Advances in Water Resources 34 (2011) 433–447

Contents lists available at ScienceDirect

Advances in Water Resources

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

Particle Filter-based assimilation algorithms for improved estimation of root-zone soil moisture under dynamic vegetation conditions

Karthik Nagarajan [a,*], Jasmeet Judge [a], Wendy D. Graham [b], Alejandro Monsivais-Huertero [c]

[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
[c] ESIME Unidad Ticoman, Insitituto Polictenico Nacional, Mexico City, Mexico

ARTICLE INFO

Article history:
Received 12 April 2010
Received in revised form 24 September 2010
Accepted 30 September 2010
Available online 4 November 2010

Keywords:
Root zone soil moisture
SVAT-vegetation models
Particle Filter
EnKF
MicroWEX-2

ABSTRACT

In this study, we implement Particle Filter (PF)-based assimilation algorithms to improve root-zone soil moisture (RZSM) estimates from a coupled SVAT-vegetation model during a growing season of sweet corn in North Central Florida. The results from four different PF algorithms were compared with those from the Ensemble Kalman Filter (EnKF) when near-surface soil moisture was assimilated every 3 days using both synthetic and field observations. In the synthetic case, the PF algorithm with the best performance used residual resampling of the states and obtained resampled parameters from a uniform distribution and provided reductions of 76% in root mean square error (RMSE) over the openloop estimates. The EnKF provided the RZSM and parameter estimates that were closer to the truth than the PF with an 84% reduction in RMSE. When field observations were assimilated, the PF algorithm that maintained maximum parameter diversity offered the largest reduction of 16% in root mean square difference (RMSD) over the openloop estimates. Minimal differences were observed in the overall performance of the EnKF and PF using field observations since errors in model physics affected both the filters in a similar manner, with maximum reductions in RMSD compared to the openloop during the mid and reproductive stages.

1. Introduction

Root-zone soil moisture (RZSM) is an important hydrologic state variable that governs moisture and energy fluxes at the land–atmospheric interface. Accurate knowledge of RZSM is essential for near-term climate predictions, hydrologic and agricultural research [24], and for effective water resources management. Typically, Soil Vegetation Atmosphere Transfer (SVAT)-vegetation growth models simulate energy and moisture transport in soil and vegetation, and estimate these fluxes at the land surface and in the root-zone over a growing season [7,40]. However, these coupled models exhibit large uncertainties in RZSM estimates due to errors in conceptualization, computation, initialization, forcings, and model parameters. Such uncertainties can be significantly reduced by assimilating in situ and/or remotely sensed observations of soil moisture [41,39,46,23]. Ensemble-based techniques such as the Ensemble Kalman Filter (EnKF) [17] and the Particle Filter (PF) [20] have been used for data assimilation in hydrology.

Both the EnKF and the PF techniques involve representing states and parameter distributions as ensembles or set of particles and are particularly suitable for estimating hydrologic states whose evolution over time are described by non-linear models. However, the EnKF uses innovations to update the prior states and parameters, while the PF uses the innovations to assign posterior weights to resample prior states and parameters. The EnKF also uses only second-order statistics for updating the states and parameters, while the PF utilizes the entire probability density function (PDF) of the states given the observations in computing the posterior weights, employing a Bayesian approach. Thus, the PF algorithms are expected to provide estimates closer to the observations than the EnKF when non-Gaussian distributions are involved or non-linear relationships exist between the estimated state and observed data. The PFs also require larger number of particles than the EnKF-based algorithms because of the need to estimate the entire PDF and are computationally more intensive. However, recent advances in high-performance computing allow further investigations into the implementation of different PF-based algorithms.

Significant research has been conducted on EnKF-based assimilation for RZSM estimation [42,46,12,37]. However, only a few studies have been performed using PF [38,45]. Also, only few studies have been conducted that compare the two techniques [54,52,23] for hydrology applications. Weerts and de Vries [52] compared the performance of the two techniques for rainfall-runoff estimation using synthetic and field observations and found that the EnKF provided discharge estimates closer to the truth than

* Corresponding author.
E-mail address: nagkart@ufl.edu (K. Nagarajan).

0309-1708/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2010.09.019

the PF when field observations were assimilated. For studies in RZSM, Zhou et al. [54] found the EnKF to perform better than the PF, while Han and Li [23] found that both the EnKF and PF offered similar performances. Both the studies assimilated empirically-derived near-surface soil moisture from remotely sensed microwave observations, where the physical model is assumed to capture the biophysics perfectly, typical for synthetic studies. However, for field studies, additional errors may be introduced due to imperfect model physics and if the error distributions do not represent the actual errors (typically not known) in the model physics, the weights computed in the PF generate estimates that are further from the truth. Since the PF is more sensitive to model errors than the EnKF [52], studies comparing the performance of the EnKF and the PF using both synthetic and field observations are necessary to understand the impact of uncertainties in model biophysics, particularly under dynamic vegetation conditions.

