4.11

A SYNTHETIC STUDY ON THE INFLUENCE OF ERROR IN SURFACE SOIL MOISTURE OBSERVATIONS ON ASSIMILATION

Jeffrey P. Walker * and Paul R. Houser
NASA Goddard Space Flight Center, Greenbelt, Maryland

1. INTRODUCTION

Defensible requirements of a remote sensing mission for the measurement of surface soil moisture are of vital importance to scientists planning such a mission. In particular, mission planners need: (i) justification for polarization, wavelength and look angle requirements of the sensor; and (ii) accuracy, temporal resolution and spatial resolution requirements of the measurement. The requirements of (i) have been fairly well defined, with horizontally polarized L-band radiometer measurements at a look angle of less than 50º yielding the greatest sensitivity to soil moisture. However, the requirements of (ii) have been less well defined. This paper seeks to address the first of those three issues; measurement accuracy requirements. The remaining two issues will be addressed in forthcoming papers.

2. MODELS

The measurement accuracy requirement is addressed in this paper through a synthetic data assimilation study. First, a land surface model is used to generate a “truth” data set that provides both the surface soil moisture “observations” and the evaluation data. The land surface forcing data and initial conditions are then degraded to simulate the uncertainties in these data and a second simulation performed. Finally, simulations are made where the observations, with various levels of error imposed, are assimilated into the simulation with degraded atmospheric forcing data and initial conditions.

2.1 Land Surface Model

The land surface model used in this study is the catchment-based land surface model of Koster *et al.* (2000), illustrated schematically in Figure 1. It uses a non-traditional land surface model framework that includes an explicit treatment of sub-grid soil moisture variability and its effect on runoff and evaporation. A key innovation in this model is the shape of the land surface element, the hydrologic watershed as defined by the topography, rather than an arbitrary grid.

This land surface model uses TOPMODEL (Beven and Kirkby, 1979) concepts to relate the water table distribution to the topography. The consideration of both the water table distribution and non-equilibrium conditions in the root zone leads to the definition of three bulk moisture prognostic variables (catchment deficit, root zone excess and surface excess) and a special treatment of moisture transfer between them.

* *Corresponding author address:* Jeffrey P. Walker, Rm. B309 Build. 33, NASA/GSFC, Mail Code 974, Greenbelt, MD USA; Tel: +1-301-614-5804; E-Mail: cejpw@land.gsfc.nasa.gov

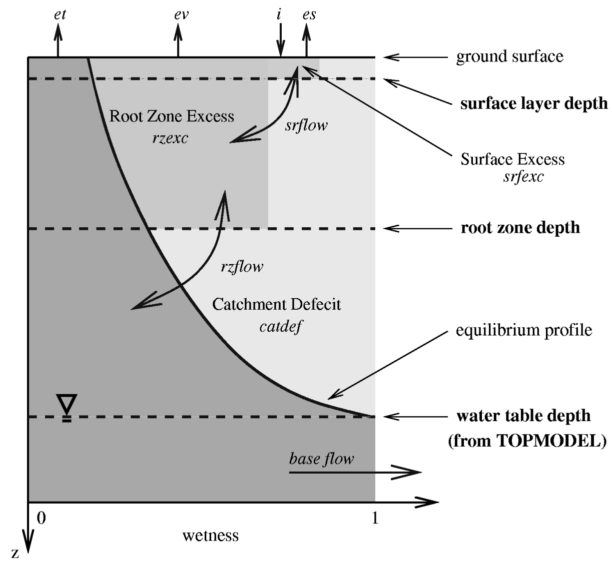

Figure 1: Schematic of the catchment-based land surface model.

Using these three prognostic variables, the catchment may be divided into regions of stressed, unstressed and saturated soil moisture regimes. The soil moisture prognostic variable forecasting equations of the catchment-based land surface model are given by:

$$srfexc^{n+1} = srfexc^{n} - srflow + i - es \tag{1}$$

$$rzexc^{n+1} = rzexc^{n} + srflow - rzflow - ev \tag{2}$$

$$catdef^{\,n+1} = catdef^{\,n} - rzflow + baseflow + et\,, \tag{3}$$

where $srfexc$ is the surface excess, $rzexc$ is the root zone excess and $catdef$ is the catchment deficit. The redistibution between the surface and root zone excesses is given by $srflow=f(srfexc,rzexc)$ and between the root zone excess and catchment deficit is given by $rzflow=f(rzexc,catdef)$. The baseflow is given by $baseflow=f(catdef)$, soil infiltration i, bare soil evaporation es, transpiration ev, evapotranspiration et and soil moisture in the surface layer (2 cm) are given by non-linear functions $f(srfexc,rzexc,\ catdef)$. A complete description of this model is given in Koster *et al.* (2000).

