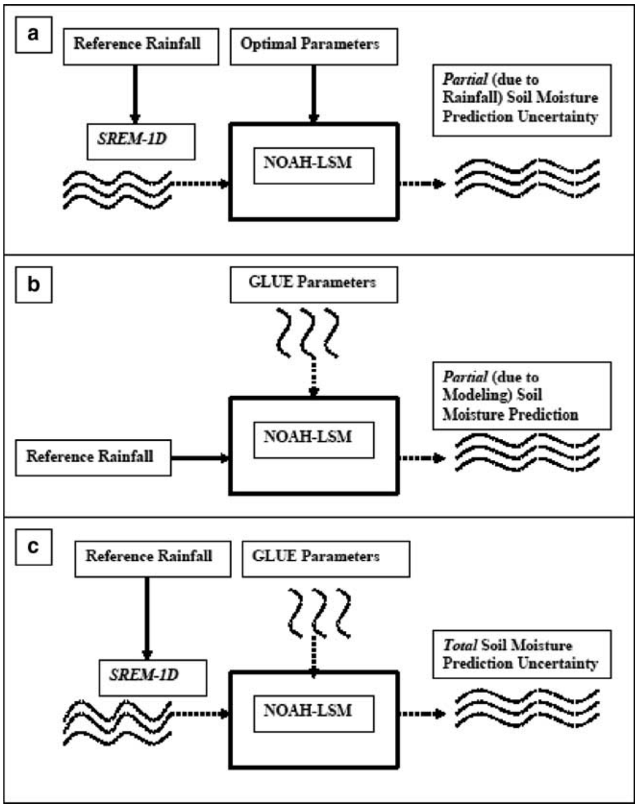


Fig. 3. Schematic representation of partial and total uncertainty in soil moisture simulation. Dotted wave-like lines represent uncertainty in the form of random realizations.

GLUE, the 500 best behavioral parameter sets were sampled from ranges shown in Table 1 with an E_{NS} greater than 0.4 (using gauge rainfall as input). For the combined uncertainty assessment, the full-blown MC uncertainty assessment comprising 250,000 (500 SREM-1D rainfall realizations times 500 GLUE parameter sets) NOAH-LSM runs was executed to identify the full range of predictive variability. In this study, the wideness of prediction quantiles in soil moisture simulation is considered a reliable measure of prediction uncertainty. This wideness, defined as uncertainty ratio (UR), is the time integrated uncertainty in soil moisture volume bounded by the quantile width (between upper and lower percentiles) normalized by the time-integrated observed soil moisture volume. The UR at n% quantile width (ranging from 10% to 90%), UR_n, is defined as follows:

$$\mathrm{UR}_n = \frac{\sum_{j=1}^{N_\mathrm{T}} (\mathrm{SM}_{j,50+n/2}^{\mathrm{sim}} - \mathrm{SM}_{j,50-n/2}^{\mathrm{sim}})}{\sum_{j=1}^{N_\mathrm{T}} \mathrm{SM}_j^{\mathrm{obs}}} \tag{2}$$

where, j is the time-step index of simulation, N_T the total number of time-steps in the simulation period. Superscripts sim and obs refer to simulated and observed soil moisture, respectively. UR represents the bulk variability in prediction expressed relatively to the magnitude of the observed variable.

In Figs. 4 and 5a–c we show each of the three inferences for CHAMPAIGN and PERKINS, respectively. The uncertainty limits of simulation are shown at the 90% quantile width. It appears that there is no significant dependency of uncertainty as a function of depth in the case of PERKINS (Fig. 5). It should be noted that the partial uncertainty due to modelling and the combined (total) uncertainty are conditioned upon the subjective threshold used to select the behavioral param-

