Advances in Water Resources 33 (2010) 201–214

Contents lists available at ScienceDirect

Advances in Water Resources

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

Towards the estimation root-zone soil moisture via the simultaneous assimilation of thermal and microwave soil moisture retrievals

Fuqin Li [a,b,*], Wade T. Crow [a], William P. Kustas [a]

[a] USDA-ARS Hydrology and Remote Sensing Lab, Bldg. 007, BARC-West, Beltsville, MD 20705, USA
[b] National Earth Observation Group, Geoscience Australia, GPO Box 378, ACT 2601, Australia

A R T I C L E I N F O

Article history:
Received 26 February 2009
Received in revised form 16 November 2009
Accepted 18 November 2009
Available online 22 November 2009

Keywords:
Data assimilation
Microwave remote sensing
Thermal remote sensing
Soil moisture

A B S T R A C T

The upcoming deployment of satellite-based microwave sensors designed specifically to retrieve surface soil moisture represents an important milestone in efforts to develop hydrologic applications for remote sensing observations. However, typical measurement depths of microwave-based soil moisture retrievals are generally considered too shallow (top 2–5 cm of the soil column) for many important water cycle and agricultural applications. Recent work has demonstrated that thermal remote sensing estimates of surface radiometric temperature provide a complementary source of land surface information that can be used to define a robust proxy for root-zone (top 1 m of the soil column) soil moisture availability. In this analysis, we examine the potential benefits of simultaneously assimilating both microwave-based surface soil moisture retrievals and thermal infrared-based root-zone soil moisture estimates into a soil water balance model using a series of synthetic twin data assimilation experiments conducted at the USDA Optimizing Production Inputs for Economic and Environmental Enhancements (OPE^3) site. Results from these experiments illustrate that, relative to a baseline case of assimilating only surface soil moisture retrievals, the assimilation of both root- and surface-zone soil moisture estimates reduces the root-mean-square difference between estimated and true root-zone soil moisture by 50% to 35% (assuming instantaneous root-zone soil moisture retrievals are obtained at an accuracy of between 0.020 and 0.030 m^3 m^{-3}). Most significantly, improvements in root-zone soil moisture accuracy are seen even for cases in which root-zone soil moisture retrievals are assumed to be relatively inaccurate (i.e. retrievals errors of up to 0.070 m^3 m^{-3}) or limited to only very sparse sampling (i.e. one instantaneous measurement every eight days). Preliminary real data results demonstrate a clear increase in the R^2 correlation coefficient with ground-based root-zone observations (from 0.51 to 0.73) upon assimilation of actual surface soil moisture and tower-based thermal infrared temperature observations made at the OPE^3 study site.

1. Introduction

Soil moisture is a key state variable governing the magnitude and variability of water and energy fluxes along the earth's surface. Given the inherent difficulty of obtaining large-scale measurements of soil moisture using ground-based networks, substantial effort has been dedicated to the development of remote sensing techniques to retrieve soil moisture patterns at continental and global scales. These efforts can be broadly categorized via their use of passive microwave [3,22], active microwave [41] or thermal [1,2,20,21] remote sensing retrieval techniques. Microwave-based approaches exploit the strong dielectric contrast between water and dry soil to directly estimate volumetric water content in the soil surface layer. In contrast, retrievals based on thermal remote sensing attempt to indirectly infer root-zone soil moisture based on an understanding of the thermal response of the vegetation canopy to limitations in available soil water [10,20,31].

Microwave and thermal-based soil moisture estimation techniques each possess their own unique set of advantages and limitations. For example, in the absence of hydrometeors, the atmosphere is approximately transparent at long microwave wavelengths (i.e. C- and L-band). As a result, microwave observations can penetrate through most cloud cover and provide almost daily coverage of the land surface. However, the value of these observations for soil moisture estimation is limited by their shallow penetration (approximately 2 cm for C-band sensors and between 3 and 5 cm for L-band) of the vertical soil column. In addition, the need for microwave observations at long wavelengths and engineering constraints surrounding the maximum size of deployable satellite antennae limits the resolution of passive microwave-based soil moisture retrievals to relatively coarse

* Corresponding author. Address: National Earth Observation Group, Geoscience Australia, GPO Box 378, ACT 2601, Australia. Tel.: +61 2 62495867; fax: +61 2 62499910.

E-mail address: Fuqin.Li@ga.gov.au (F. Li).

0309-1708/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2009.11.007

spatial resolutions (>10 km). Active microwave radar can provide finer-scale retrievals but at a generally lower accuracy level.

In contrast, thermal-derived soil moisture estimates are obtainable at lower temporal frequency since retrieval is not possible in the presence of cloud cover. However, because they measure plant response to soil moisture limitation throughout the root-zone, they are capable of inferring soil moisture variations over a deeper portion of the soil column. In addition, thermal remote sensing observations are potentially obtainable at much finer spatial resolutions than microwave observations.

Over the last decade, the assimilation of microwave-derived surface soil moisture into land surface models has been an active area of research [11,13,19,29,35,36,42]. Likewise, a number of past studies have examined the direct assimilation of surface radiometric temperatures (derived from thermal remote sensing observations) into a land model [9,23,26]. Recent work has also focused on the assimilation of energy flux estimates derived from surface radiometric retrievals [34,39] rather than the temperature retrievals themselves. In particular, [10] demonstrates the benefit of assimilating a root-zone soil moisture proxy (based on surface energy flux estimates obtained from thermal remote sensing) relative to the alternative strategy of directly assimilating surface radiometric temperature. However, relatively little work has focused on the simultaneous assimilation of mutually independent soil moisture information obtained from microwave and thermal remote sensing sources.