2.2 Kalman Filter

The Kalman filter algorithm tracks the conditional mean of a statistically optimal estimate of a state vector $\mathbf{X}$ and its covariance matrix $\Sigma_{\mathbf{X}}$, through a series of forecasting and update steps. The forecasting equations are:

$$\hat{\mathbf{X}}^{n+1/n} = \mathbf{A}^{n} \cdot \hat{\mathbf{X}}^{n/n} + \mathbf{U}^{n} + \left(\mathbf{w}^{n}\right) \tag{4}$$

$$\mathbf{\Sigma}_{\mathrm{x}}^{n+1/n} = \mathbf{A}^{n} \cdot \mathbf{\Sigma}_{\mathrm{x}}^{n/n} \cdot \mathbf{A}^{n^{\mathrm{T}}} + \mathbf{Q}^{n}\,, \tag{5}$$

where $\mathbf{A}$ is the state propagation matrix relating the system states at times $n+1$ and n, $\mathbf{U}$ is a vector of forcing, $\mathbf{w}$ is the model error and $\mathbf{Q}$ is the covariance matrix of the system noise (model error), defined as $E[\mathbf{w}\cdot\mathbf{w}^T]$. The notation $n+1/n$ refers to the system state estimate at time $n+1$ from a forecasting step, and n/n refers to the system state estimate from either a forecasting or updating step at time n. The update equations can be found in Bras and Rodriguez-Iturbe (1985).

In this study, we have used a one-dimensional Kalman filter for updating the soil moisture prognostic variables of the land surface model. A one-dimensional Kalman filter was used because of its computational efficiency and the fact that at the scale of catchments used, correlation between the soil moisture prognostic variables of adjacent catchments is only through the large-scale correlation of atmospheric forcing. Moreover, all calculations for soil moisture in the land surface model are performed independent of the soil moisture in adjacent catchments.

Forecasting of the soil moisture prognostic variables covariance matrix was achieved through linearization of the soil moisture forecasting equations. The linearization was performed by a first order Taylor series expansion of the non-linear forecasting equations (1-3). Using this approach, the covariance forecasting matrix is given by

$$\mathbf{A}=\begin{bmatrix} \dfrac{\partial srfexc^{n+1}}{\partial srfexc^{n}} & \dfrac{\partial srfexc^{n+1}}{\partial rzexc^{n}} & \dfrac{\partial srfexc^{n+1}}{\partial catdef^{n}} \\ \dfrac{\partial rzexc^{n+1}}{\partial srfexc^{n}} & \dfrac{\partial rzexc^{n+1}}{\partial rzexc^{n}} & \dfrac{\partial rzexc^{n+1}}{\partial catdef^{n}} \\ \dfrac{\partial catdef^{n+1}}{\partial srfexc^{n}} & \dfrac{\partial catdef^{n+1}}{\partial rzexc^{n}} & \dfrac{\partial catdef^{n+1}}{\partial catdef^{n}} \end{bmatrix}. \qquad (6)$$

For the initial covariance matrix, diagonal terms were specified to have a standard deviation of the maximum difference between the initial prognostic state value and the upper and lower limits, with off diagonal terms specified as zero. The diagonal terms of the forecast model error covariance matrix $\mathbf{Q}$ were taken to be the predefined values of 0.0025, 0.025 and 0.25 mm/min for *srfexc*, *rzexc* and *catdef* respectively, with the off diagonal terms taken to be zero.

3. SYNTHETIC EXPERIMENTS

To demonstrate the effect of observation error on retrieval of the soil moisture profile by assimilation, a set of synthetic experiments have been undertaken for the entire North American continent.

3.1 Model Input Data

In this study, atmospheric forcing data and soil and vegetation properties from the first International Satellite Land Surface Climatology Project (ISLSCP) initiative (Sellers *et al.,* 1996) have been used as model input for the year 1987. Soil properties not defined by ISLSCP were assumed uniform with the values in Table 1. Total soil depth had a variation of 1 to 3.6 m. Initial model states were derived by driving the model to equilibrium at the beginning of 1987.

Table 1: Uniform soil properties specified for North America.

saturated surface hydraulic conductivity	2.2×10^{-3} m s^{-1}
transmittivity decay factor	3.26 m^{-1}
saturated soil matric potential	–0.281 m
Clapp and Hornberger (1978) *b*	4
root zone depth	1 m
wilting point	14.8 % v/v

3.2 Observation and Evaluation Data

Using the catchment-based land surface model of Koster *et al.* (2000), the initial conditions from spin-up and the model input data described above, the temporal and spatial variation of soil moisture across the North American continent was forecast for 1987. The forecasts of surface soil moisture were output once every 3 days to represent the soil moisture that would be measured by a remote sensing satellite. In addition to surface soil moisture observation data, this simulation provided the data for evaluation of degraded simulations.