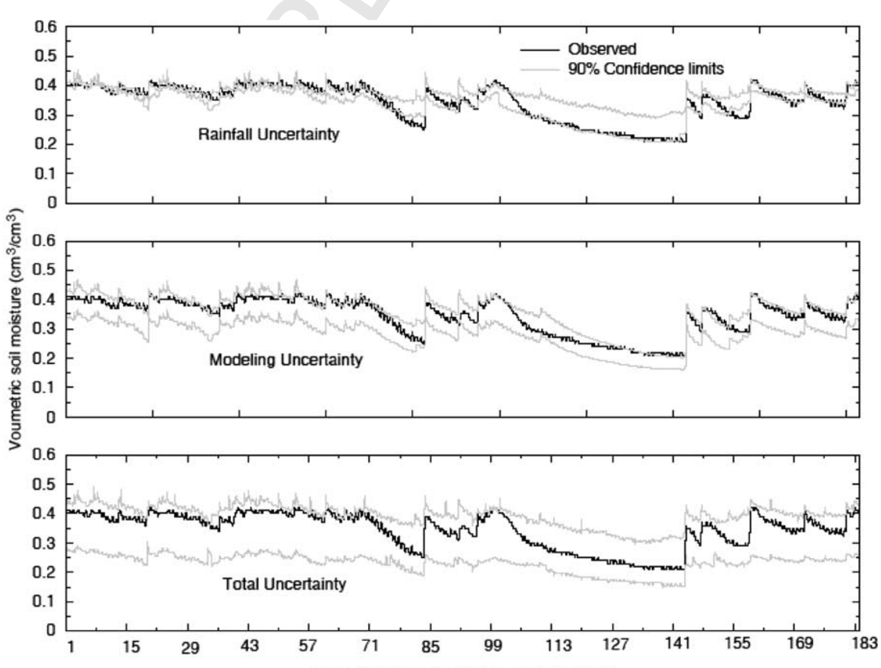


Fig. 4. Comparisons of soil moisture prediction uncertainty (partial and total) for CHAMPAIGN.

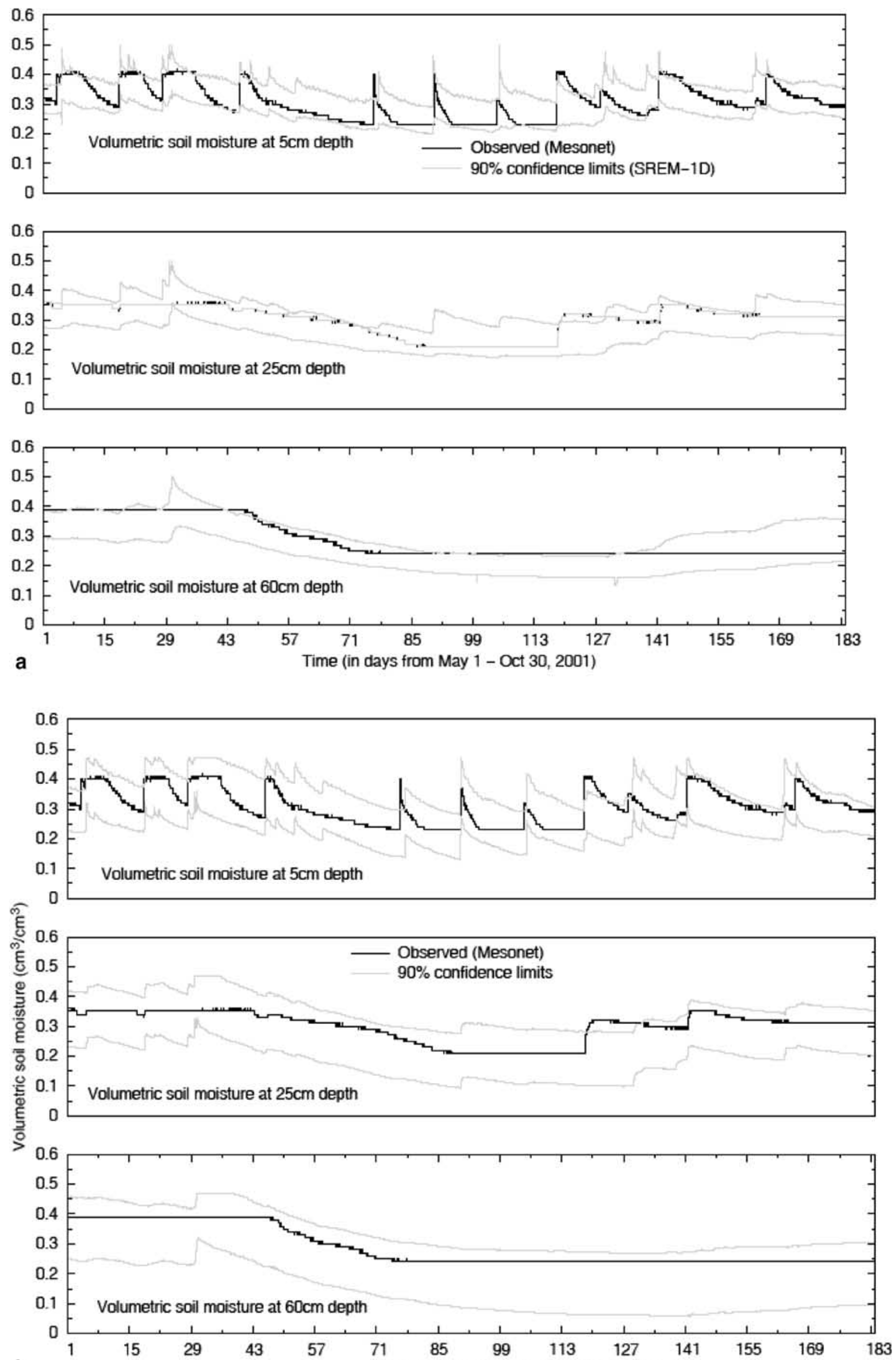
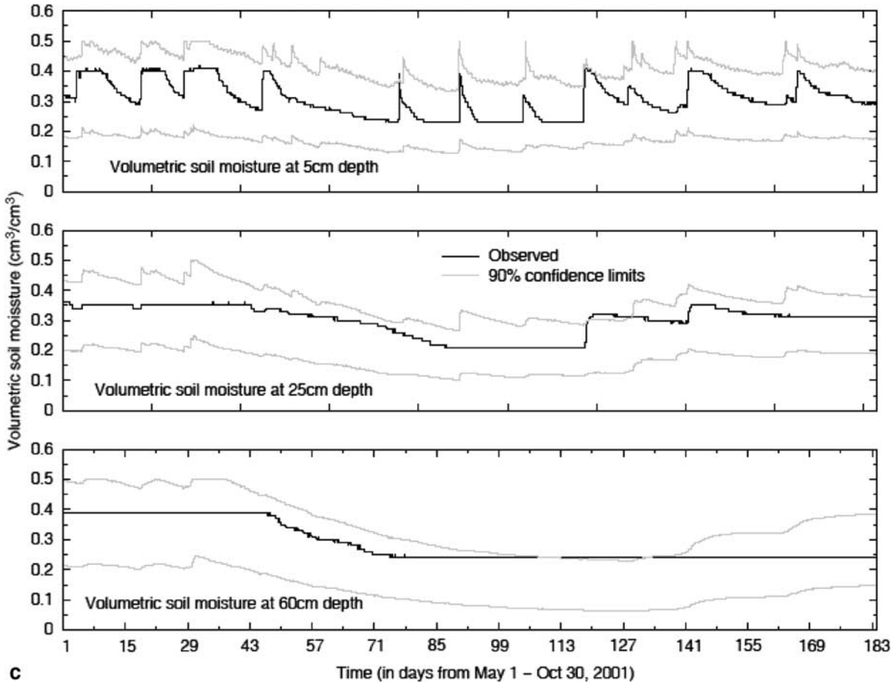


