J9.3 Integration of spaceborne precipitation and surface brightness temperature measurements using an Ensemble Kalman filter

W.T. Crow*
Hydrology and Remote Sensing Laboratory, ARS-USDA,
Beltsville, Maryland

1. Introduction

Efforts to observe components of the terrestrial hydrologic cycle from space are typically frustrated by the complex spatial and temporal patterns observed in these processes in concert with the sampling and resolution limitations of the spaceborne sensor. Two examples of this difficultly are the estimation of rainfall accumulations at short time scales ($<$ 1 week) using sparse (2 to 12 samples per day) retrievals of rainfall rates from space and the inability of passive microwave remote sensors to estimate soil moisture beyond a shallow surface layer (5 cm).

One strategy for addressing these shortcomings is the design of data assimilation systems to integrate numerical models of the land surface with remote observations. A well designed data assimilation system should be capable of combining land surface information in such a way that shortcomings in observational and computational tools are mutually compensated. For instance, surface soil moisture retrievals from space provides an indirect measure of antecedent rainfall that, if properly interpreted, could correct rainfall estimates derived from sparse temporal sampling of a rainfall event. Likewise, spaceborne rainfall measurements, combined with an accurate model of vertical water movement and evapotranspiration from the soil column could be used to vertically extrapolate surface soil moisture retrievals throughout the root zone (*Enekhabi et al.* 1994).

*Wade T. Crow, Blg. 007, Rm. 104, BARC-W, Beltsville, MD, 20705, wcrow@hydrolab.arsusda.gov

The purpose of this analysis is to examine the potential of a particular data assimilation approach - the Ensemble Kalman filter (EnKF) - to combine temporally sparse spaceborne precipitation estimates with shallow soil moisture retrievals. The approach is based on using an ensemble of precipitation realizations (consistent with both climatological expectations and sparse spaceborne measurements) to create an ensemble of land surface model predictions. These prediction are, in turn, updated using a remotely observed soil brightness temperature value and the EnKF. As a first approach for proof of concept, results presented here are generated using synthetically derived observations.

2. Methodology

a. Land Surface Modeling

Numerical modeling of the land surface was based on TOPographically-based Land Atmosphere Transfer Scheme (TOPLATS) simulations (*Peters-Lidard et al.* 1997) (*Famiglietti and Wood* 1994). Surface (5-cm) state predictions made by TOPLATS were processed through the Land Surface Microwave Emission Model (LSMEM) to produce corresponding estimates of L-band surface brightness temperature (T_B). Modeling was performed From 1 April 1997 to 31 March 1998 at the National Oceanic and Atmospheric Administration/Atmospheric Turbulence and Diffusion Division (NOAA/ATDD) Little Washita Watershed site near Chickasha, Oklahoma. Land cover at the site is grassland/rangeland. Validation of TOPLATS and LSMEM predictions was performed using flux tower data at the site and observations made during the 1997 Southern Great Plains Hydrology Experiment (SGP97).

b. Conditioned Rainfall Distributions

Using 15-minute rain gauge data collected at Oklahoma Mesonet stations during 1997 and 1999, daily rainfall accumulation estimates $\hat{R}$ were constructed based on the sub-sampling of ν 15-minutes rainfall rates per day. This data was used to mimic sampling errors associated with the sampling limitations of daily rainfall accumulations derived from spaceborne radar. Each estimate $\hat{R}$ was matched with a corresponding actual rainfall value R derived from all 96 daily 15-minute values. In this way, a series of conditional distributions $f_\nu(R/\hat{R})$ were constructed for various discrete ranges of $\hat{R}$. These distributions describe the manner in which knowledge of $\hat{R}$ conditions expectations about actual rainfall accumulations. This conditioning, of course, depends on the frequency of sampling which support the estimation. Sets of conditional distribution were constructed for ν values of 2, 4, 6, 8, and 12 day^{-1}.

c. The Ensemble Kalman Filter

The Ensemble Kalman filter (EnKF) is based on the generation of an ensemble of model predictions to estimate the error/covariance information required by the standard Kalman filter (KF) for the updating of model predictions with observations (*Reichle et al.* 2002) (*Evensen* 1994). The EnKF can be generalized using a state space representation of prediction and observation operators. Take $\mathbf{Y}(t)$ to be a vector of land surface state variables at time t. The equation describing the evolution of these states, as determined by a potentially nonlinear land surface model $\mathbf{f}$, is given by:

$$\frac{d\mathbf{Y}}{dt} = \mathbf{f}(\mathbf{Y}, \mathbf{w}) \tag{1}$$

where $\mathbf{w}$ relates errors in model physics, parameterization, and/or forcing data and is taken to be mean zero with a covariance $\mathbf{C}_w$. The goal of the filtering problem is to constrain these predictions using a set of observations which are related to the model states contained in $\mathbf{Y}$. Let the operator $\mathbf{M}$ represent the observation process which relates $\mathbf{Y}$ to the actual measurements taken at time t_k:

$$\mathbf{Z}_k = \mathbf{M}(\mathbf{Y}(t_k), \mathbf{v}_k) \tag{2}$$

where $\mathbf{v}_k$ represents Gaussian measurement error with covariance $\mathbf{C}_{vk}$. The EnKF is initialized by the introduction of synthetic Gaussian error into initial conditions and generating an ensemble of model predictions using equation (1). At the time of measurement predictions made by the ith model replicate are referred to as the state forecast $\mathbf{Y^i_-}$. If $\mathbf{f}$ is linear and all errors are additive, independent and Gaussian, the optimal updating of $\mathbf{Y^i_-}$ by the measurement $\mathbf{Z}_k$ is given by:

$$\mathbf{Y^i_+} = \mathbf{Y^i_-} + \mathbf{K}_k[\mathbf{Z}_k - \mathbf{M}_k(\mathbf{Y^i_-})] \tag{3}$$

and:

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

where $\mathbf{C}_M$ is the error covariance matrix of the measurement forecasts $\mathbf{M}_k(\mathbf{Y^i_-})$ and $\mathbf{C}_{YM}$ is the cross-covariance matrix linking the predicted measurements with the state variables contained in $\mathbf{Y^i_-}$. All covariance values are statistically estimated around the ensemble mean. Here $\mathbf{Y^i_+}$ signifies the updated or analysis state representation.

d. Fraternal Twin Experiment

The overall methodology of the analysis was based on a fraternal twin data assimilation experiment. First, a single TOPLATS/LSMEM simulation, validated by independent observations and forced by all available meteorological data, was designated as truth. "True" T_B observations were then perturbed with random error ($\mathbf{C}_{vk}$) to form the set of observations $\mathbf{Z}_k$. Based on $\hat{R}$ values derived from 15-minute rainfall data, daily rainfall ensembles were generated by sampling from the appropriate $f_\nu(R/\hat{R})$ distribution. Sampled daily rainfall amount were downscaled to hourly based on the fraction of daily precipitation rate samples where rainfall was detected. These hourly rainfall realizations were, in turn, used to generate an ensemble of TOPLATS state ($\mathbf{Y}$) and observation ($\mathbf{M}$) forecasts from which the prior estimates of model errors ($\mathbf{C}_{YM}$ and $\mathbf{C}_M$) could be obtained. Each member of the ensemble was then updated using the perturbed observations in $\mathbf{Z}_k$ and the EnKF update equation given in (3).

An analogous set of updated open loop TOPLATS simulations were constructed using a sub-sampling procedure on 15-minute rainfall gauge data at the study site. For a daily sampling frequency of ν, rainfall rates were assumed constant and equal to the observed 15-minute gauge-derived rate for the 24ν^{-1} hour period centered on each observation. Open loop simulations describe the accuracy of model results in the absence of updating with T_B observations. Both set of results (EnKF and open loop) were repeated using each of 15-minute rainfall observations in the first 24ν^{-1} hour sampling interval as the simulation start time.

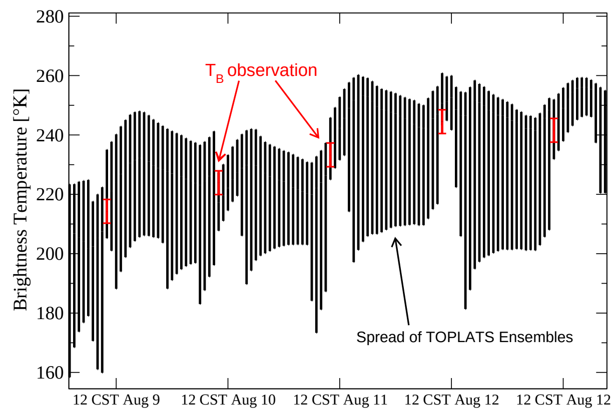

Figure 1: Demonstration of the EnKF assimilation procedure.

3. Results

Figure 1 shows the sequential operation of the EnKF for a five day period during summer 1997. Ensemble spreads for T_B originate from forcing TOPLATS with rainfall realizations sampled from the appropriate $f_\nu(R/\hat{R})$ distribution. T_B measurements (in red) update each member of the ensemble using error statistics sampled from the ensembles and (3). This updating adjusts particular realizations to compensate for rainfall forcing errors and reduces the total spread of the ensemble.

Figure 2b compares EnKF and open loop results for integrated soil moisture values within the top 40-cm of the soil column. EnKF results are for an ensemble size of 25, an assumed $\mathbf{C}_v$ of 16 K^2, daily observations of T_B and 8 rainfall rate samples per day. Open loop results are for 8 sample of rainfall rate per day and no T_B observations. It is important to note that 40-cm is roughly 8 times the penetration depth of L-band T_B observations and far beyond the depth that can be accurately updated through direct insertion of surface soil moisture retrievals.

Open loop and EnKF results in Figure 1b are for a rainfall observation start time of 0000 CST. Analogous results can be obtained for start times corresponding to each of 15-minute intervals between 0000 and 0300 CST. Figure 1c shows root-mean-square (RMS) errors for results derived from all possible start times. Throughout the growing season the EnKF-based integration of T_B measurements substantially improves the representation of 40-cm soil moisture dynamics.

Figure 3 gives RMS results averaged over the entire simulation period for a range of rainfall and T_B sampling frequencies. Generally, the assimilation of T_B measurements reduces the RMS error in 40-cm soil moisture predictions by more than 60%. The correction associated with assimilated T_B observations is substantial even for T_B measurement frequencies as low as once every five days. As T_B measurements become more frequent, the observed sensitivity of soil moisture errors to rainfall observation frequency decreases. To the point where assimilation of daily T_B measurements essentially eliminates the sensitivity of soil moisture results to the frequency of rainfall observations.

4. Discussion

Results demonstrate the basic feasibility and effectiveness of the EnKF as a framework for integrating daily rainfall estimates derived from temporally sparse rate estimates with surface T_B observations for the monitoring of root-zone soil moisture. Results with synthetically generated data show that the approach can theoretically correct for a substantial fraction of soil moisture error asso-

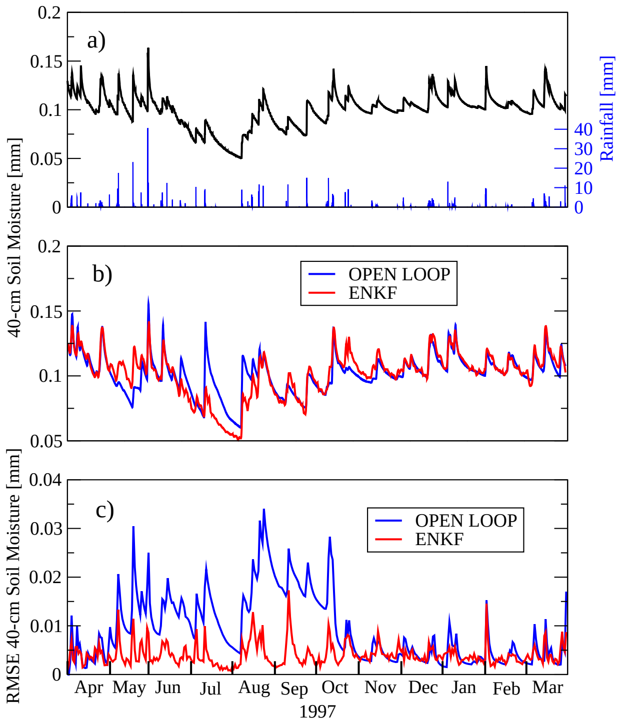

Figure 2: a) Times series of rainfall and benchmark 40-cm soil moisture predictions. b) Open loop and EnKF results for 40-cm soil moisture given 8 rainfall samples per day and daily T_B observations. c) Open loop and EnKF 40-cm soil moisture RMS errors for all possible rainfall observation starting times.

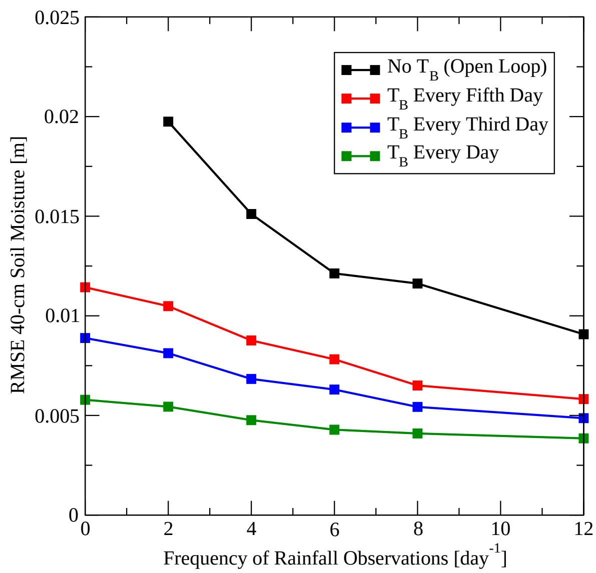

Figure 3: RMS errors for 40-cm soil moisture results derived from various combinations of T_B and rainfall measurement frequencies.

ciated with the sparse temporal sampling of rainfall. Current work focuses on demonstrating analogous results with actual observations and more precise assessments of filter performance.

References

Entekhabi, D., H. Nakamura, and E.G. Njoku, Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations, IEEE Trans. Geosci. Remote Sens., 32:438-447, 1994.

Evensen, G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. of Geophys. Res., 99:10143-10162, 1994.

Famiglietti, J.F., and E.F. Wood, Multiscale modeling of spatially variable water and energy balance processes, Water Resour. Res., 30:3061-3078, 1994.

Peters-Lidard, C.D., M.S. Zion, and E.F. Wood, A soil-vegetation-atmosphere transfer scheme for modeling spatially variable water and energy balance processes, J. Geophys. Res., 102:4303-4324, 1997.

Reichle, R.H., D.B. McLaughlin, and D. Entekhabi, Hydrologic data assimilation with the Ensemble Kalman filter, Mon. Weather Rev., 130:103-114, 2002.