Advances in Water Resources 86 (2015) 425–438

Contents lists available at ScienceDirect

Advances in Water Resources

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

Comparison of ensemble-based state and parameter estimation methods for soil moisture data assimilation

Weijing Chen [a,b], Chunlin Huang [b,c,*], Huanfeng Shen [a], Xin Li [b,d]

[a] School of Resource and Environmental Science, Wuhan University, Wuhan, Hubei 430079, China
[b] Key Laboratory of Remote Sensing of Gansu Province, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu 730000, China
[c] Heihe Remote Sensing Experimental Research Station, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu 730000, China
[d] CAS Center for Excellence in Tibetan Plateau Earth Sciences, Chinese Academy of Sciences, Beijing 100101, China

ARTICLE INFO

Article history:
Available online 10 August 2015

Keywords:
Data assimilation
Soil moisture
Brightness temperature
State-parameter estimation
Common Land Model

ABSTRACT

Model parameters are a source of uncertainty that can easily cause systematic deviation and significantly affect the accuracy of soil moisture generation in assimilation systems. This study addresses the issue of retrieving model parameters related to soil moisture via the simultaneous estimation of states and parameters based on the Common Land Model (CoLM). The state-parameter estimation algorithms AEnKF (Augmented Ensemble Kalman Filter), DEnKF (Dual Ensemble Kalman Filter) and SODA (Simultaneous optimization and data assimilation) are entirely implemented within an EnKF framework to investigate how the three algorithms can correct model parameters and improve the accuracy of soil moisture estimation. The analysis is illustrated by assimilating the surface soil moisture levels from varying observation intervals using data from Mongolian plateau sites. Furthermore, a radiation transfer model is introduced as an observation operator to analyze the influence of brightness temperature assimilation on states and parameters that are estimated at different microwave signal frequencies. Three cases were analyzed for both soil moisture and brightness temperature assimilation, focusing on the progressive incorporation of parameter uncertainty, forcing data uncertainty and model uncertainty. It has been demonstrated that EnKF is outperformed by all other methods, as it consistently maintains a bias. State-parameter estimation algorithms can provide a more accurate estimation of soil moisture than EnKF. AEnKF is the most robust method, with the lowest RMSE values for retrieving states and parameters dealing only with parameter uncertainty, but it possesses disadvantages related to increasing sources of uncertainty and decreasing numbers of observations. SODA performs well under the complex situations in which DEnKF shows slight disadvantages in terms of statistical indicators; however, the former consumes far more memory and time than the latter.

1. Introduction

Soil moisture is a key variable in understanding land surface hydrological processes that partition precipitation into runoff and infiltration and that control water storage and drainage [20]. As a vital element in the water and energy cycle, soil moisture forms the foundation of meteorological research, water resource regulation and agricultural management [9,13,37,54]. Modeling provides temporally and spatially continuous simulations and predictions of soil moisture but lacks precision. Meanwhile, many types of observations have uncertain accuracy and poor resolution, which is due to limited finances and the capabilities of the utilized instruments themselves. Many studies have noted that data assimilation has the potential to produce continuous and accurate soil moisture data sets that are reconciled in temporal and spatial resolution [6,18,22,27,28,41,52].

The assimilation of data originating from atmospheric and oceanographic sciences [15,21] takes full advantage of imperfect models and finite data in an optimal way by merging the information embodied in remote-sensing or ground-based networks into a dynamic model to improve forecast trajectory. Many experiments have been conducted to improve soil moisture estimation using *in situ* observations at the beginning of the development of land or hydrology data assimilation [5,19,26,55]. However, remote-sensing techniques dominate over *in situ* measurements in terms of the scope of an observed area. Low-frequency microwave brightness temperature is highly related to near-surface soil moisture and is only weakly affected by the atmosphere and clouds. Recently, many soil

* Corresponding author at: Key Laboratory of Remote Sensing of Gansu Province, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu 730000, China. Tel.: +869314967975.
E-mail address: huangcl@lzb.ac.cn (C. Huang).

http://dx.doi.org/10.1016/j.advwatres.2015.08.003
0309-1708/

moisture products have been applied to enhance model simulations on a regional scale because of the successful launch of a series of satellites with microwave sensors, such as AMSR-E (Advanced Microwave Scanning Radiometer for Earth Observing System), ASCAT (Advanced Scatterometer), and SMOS (Soil Moisture and Ocean Salinity) [3,4,12,39]. However, the large uncertainties that exist in retrieval algorithms may contaminate the quality of soil moisture products, which are expected to be used during assimilation. An alternative method is to directly assimilate brightness temperature into land surface or hydrological models to yield optimal soil moisture estimation [8,22,24,29,40].

As a fusion method, data assimilation improves model simulation by effectively handling background error, which is not the sole factor that influences the capability of data assimilation. Most applications of state assimilation focus on managing the errors that are embodied in the background, on observations and on forcing data, as opposed to employing model structure or parameters [1,25,48,53]. However, the outcome of data assimilation relies on an unbiased prediction of a model state, which is largely dependent on the effectiveness and representativeness of the model. Discrepancies between model parameterization and actual land surface processes account for model errors, but many studies are conducted under the assumption of a state-of-the-art model. Considering that parameter uncertainty affects state estimation to a large extent, the proper specification of model parameters as functions of variables when characterizing a state has become a crucial aspect of recent studies.

It is generally recognized that parameter calibration can diminish long-term bias, while state updating can weaken stochastic error. Thus, a calibration period is usually necessary to optimize the parameters of a hydrological model. Recently, many scientists have focused their attention on minimizing parameter errors in land data assimilation by performing simultaneous estimations of states and parameters [31,32,34,38,46,47,56]. This joint method expands the data assimilation framework from only updating model states to updating both model states and parameters.

Within the framework of EnKF-based assimilation, three types of algorithms are typically used for simultaneous state and parameter estimation. EnKF, which was originally proposed by Evensen [14], is a commonly used sequential algorithm for data assimilation and has shown strength in dealing with non-linear models because of its reliance on the propagation of a random ensemble of retrieved variables. EnKF is also an advantageous approach for highly dimensional applications, mainly because it captures the relevant parts of an error structure by means of a comparatively small ensemble of model trajectories, including (1) The state augmentation approach [2,17,33]. Monsivais-Huertero et al. [33] employed both synthetic and field observations to understand the effects of simultaneous state-parameter estimation using an augmented state vector, spatial and temporal update frequency and forcing data uncertainty in root-zone soil moisture. (2) The dual filter approach ([30,35,44]; Lü et al., 2008). Moradkhani et al. [35] presented a dual state-parameter estimation approach for the sequential estimation of parameters and states in a conceptual rainfall-runoff model (HyMOD) using observed streamflow. The algorithm is recursive, updating parameters and states in turn, and is mutually affective. (3) The parameter optimization and state assimilation approach [45,50]. Vrugt et al. [50] proposed the combined usage of parameter optimization and sequential data assimilation to facilitate valid treatment of input, output, parameter, and model structural errors in a Sacramento model, which was designated as the simultaneous optimization and data assimilation method (SODA).

Given the abundance of studies on joint state and parameter estimation and the paucity of investigations about algorithm applicability, the main objective of this study was to evaluate the performance abilities of all three of the above-discussed algorithms in a series of comparative experiments. We developed a data assimilation framework based on the common land model (CoLM), with soil moisture as the state variable of concern. First, we utilized *in situ* soil moisture to diagnose the performance of state assimilation at different observation intervals; at the same time, we examined the applicability of retrieving information regarding three soil property parameters (volume percentage of sand, volume percentage of clay and porosity). Second, we coupled the land surface model to a radiative transfer model (RTM), which acted as an observation operator, and added the standard deviation of the surface height into the parameter space. Brightness temperature was assimilated at different frequency combinations to judge the validity of each method. All of the soil moisture and brightness temperature experiments were implemented for three different cases: parameter uncertainty, atmospheric forcing data uncertainty and model uncertainty.

This paper is structurally organized as follows: models and methods are introduced in Section 2, in which the study area and experimental design are also described. The results of and discussion about the experiments are explained in Section 3. Section 4 presents other related discussions and the final conclusions.

2. Data assimilation scheme

2.1. Land surface model

The CoLM [10] is an improved version of the Community Land Model, with one vegetation layer, 10 unevenly spaced vertical soil layers, and up to 5 snow layers (depending on the total snow depth). We employed the CoLM as a dynamic model (model operator) to maintain its prognostic variables, which represent soil moisture in this work. The soil water equation is:

$$\frac{\Delta z_j}{\Delta t}\Delta\theta_j = [q_{j-1} - q_j] - f_{root,j*}E_{tr} \tag{1}$$

where $\Delta\theta_j$ is the change in water content as a result of the last time step in layer j, and Δz_j is the thickness of layer j. $f_{root,j*}$ and E_{tr} represent effective root fraction and transpiration, respectively. q_j is the water flow at the depth of the $z_{h,j}$ interface between layer j and layer j+1, as calculated by Darcy's law:

$$q = -K\left(\frac{\partial\psi}{\partial z} - 1\right) \tag{2}$$

K and ψ are the hydraulic conductivity and matric potential of soil, respectively, which vary with soil water content, θ, and soil texture based on the scheme proposed by Clapp and Hornberger [7].

(a) The hydraulic conductivity of soil, K, is:

$$K = K_{sat}s^{2B+3} \tag{3}$$

where the wetness (liquid water degree of saturation) is defined as:

$$s = \left[\frac{\theta_1}{1-\theta_d-\theta_i}\right] \tag{4}$$

where$1-\theta_{\mathrm{d}}$represents porosity, and the exponent B is defined as $B = 2.91 + 0.159(\%clay)$. For numerical reasons, when the effective porosity, $(1-\theta_d-\theta_i)$, is less than 0.05 in any of two neighboring layers, or when the liquid content is less than 0.001, then $K = 0$.

(b) The matric potential of soil is ψ, and the matric potential of unfrozen soil is:

$$\psi = \psi_{sat}s^{-B} \tag{5}$$

CoLM establishes the relationship between soil texture and soil thermal and hydraulic parameters as follows.