Fig. 5. (a) Partial uncertainty in soil moisture prediction due to precipitation uncertainty for PERKINS. (b) Partial uncertainty in soil moisture simulation due to modeling uncertainty for PERKINS. (c) Total uncertainty in soil moisture prediction due to uncertainties in precipitation measurement and modeling for PERKINS.

eter sets (which was fixed at $E_{NS} > 0.4$). The partial uncertainties due to rainfall estimation and modelling are considerably higher in PERKINS (compare the uppermost and middle panels of Figs. 4 and 5). Due to the numerical nature of our investigation, we can only speculate that the vegetation and hydraulic properties may be one of the many potential catalysts for the increased error interaction. We support our speculation numerically with a stand-alone sensitivity study described next.

The increasing sensitivity to precipitation error as parameter values transformed from CHAMPAIGN to PERKINS vegetation/soil type environment is shown in Fig. 6a. The simulation experiment that resulted to



c

Fig. 5 (*continued*)

this figure used meso-forcing meteorological data of CHAMPAIGN, while the soil-hydraulic parameters were varied from the optimal values of CHAMPAIGN to the optimal value of PERKINS (shown in Table 1). In the figure, the varied parameter values are shown re-scaled between 0 and 1, where the lower and upper ends represent the environment for CHAMPAIGN and PERKINS, respectively (here, we use the term 'regime scale'). The vertical axis of this figure shows the relative increase (in %) of the UR evaluated in terms of the 90th quantile width (UR_{90}, Eq. (2)). Similarly, Fig. 6b shows the relative level of error propagation from rainfall input to soil moisture as function of quantile width assuming optimal model performance (run at optimal parameter sets shown in Table 1). The UR_n where n is varied from 10% to 90% is used here to characterize the level of uncertainty in precipitation and the predicted soil moisture. What is evident is that soil moisture uncertainty is significantly dampened in the rainfall–soil moisture transformation process in a highly non-linear fashion with porosity controlling the upper bound of variability. It is shown that the satellite IR hourly rain input uncertainty increases exponentially with quantile width to over than twice the magnitude of the estimated rainfall (which is commonly expected for IR retrievals at high resolution). The corresponding IR rain estimation error propagation to soil moisture prediction is, though, associated with a significant non-linear dampening: i.e., the UR converges to values well below 0.4 (i.e., >85% error reduction). This dampening is notably more significant for CHAMPAIGN (90%) than PERKINS (85%) where the error propagation is enhanced due to the vegetated environment.