Using an Ensemble Kalman filter (EnKF), this paper conducts a series of synthetic data assimilation experiments at the USDA Economic and Environmental Enhancement (OPE^3) experimental site located in Beltsville, Maryland (http://www.ars.usda.gov/Research/docs.htm?docid=8438) to evaluate the relative utility of surface microwave and root-zone thermal soil moisture estimates for constraining model root-zone soil moisture predictions. These experiments represent the appropriate first step in the development of data assimilation systems based on new observation types [14,36,40] and should be completed prior to the assimilation of real measurements. In particular, synthetic experiments allow the performance of data assimilation systems to be assessed in tightly controlled circumstances where the nature of modeling and observation errors is known. For synthetic twin experiments presented here, we place special emphasis on accurately capturing temporal frequency and accuracy limitations in thermal-based soil moisture estimates and, therefore, realistically describing the added value of simultaneously assimilating both microwave and thermal-based soil moisture estimates relative to a baseline case of assimilating only microwave-based surface retrievals. The sensitivities of this added value to variations in the magnitude of modeling and observing errors, local soil texture, root-zone retrieval frequency, and the vertical correlation structure of modeling errors are examined through a series of synthetic experiments. While such synthetic data assimilation results represent an important first step, their results should be subsequently validated via data assimilation experiments involving real observations. As a first step in this direction, we also present results for the assimilation of real observations obtained at the OPE^3 site.

2. The ensemble Kalman filter

The EnKF is based on the generation of an ensemble of model realizations, each perturbed in a Monte Carlo manner, to estimate the state forecast error covariance information required by the standard Kalman filter update equation [4,15]. A decade of application in a variety of earth science fields have demonstrated that it is an effective approach for assimilating observations into moderately non-linear models [16].

Assuming that $Y(t)$ is vector of land surface state variables at time t, and $\mathbf{F}$ is a potentially non-linear land surface model, the updating of $Y(t)$ via $\mathbf{F}$ can be expressed as:

$$\frac{dY(t)}{dt} = \mathbf{F}[Y(t) + w] \tag{1}$$

where w captures perturbations due to model error arising from any potential combination of inadequate model physics, poorly calibrated parameters, and noisy forcing data. Here such random perturbations are assumed to be mean-zero. The EnKF is based on minimizing the impact of w via the consideration of independent observations related to land surface states contained in Y. Assuming $\mathbf{M}$ to be the true observation operator, observations Z (at a discrete time indexed by k) can be expressed as:

$$Z_k = \mathbf{M}_k[Y(t_k)] + v_k \tag{2}$$

where the observation perturbation v, a mean-zero, Gaussian random variable with covariance $\mathbf{C}_v$, represents degradation in the observations due to measurement noise or incomplete knowledge of $\mathbf{M}$. These perturbations are assumed to be statistically independent from the model perturbations introduced in (1).

The EnKF is initialized by the introduction of synthetic Gaussian error into initial conditions and generating an ensemble of model predictions using (1). If $\mathbf{F}$ and $\mathbf{M}$ are linear and all stated assumption 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:

$$Y_k^{i+} = Y_k^{i-} + \mathbf{K}_k[Z_k + v_k^i - \mathbf{M}_k(Y_k^{i-})] \tag{3}$$

where:

$$\mathbf{K}_k = [\mathbf{C}_{YM}(\mathbf{C}_M + \mathbf{C}_v)^{-1}]_{t=t_k} \tag{4}$$

State vectors Y^{i+} and Y^{i-} refer to the ith ensemble member before and after updating, respectively. $\mathbf{K}$ in (3) is the Kalman filter gain and defined as a function of: the cross-covariance between $\mathbf{M}(Y^{i-})$ and Y^{i-} ($\mathbf{C}_{YM}$), the covariance matrix of $\mathbf{M}(Y^{i-})$($\mathbf{C}_M$), and $\mathbf{C}_v$ via (4). The extra perturbation term v_k^i in (3) is required to ensure the proper spread in the post update ensemble [4]. Both $\mathbf{C}_{YM}$ and $\mathbf{C}_M$ are statistically estimated from all individual ensemble realizations and calculated around the ensemble mean. At any point in time, EnKF-based state estimates are obtained by averaging across the resulting ensemble of Y replicates. Past work has indicated that 30 ensemble members are generally appropriate for the application of the EnKF to one-dimensional land surface models [11,35]. However, in order to be conservative and ensure optimal performance, we use a relatively large ensemble size of 150 for all EnKF simulations presented here.

3. Models

Two separate land surface models are used in this analysis: (i) the thermal remote sensing two-source model (TSM) [25,27,33] and (ii) a one-dimensional water and energy balance (WEB) soil–vegetation–atmosphere transfer model (WEB-SVAT) [10] based on a force-restore concept for the soil water balance [12,32]. The WEB-SVAT model will be used as the basis for the synthetic twin data assimilation experiments described in Section 5. The TSM is used to estimate a root-zone soil moisture proxy based on tower-based thermal infrared temperature observations which are subsequently assimilated into the WEB-SVAT model during the real data analysis described in Section 6.

3.1. The two-source model (TSM)

In the TSM, remotely-sensed surface radiometric temperature (T_R) is expressed as the composite temperature of soil and canopy contributions: