Advances in Water Resources 34 (2011) 526–536

Contents lists available at ScienceDirect

Advances in Water Resources

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

Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture

Fan Chen [a], Wade T. Crow [b,*], Patrick J. Starks [c], Daniel N. Moriasi [c]

[a] Science Systems and Applications, Inc./USDA ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD 20705, United States
[b] USDA ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD 20705, United States
[c] USDA ARS Grazinglands Research Laboratory, El Reno, OK 73036, United States

ARTICLE INFO

Article history:
Received 1 October 2010
Received in revised form 25 January 2011
Accepted 26 January 2011
Available online 4 February 2011

Keywords:
Soil moisture
Hydrologic modeling
Data assimilation
Remote sensing

ABSTRACT

This paper examines the potential for improving Soil and Water Assessment Tool (SWAT) hydrologic predictions of root-zone soil moisture, evapotranspiration, and stream flow within the 341 km^2 Cobb Creek Watershed in southwestern Oklahoma through the assimilation of surface soil moisture observations using an Ensemble Kalman filter (EnKF). In a series of synthetic twin experiments assimilating surface soil moisture is shown to effectively update SWAT upper-layer soil moisture predictions and provide moderate improvement to lower layer soil moisture and evapotranspiration estimates. However, insufficient SWAT-predicted vertical coupling results in limited updating of deep soil moisture, regardless of the SWAT parameterization chosen for root-water extraction. Likewise, a real data assimilation experiment using ground-based soil moisture observations has only limited success in updating upper-layer soil moisture and is generally unsuccessful in enhancing SWAT stream flow predictions. Comparisons against ground-based observations suggest that SWAT significantly under-predicts the magnitude of vertical soil water coupling at the site, and this lack of coupling impedes the ability of the EnKF to effectively update deep soil moisture, groundwater flow and surface runoff. The failed attempt to improve stream flow prediction is also attributed to the inability of the EnKF to correct for existing biases in SWAT-predicted stream flow components.

Published by Elsevier Ltd.

1. Introduction

Soil moisture plays an essential role in the exchange of energy and water within the soil–vegetation–atmosphere continuum. Successful initialization and modeling of soil moisture is crucial for the prediction of hydrologic processes including runoff, ground water recharge and evapotranspiration. Nevertheless, accurate estimation of soil moisture is typically limited by uncertainties in model inputs, parameter values and imperfect model physics regarding subsurface processes. Given the lack of a dense soil monitoring network in most regions, satellite observations are the most viable solution to improving the representation of soil moisture states in land surface and hydrologic models.

During the past decade a range of data assimilation techniques have been developed to optimally merge land model estimates with satellite observations to reduce modeling errors arising from various sources (e.g. [1–3]). At their core, these approaches provide a methodology for properly updating error-prone model predictions with incomplete and uncertain observations of model states. A variety of assimilation approaches have been proposed for this task. However, in recent years the Ensemble Kalman filter (EnKF) has emerged as (arguably) the most popular choice for land data assimilation. The EnKF is based on generating a Monte Carlo ensemble of model predictions in order to propagate the background uncertainty information required by the Kalman filter update equations (see Section 2.1 below for further details). Relative to competing approaches, the EnKF offers the benefits of easy implementation, flexibility regarding the nature of modeling error, computational efficiency and demonstrated robustness when applied to land surface models [4,5]. However, most hydrologic EnKF applications have focused on the estimation of soil moisture profiles and surface energy fluxes in land surface models used in numerical weather prediction. In contrast, relatively little data assimilation work has been conducted for rainfall-runoff and/or stream flow models commonly applied to water resource quantity and quality studies. The few studies that have been completed generally show some potential for improving runoff prediction by assimilating surface soil moisture and/or stream flow observations (e.g. [6–10]).

Studies examining the assimilation of surface soil moisture are highly relevant given the expected wealth of global soil moisture data products created by the current ESA Soil Moisture Ocean Salinity mission (SMOS) [11] and the upcoming NASA Soil Moisture

* Corresponding author.
E-mail address: Wade.Crow@ars.usda.gov (W.T. Crow).

0309-1708/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.advwatres.2011.01.011

Active/Passive (SMAP) mission [12]. Both instruments will provide near-daily estimates of surface (0–5 cm) soil moisture – albeit at a relatively coarse spatial resolution of between 10 and 40 km. The ultimate value of these data products for improving water quality and quantity modeling is currently unknown.

The Soil and Water Assessment Tool (SWAT) is a physically-based, semi-distributed continuous watershed model developed to predict the impact of land management practices and climatic change on water, sediment and agricultural chemical yields over long periods of time [13,14]. SWAT has been widely applied to hydrologic (e.g. flow prediction, snow/runoff/groundwater/soil water dynamics, irrigation management) and water quality assessment (non-point source modeling, sediment yield, pollutant fate, best agricultural management practices, conservation effects) problems. Gassman et al. [15] provides a detailed review of the development and applications of SWAT. Despite its widespread and successful application to a number of critical water resources applications, SWAT is based on a much simpler representation of surface energy processes and the vertical redistribution of water within the soil column than land surface models used in past EnKF applications (see e.g. [16,17,2]). Given the importance of vertical processes that couple the surface to deeper model states in surface soil moisture data assimilation [18,19], it is unclear how effective existing land data assimilation techniques are when applied to SWAT. These issues must be addressed before SMOS and SMAP data products can be leveraged to enhance water resource applications currently addressed by SWAT modeling.