5.2. The impact of model uncertainty

Next we study the relative significance of the two uncertainty sources (precipitation versus modeling), which warrants a more detailed characterization of the role of the behavioral threshold for parameter sets used in GLUE. Since this threshold is essentially subjective, it is important to recognize that its value may increase or decrease (from $E_{NS} = 0.4$) to represent various levels of parametric uncertainty (or model accuracy) at the operational scenario. Consequently, we grouped the behavioral parameter sets (all having $E_{NS} > 0.4$) into three model performance categories—(1) HIGH (high modeling accuracy: $E_{NS} \geqslant 0.75$); (2) MEDIUM (moderate modeling accuracy: $0.5 \leqslant E_{NS} < 0.75$), and (3) LOW (low modeling accuracy: $0.4 < E_{NS} < 0.5$). In each category group, the norm distance of its parameter sets (θ^j) from the optimal parameter set (shown in Table 1) was determined as follows:

$$\mathrm{Dis}_{\theta_j} = \sum_{i=1}^{5} (\theta_i^{\mathrm{opt}} - \theta_i^j)^2 \tag{3}$$

where, i is the parameter set index. Fig. 7a shows the dispersion of behavioral parameter sets from the optimal set versus model performance, while Fig. 7b shows the corresponding cumulative density functions of Dis_θ for each model performance category. From each of the

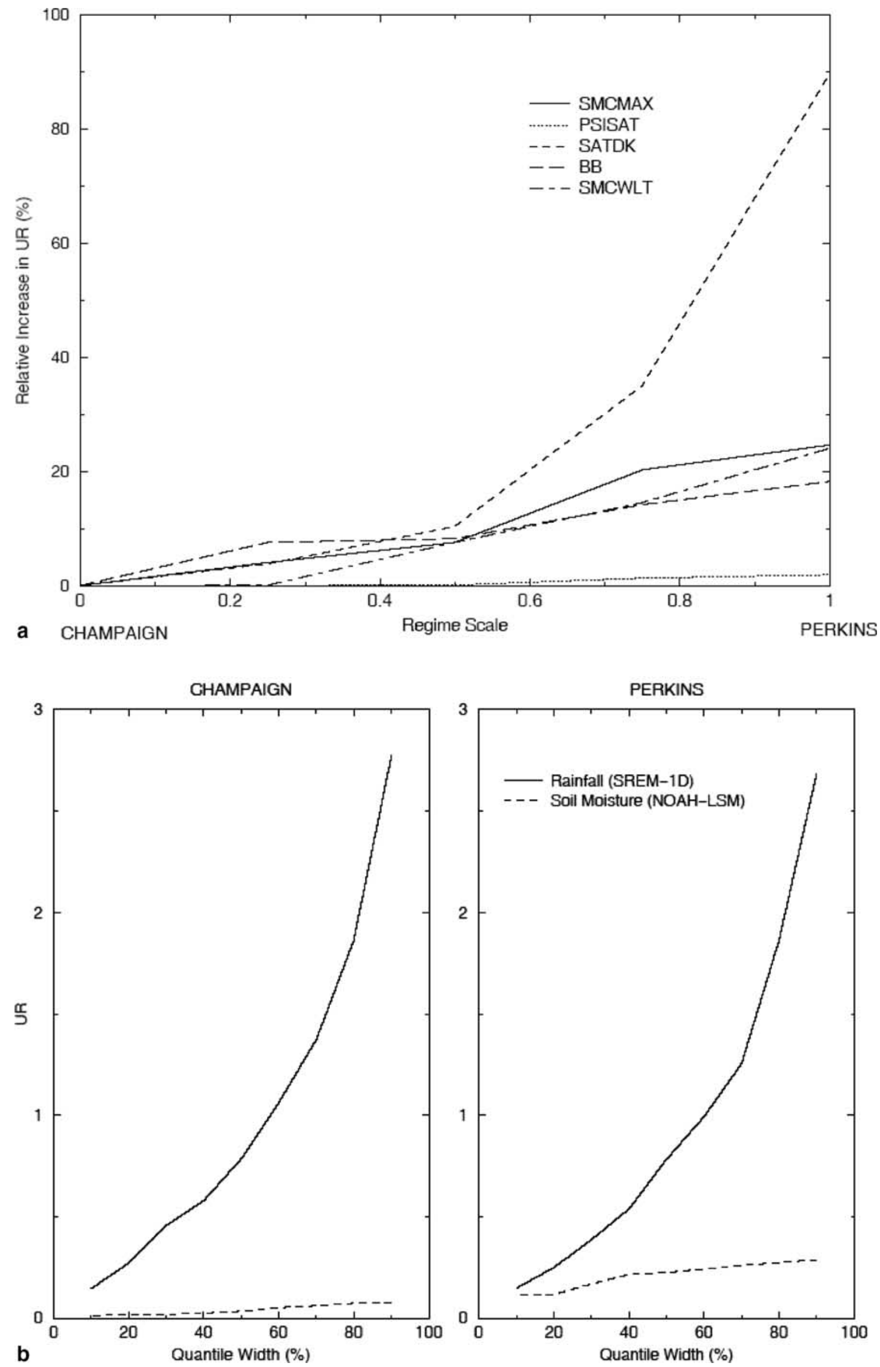


Fig. 6. (a) The sensitivity (stand-alone) of soil moisture prediction uncertainty to precipitation uncertainty as the parameter values change from a CHAMPAIGN-like environment (left-hand side) to a more PERKINS-like environment (right-hand side). The *y*-axis represents the relative increase of UR in soil moisture prediction (5 cm depth) at the 90% quantile width. The meso-forcing data pertained to CHAMPAIGN. (b) Error propagation from rainfall to soil moisture for two regimes when model performs at optimal level. The soil moisture simulations are at the 5 cm depth.