The PF-based assimilations typically include resampling algorithms to regenerate values of states and parameters from the posterior distributions. Resampling is desired because PFs implemented without it, such as the Sequential Importance Sampling (SIS), can potentially lead to severe depletion of samples resulting in sample impoverishment [16]. Studies that used PF for RZSM estimation have incorporated the Sampling Importance Resampling (SIR) [38,23] algorithm. However, Liu and Chen [33] proposed residual resampling (RR) as an improvement to the SIR algorithm as it is computationally more efficient and provides lower variance of the posterior estimates. Weerts and El Serafy [52] compared the SIR and RR algorithms in rainfall-runoff modeling and found that the RR algorithm offered lower RMSEs. To date, the RR algorithm has not been applied in RZSM studies.

As mentioned earlier, for synthetic studies, the model error is assumed to be zero. However, when field observations are assimilated, additional bias may be introduced, particularly under dynamic land-surface conditions. Simultaneous estimation of parameters with states [9] allows for addressing bias in the model [14]. Applying the SIR or the RR algorithm without parameter perturbation may lead to repeated resampling of a few highly likely parameters resulting in sample impoverishment and inaccurate RZSM predictions after several updates. Hence, the resampling algorithms in PFs should be augmented by techniques that allow parameter values to maintain parameter diversity to obtain accurate estimates of RZSM in such cases. Moradkhani et al. [39] used a random walk model to perturb parameter estimates and maintain diversity. Because accurate estimation of model parameters play a significant role in accurate estimation of RZSM, an analyses of the impact of different parameter resampling algorithms on RZSM estimates is needed.

According to the hypothesis of equifinality [2], estimating states and parameters from observations related to RZSM imposes an ill-posed problem, wherein multiple parameter combinations may offer similar RZSM estimates. Both the EnKF and the PF (with resampling) narrow the PDF of parameters during assimilation. In this study, a comparison using parameter values estimated by different resampling methods within PF, including an algorithm with no resampling, is performed to investigate equifinality in RZSM estimation. The PF algorithm implemented without resampling would resemble the Generalized Likelihood Uncertainty Estimation (GLUE) method which has been used to study equifinality [2].

The goal of this study is to understand the impact of different parameter resampling algorithms in PF on the RZSM estimates while simultaneously updating states and parameters under dynamic vegetation conditions. A coupled SVAT-vegetation model is used to estimate RZSM during a growing season of sweet corn in North Central Florida. The estimates from the PF using four different algorithms are compared with those from the EnKF when synthetic and field-based observations of near-surface (0–5 cm) soil moisture are assimilated every 3 days. The four algorithms include, a PF without resampling or memory (PF-NRNM), a PF implemented using the RR algorithm (PF-RR), and PFs implemented using the RR algorithm with Gaussian (RR2) and uniform (RR3) parameter perturbations. Convergence of the RZSM estimates are investigated using two PF algorithms, one that includes resampling and the other that does not, to determine the number of particles needed in the PF for our application. Parameter convergence and root mean square errors (RMSE) in RZSM estimates are analyzed over the whole season and also during different growth stages to understand seasonal differences in algorithm performance.

In the next section, we briefly describe the field observations, the coupled SVAT-Vegetation model, and the PF and the EnKF algorithms used in this study.

2. Experiment, model, and assimilation

2.1. MicroWEX-2

The Second Microwave, Water and Energy Balance Experiment (MicroWEX-2) was conducted in North Central Florida (29.41 N, 82.18 W) during a growing season of sweet corn from Day of Year (DoY) 78 (March 18) to DoY 154 (June 2) in 2004 [28]. The experimental site was 3.6 h (9 acre). The soils at the site were primarily sand (89%) and the crop was heavily irrigated. During the MicroWEX-2, observations were conducted for soil moisture and temperature at depths of 2, 4, 8, 32, 64, and 100 cm, every 15 min along with observations of up and downwelling solar and longwave radiation, air temperature, relative humidity, precipitation, and irrigation. Judge et al. [28] provides details of observations conducted during MicroWEX-2 including microwave, soil, and vegetation observations. Table 1 shows the different growth stages of corn and their associated vegetation characteristics.

2.2. LSP-DSSAT model

The SVAT model used in this study is the Land Surface process (LSP) model [30]. 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 longwave radiation, irrigation/precipitation, and windspeed. The model has been rigorously tested [27] and extended to wheat-stubble [29] and brome-grass [31] in the Great Plains, prairie wetlands in Florida [53], to tundra in the Arctic [10], and to growing crop [7]. The LSP model includes 16 parameters as shown in Table 2. The vegetation energy balance is calculated using the model developed by Verseghy et al. [51] for the water drainage from canopy, the bulk transfer approach for the sensible heat flux from Trenberth [49], and the latent heat flux following [43]. The coupled energy and water balance in soil is calculated from Philip and de Vries [44,15]. A block-centered finite-difference scheme is employed to solve the coupled governing equations 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 simulates crop growth and development at a daily step using modules for soil, soil–plant–atmosphere, weather, management, including irrigation and fertilization [26]. The DSSAT model includes modules for over 25 types of crops, including corn, soybeans, wheat, cotton, and different grass types for pasture. The model has been extensively tested in different