The objective of this study is to evaluate the potential of improving SWAT's hydrologic predictions (i.e. root-zone soil moisture, evapotranspiration, runoff and stream flow) within the 341 km^2 Cobb Creek Watershed in southwestern Oklahoma via the EnKF-based assimilation of surface soil moisture observations. The organization of this paper is as follows. Section 2 presents a review of the SWAT model and EnKF methodology, as well as details of the data used and a description of the design for the data assimilation experiments. Subsequent results are presented for two separate data assimilation experiments. Results in Section 3.1 are derived from a set of synthetic twin data assimilation experiments in which artificial observations are generated using the SWAT model. Results in Section 3.2 are analogous except for the more demanding case of assimilating actual soil moisture observations obtained within the Cobb Creek Watershed. Section 4 provides a brief summary and discussion of key results.

2. Methods and data

This section gives a brief description of the EnKF algorithm and summarizes basic SWAT physics with an emphasis on processes controlling runoff generation and the vertical redistribution of soil water. Methodologies for both the synthetic twin and real-data assimilation experiments are also presented.

2.1. Ensemble Kalman filter

As discussed above, the EnKF is a sequential data assimilation method evolved from the standard Kalman filter [20] that has been demonstrated to efficiently handle the assimilation of observations into moderately nonlinear models [5]. It is based on an ensemble generation of model states produced by adding Monte Carlo noise to model states and/or forcings to approximate the model forecast state error covariance matrix in order to optimally merge model predictions with observations.

Letting $Y(t)$ be a vector of background model states at time t and F a potentially non-linear land surface model, the continuous forecasting of $Y(t)$ via F can be expressed as:

$$\frac{dY(t)}{dt} = F[Y(t), w] \quad (1)$$

where the random noise term w represents the aggregate impact of modeling errors arising from various sources including: inadequate model physics, poorly calibrated parameters, and noisy forcing data.

Conversely, let Z_k be the observation vector collected at discrete time t_k and the observation process is derived as:

$$Z_k = M_k[Y(t_k)] + v_k \quad (2)$$

where M is the observation operator that relates the true state to the measured variable and v reflects the observation noise. The noise term v is assumed to be a mean-zero, Gaussian random variables with variance C_v and statistically independent of w.

The EnKF is based on minimizing the impact of w via the consideration of independent observations Z related to land surface states contained in Y. If F and M are linear and stated assumptions concerning v and w are met, then the optimal updating of Y replicates given the presence of an observation Z at time k can be expressed as:

$$Y_k^{i+} = Y_k^{i-} + K_k\left[Z_k + \varepsilon_k^i - M_k\left(Y_k^{i-}\right)\right] \quad (3)$$

where:

$$K_k = [C_{YM}(C_M + C_v)^{-1}]_{t=t_k} \quad (4)$$

and ε is a mean-zero, random variable independently sampled (for each ensemble member) from a mean-zero, Gaussian distribution with variance C_v (see [21]). Variables Y^{i+} and Y^{i-} in (3) are state vectors for the ith ensemble member before and after updating, respectively. K_k in (4) is the Kalman gain that defines the weights of measurement and model estimation and is calculated from the forecast error covariance matrix C_M of the measurement predictions $M_k[Y(t_k)]$ and the forecast cross covariance C_{YM} between any given state and $M_k[Y(t_k)]$. A single deterministic EnKF prediction (i.e. the "analysis") is then acquired by averaging model state predictions across the ensemble. The analysis of other model forecast variables (e.g. stream flow) is defined in the same manner.

2.2. Model description

SWAT is a physically-based, semi-distributed watershed model widely used to assess the impact of land management practices and climatic changes on long-term water, sediment and pollutant yields. A watershed is geographically delineated into a number of smaller sub-basins where flow routing is simulated. The sub-basins are further subdivided into hydrologic response units (HRU's) that consist of uniform land use, soil and management practices. While the area-fraction of a sub-basin covered by each HRU is accounted for, the exact location of each HRU is not explicitly represented. The HRU is a basic unit in SWAT where fundamental surface processes such as flow generation, soil water dynamics, crop growth, evapotranspiration, sediment and nutrient transport are simulated.

2.2.1. Flow generation

Total SWAT stream flow is calculated as

$$Q = Q_{surf} + Q_{lat} + Q_{gw} \quad (5)$$

where Q is total stream flow of the day (mm H_2O), Q_{surf} is surface runoff (mm H_2O), Q_{lat} is subsurface lateral flow (mm H_2O) and Q_{gw} is groundwater flow (mm H_2O). Surface runoff, lateral flow and groundwater flow are generated from each HRU and aggregated at the main channel of each sub-basin, then routed to obtain the total stream flow for the watershed.