three categories, we sampled a set of 100 parameter sets to evaluate modeling uncertainty. The procedure was as follows. To start with, 100 uniformly distributed, $U[0, 1]$, random numbers were generated. Each random number represented a cumulative density value for Dis_θ. Projecting this value through the CDF function shown in Fig. 7b we evaluated the corresponding Dis_θ value (quantile). The selected Dis_θ values were then used to get the representative parameter sets, θ, of the group. On the basis of the 100 selected GLUE parameters we performed partial-modeling uncertainty evaluation. Combining the 100 GLUE parameters with 500 SREM-1D random ensembles (total: 50,000 LSM realizations) we evaluated the combined precipitation-modeling uncertainty. In Fig. 8 we show the UR values at the 5 cm depth for each site as a function of model performance category (high,

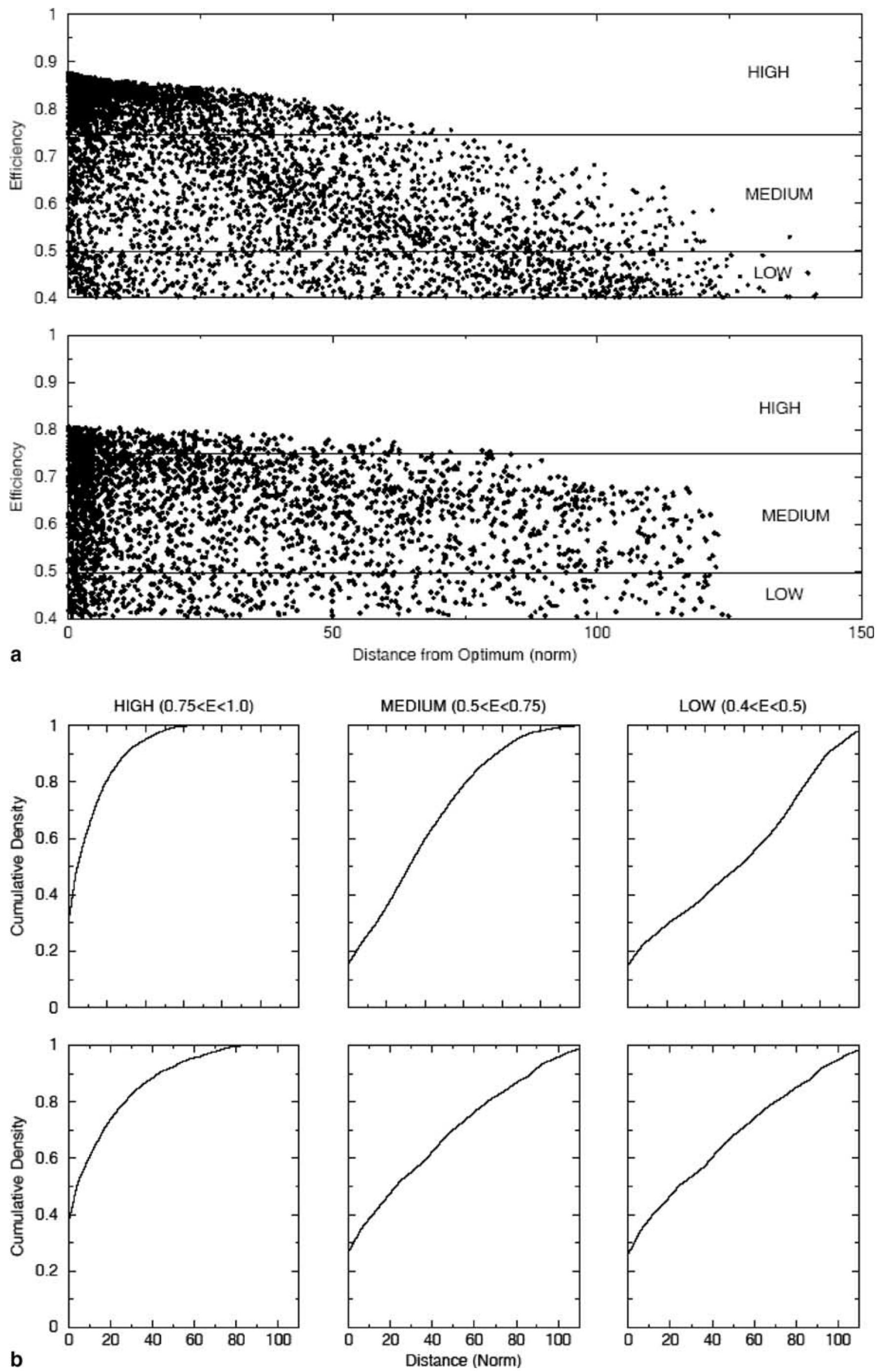


Fig. 7. (a) The ranking of model parameter sets based on accuracy levels as—HIGH ($0.75 < E_{NS} < 1.0$), MEDIUM ($0.5 < E_{NS} < 0.75$) and LOW ($0.4 < E_{NS} < 0.5$). Upper panel—CHAMPAIGN; Lower panel—PERKINS. *X*-axis represents the distance of each parameter set from the optimal value based on Eq. (3). (b) The cumulative density function of the distance of model parameter sets (from optimal set) ranked according the model performance categories (high, medium and low). Upper panel—CHAMPAIGN; Lower panel—PERKINS.

medium, low) and quantile width. The partial uncertainty due to precipitation (at optimal model performance) is also shown in the form of long-dashed line in each plot for comparison of the dependencies. The following are the most notable observations from this figure: (1) UR values at 90% quantile width (total and modeling-partial uncertainty) for PERKINS are in the range of 50–100% higher than those for CHAMPAIGN; (2) The interaction of modeling uncertainty with precipitation uncertainty increases as a function of modeling uncertainty—furthermore, this interaction is greater for PERKINS; (3) The partial uncertainty in soil moisture prediction arising due to precipitation uncertainty only (considering optimal model performance) signifi-

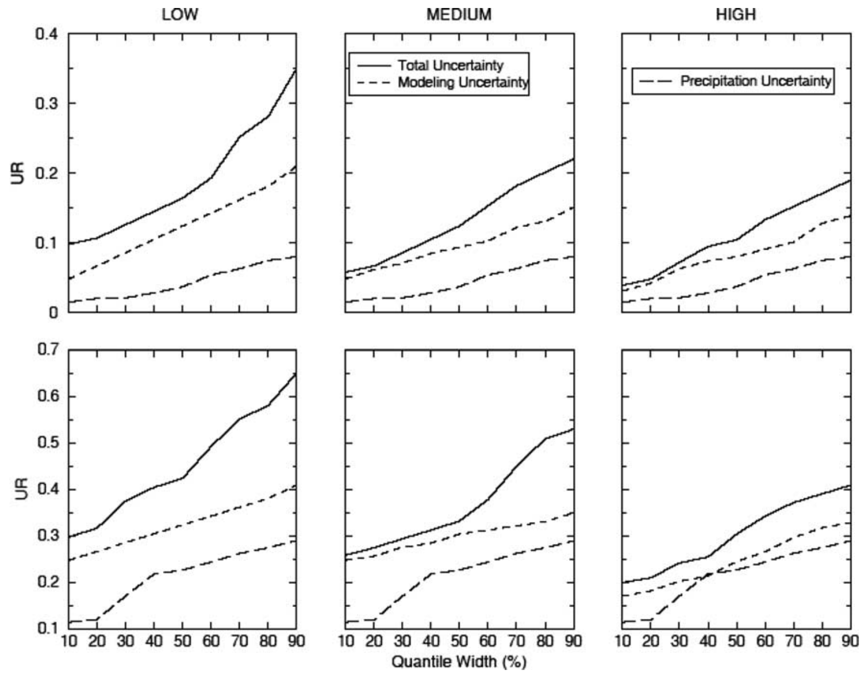


Fig. 8. Total and partial soil moisture prediction uncertainty at 5 cm depth as a function of model performance levels (behavioral thresholds). Upper panel—CHAMPAIGN; Lower panel—PERKINS.

cantly under-represents the overall uncertainty—this underestimation reduces as the model performance levels improves (from LOW to HIGH).

The global picture emerging from this analysis is that proper characterization of error propagation in hydrologic prediction in soil moisture would require the study of non-linear error interaction between modeling error and error in forcing variables (precipitation, and other meteorological radiation parameters). While this may be a recognized issue in current literature, our consideration of satellites as the primary rainfall source and its comprehensive error modeling represents, what we believe, a new agenda in anticipation of future hydrologic missions (GPM and HYDROS). As shown above, the precipitation uncertainty associated with satellite IR retrievals would explain about half of the total uncertainty in soil moisture prediction for a high model accuracy scenario, while less than 30% in the case of low modeling accuracy. This indicates that via an understanding of the retrieval-modeling error interaction in hydrologic prediction, we should attempt investigating the optimality criteria for integrating satellite rain retrievals in land data assimilation systems.

6. Conclusions

This study focused on the sensitivity of soil moisture prediction accuracy to the interaction of two types of error sources considered relevant for emerging assimilation systems: the precipitation input from satellites and land surface model parametric uncertainties. The moisture prediction was limited to 1-D vertical simulation neglecting horizontal advection and spatial heterogeneity, which should therefore be considered as an inherent limitation of our study. The modeling uncertainty was represented by GLUE technique that characterized the non-uniqueness of model parameters yielding similar model performance assessment. A satellite rainfall error model (SREM-1D) was devised to characterize uncertainty in satellite rain retrieval. Satellite rainfall estimates pertained to hourly averaged satellite infra-red (IR) estimates. The combined assessment of uncertainty—namely, rainfall input and modeling (SREM-1D + GLUE)—was compared with the partial assessment that accounted for modeling (GLUE) or IR rain retrieval uncertainty (SREM-1D). Comparisons were also made on two distinct sites: (1) a site with sparse farmland vegetation (in Champaign, Illinois); and (2) a site with denser vegetation (in Perkins, Oklahoma). Soil moisture prediction uncertainty was found to be about 50–100% larger for the more vegetated site. Current IR rain retrievals are shown to contribute between 20% and 60% of the total uncertainty in soil moisture prediction. The lower (upper) limit corresponds to high (low) modeling accuracies. The study indicates that a rigorous assessment of satellite rain retrievals in terms of hydrologic predictions requires an understanding of the role played by modeling uncertainty in error interaction.

While the above findings represent a useful first step, it is not until a number of similar studies from a range of research objectives are undertaken to achieve a firm understanding of the optimality criteria for integration of remotely sensed data in LSMs. Towards that end, we highlight the following as natural extensions to address limitations of our current work. To improve the hydrologic application of satellite rain estimation, merging of IR rainfall fields with the less-frequent, but more definitive, passive microwave (PM) rainfall estimates should be explored. A recent study on the basis of an experimental error assessment framework by Anagnostou [2] has shown that optimal merging of IR and PM rainfall fields reduces hydrological prediction error statistics (both marginal and conditional). Other studies related to runoff prediction have shown that PM-IR merging can reduce uncertainty of certain runoff parameters (e.g., runoff volume for water balance studies) [19]. This error propagation framework needs to be augmented incorporating other LSM schemes, such as those currently used in LDAS. Another aspect worth addressing as a future extension is the spatial structure of error (2-D simulations). More useful analyses for uncertainty and data assimilation techniques can be expected when the spatial structure of satellite retrieval error and soil moisture simulation are considered involving some down-scaling approaches to address scale mismatch between rainfall observations and model predictions. To address long memory effects of soil moisture longer time series need to be studied that is commensurate with current computational resources. We are currently working on expanding SREM-1D to simulate the spatial variability of satellite rainfall fields' estimation error (Hossain and Anagnostou, 2005) and we hope to report findings on the hydrologic implications in the near future.

7. Uncited references

[18,22].

Acknowledgments

This study was supported by NASA's Global Water and Energy Cycle program (NAG5-11527). The first author was supported by a NASA Earth System Science Fellowship.

References

[1] Albertson JD, Montaldo N. Temporal dynamics of soil moisture variability: 1 Theoretical basis. Water Resour Res 2003;39(10): 1274. doi:10.1029/2002WR001616.
[2] Anagnostou EN. Assessment of satellite rain retrieval error propagation in the prediction of land surface hydrologic variables. In: Levizzani V, Bauer P, Turk FJ, editors. Book Chapter in Measuring Precipitation from Space: EURAINSAT and the Future. Kluwer Academic Publishers, in press.
[3] Bates BC, Campbell EP. A Markov Chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. Water Resour Res 2001;37(4):937–47.
[4] Beven KJ, Binley A. The future of distributed models: Model calibration and uncertainty prediction. Hydrol Process 1992;6: 279–298.
[5] Beven KJ, Freer J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. J Hydrol 2001;249:11–29.
[6] Bidwell S, Turk J, Flaming M, Mendelsohn C, Everett D, Adams J, et al. Calibration plans for the Global Precipitation Measurement, Paper presented at Joint 2nd Int'l Microwave Radiometer Calibration Workshop and CEOS working group on Calibration and Validation, Barcelona, Spain, 9–11 October 2002.
[7] Bras RL. Hydrology: an introduction to hydrologic science. US: Addison-Wesley; 1990.
[8] Brock F, Crawford KC, Elliott GW, Cuperus SJ, Stadler SJ, Johnson HL, et al. The Oklahoma Mesonet: A technical overview. J Atmos Oceanic Technol 1995;21(1):5–19.
[9] Butts MB, Payne JT, Kristensen M, Madsen H. An evaluation of the impact of model structure on hydrological modeling uncertainty for streamflow simulation. J Hydrol 2004;298:242–66.
[10] Carpenter TM, Georgakakos KP. Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model. J Hydrol 2004;298:202–21.
[11] Chen F, Mitchell K, Schaake JC, Xue Y, Pan H-L, Koren V, et al. Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. J Geophys Res 1996; 101:7251–68.
[12] Clapp RB, Hornberger GM. Empirical equations for some soil hydraulic properties. Water Resour Res 1978;14:601–4.
[13] Cosby BJ, Hornberger GM, Clapp RB, Ginn TR. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resour Res 1984;20(6): 682–690.
[14] Cosgrove BA. Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS). J Geophys Res 2003;108(D22):8845. doi:10.1029/2002JD003316.
[15] Elliot RL, Brock FV, Stone ML, Harp SL. Configuration decisions for automated weather station network. Appl Eng Agri 1994;10(1):45–51.
[16] Entekhabi D, Rodriguez-Iturbe I. Analytical framework for the characterization of the space-time variability of soil moisture. Adv Water Resour 1994;17:35–45.
[17] Freer J, Beven KJ, Ambroise B. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. Water Resour Res 1996;32(7):2161–73.
[18] Georgakakos KP, Seo D-J, Gupta H, Schaake JC, Butts M. Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. J Hydrol 2004;298:223–41.
[19] Hossain F, Anagnostou EN. Assessment of current passive-microwave and infra-red based satellite rainfall remote sensing for flood prediction. J Geophys Res 2004;109(D07012). doi:10.1029/2003JD003986.
[20] Hossain F, Anagnostou EN, Borga M, Dinku T. Hydrological model sensitivity to parameter and radar rainfall estimation uncertainty. Hydrol Process 2004;18(17):3277–99. doi:10.1002/hyp.5659.
[21] Hossain F, Anagnostou EN, Lee K-H. A stochastic response surface method for Bayesian estimation of uncertainty in soil moisture simulation from a land surface model. Special Issue in Non-Linear Processes in Geophysics, 2004. p. 11.

[22] Hossain F, Anagnostou EN. A two-dimensional satellite rainfall error model. IEEE Trans Geosci Remote Sensing, in press.
[23] 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.
[24] Huffman GJ, Adler RF, Stocker EF, Bolvin DT, Nelkin EJ. Analysis of TRMM 3-hourly multi-satellite precipitation estimates computed in both real and post-real time. Paper presented at 12th Conference on Satellite Meteorology and Oceanography, Long Beach, CA, 9–13 February 2003.
[25] Idso SB. A set of equations for full spectrum of 8- to 14-μm and 10.5 to 12.5-μm, thermal radiation from cloudless skies. Water Resour Res 1981;17(2):295–304.
[26] Kummerow C, Hong Y, Olson WS, Yang S, Adler RF, McCollum J, et al. The status of the tropical rainfall measuring mission (TRMM) after two years in orbit. J Appl Meteorol 2000;39(12):1965–82.
[27] Luo L, Robock A, Mitchell KE, Houser PR, Wood EF, Schaake JC, et al. Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains. J Geophys Res 2003;108(D22):8843. doi:10.1029/2002JD003246.
[28] Margulis SA, McLaughin D, Entekhabi D, Dunne S. Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. Water Resour Res 2002;38(12).
[29] McLaughin D. An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering. Adv Water Resour 2002;25:1275–86.
[30] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models, 1. A discussion of principles. J Hydrol 1970;10:282–90.
[31] Pan H-L, Mahrt L. Interaction between soil hydrology and boundary layer development. Bound Layer Meteorol 1987;38:185–202.
[32] Pan F, Peters-Lidard CD, Sale MJ. An analytical method for predicting surface soil moisture from rainfall observations. Water Resour Res 2003;39(11):1314. doi:10.1029/2003WR002142.
[33] Robock A, Luo L, Wood EF, Wen F, Mitchell KE, Mitchell, et al. Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season. J Geophys Res 2003;28(D22):8846. doi:10.1029/2002JD003245.
[34] Roddell M, Houser PR, Jambor U, Gottschalck J, Mitchell K, Meng C-J, et al. The Global Land Data Assimilation System. Bull Amer Meteorol Soc 2004;85(3):381–94.
[35] Schaake JC, et al. An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS). J Geophys Res 2004;109(D01S90). doi:10.1029/2002JD003309.
[36] Smith EA. Satellites, orbits and coverages, Paper presented at IGARSS 2001 (International Geoscience and Remote Sensing Symposium), Sydney, Australia, 9–13 July 2001.
[37] Smith MB, Georgakakos KP, Xu L. The distributed model intercomparison project (DMIP). J Hydrol 2004;298:1–3.
[38] Sridhar V, Elliot RL, Chen F, Botzge JA. Validation of the NOAH-OSU land surface model using surface flux measurements in Oklahoma. J Geophys Res 2002;107(D20).
[39] Walker JP, Houser PR. Requirements of a global near-surface soil moisture satellite mission: accuracy, repeat time, and spatial resolution. Adv Water Resour 2004;27:785–801.