3.3 Degraded Simulation

To represent the errors associated with forecast land surface states in a typical land surface model simulation as a result of poor initial conditions and errors in atmospheric forcing data, both the initial conditions and forcing data were degraded. The initial conditions were degraded by applying zero mean normally distributed random perturbations with the standard deviations given in Table 2, to each of the three soil moisture prognostic variables from the original spin-up data. The forcing data were similarly degraded using the standard deviations in Table 2 to represent the uncertainty associated with atmospheric forcing data, as a result of both measurement and interpolation error.

Applying perturbations to precipitation was more difficult than other forcing parameters, as the occurrence of precipitation is an intermittent process. Hence, precipitation was perturbed by a fraction of the precipitation rate to account for spatial variability. To account for the fact that precipitation could have occurred even when the data suggested there was none, a perturbation to precipitation was added whenever a randomly distributed zero mean number greater than three times its standard deviation was generated. Under this situation, the standard deviation

Table 2: Standard deviations used for applying a normally distributed random perturbation to the initial conditions and atmospheric forcing data.

srfex	1 mm
rzex	10 mm
catdef	100 mm
convective precipitation	50% or 0.1 to 8 mm hr^{-1}
total precipitation	50% or 0.1 to 8 mm hr^{-1}
2 m air temperature	5 ° C
2 m dewpoint temperature	5 ° C
downward longwave radiation	25 w m^{-2}
downward shortwave radiation	50 w m^{-2}
surface pressure	1 kPa
10 m wind speed	1 m s^{-1}

for the perturbation was taken as 1 mm hr^{-1}, multiplied by the ratio of mean annual precipitation for the catchment (55 mm to 4595 mm) to the average mean annual precipitation (595 mm) for the North American continent.

As wind speed, downward radiation and precipitation cannot be negative, negative values after perturbation were truncated to zero. A time series histogram of the error introduced into the precipitation is given in Figure 2a and the error introduced in soil moisture forecasts for the entire soil profile is given in Figure 2b. Both of these figures indicate a small wet bias in the degraded simulation during summer months.

In view of the fact that this study assumed a perfect model, significant error in the degraded simulation was ensured by degradation in both initial conditions and atmospheric forcing, namely precipitation. Since the catchment-based land surface model operated at the dry end of the scale over much of the North American continent in the evaluation simulation, any significant error introduced in the model must be on the wet end of the scale. Hence the degraded simulation resulted in a wet bias, both in the initial condition and forecasts of the pursuing summer months. Model error could have been accounted for by generating the observation and evaluation data from a different land surface model scheme, but that would not have had a significant influence on the resulting conclusions drawn from this study.

3.4 Degraded Observations

The effect of error in surface soil moisture observations is demonstrated by adding zero mean normally distributed perturbations to the surface soil moisture observation data set described above. Standard deviations used for generating perturbations were 1, 2, 3, 4, 5 and 10% v/v.

3.5 Effect of Observation Error

To demonstrate the effect of observation error on soil moisture profile retrieval, individual simulations were made where the degraded observation data were assimilated into the degraded simulation described above. The resulting time series histogram for errors in the retrieved soil moisture for the entire soil profile is given in Figure 2c for perfect observations, and Figure 2d for observations with 4% v/v error. In both situations the bias in the soil moisture forecast has been improved, but the amount of error in the forecast has increased for the latter.

The effect of observation error on soil moisture profile retrieval can be seen further in Figure 3. This figure shows that the root mean square (rms) error in both soil moisture retrieval and evapotranspiration forecasts increased with observation error, with an observation error of less than 3% v/v required for soil moisture retrieval to be better than the original degraded simulation. However, the rms error in evapotranspiration forecasts was always greater than for the original degraded simulation, but provided the observation error was less than 3% v/v there was an improvement in the evapotranspiration bias. For soil moisture however, provided the observation error was less than 5½% v/v there was an improvement in the bias.

3.6 Effect of Forcing Bias

Figure 2 showed that soil moisture forecasts for the entire soil profile had a wet bias in the original degraded simulation, as a result of a wet bias in the precipitation forcing. However, this switched to a dry bias in the retrieved soil moisture profile for perfect observations, while the simulation with greater observation error still had a slight wet bias. This switching in the bias can be seen more clearly in Figure 3b.

The bias in retrieved soil moisture for the entire soil profile is the result of a violation of one of the key assumptions of the Kalman filter; that the continuous time error process $\mathbf{w}$ is a zero mean Gaussian white noise stochastic process. Since the precipitation field was wet biased, the surface soil moisture was always wet biased. The Kalman filter recognized that the soil moisture in the surface layer had a strong link with the sub-surface, so a dry bias was introduced in the deep layer to counteract this wet bias in the surface layer. As the observation error was increased, the weight given to observations relative to model forecasts was decreased,

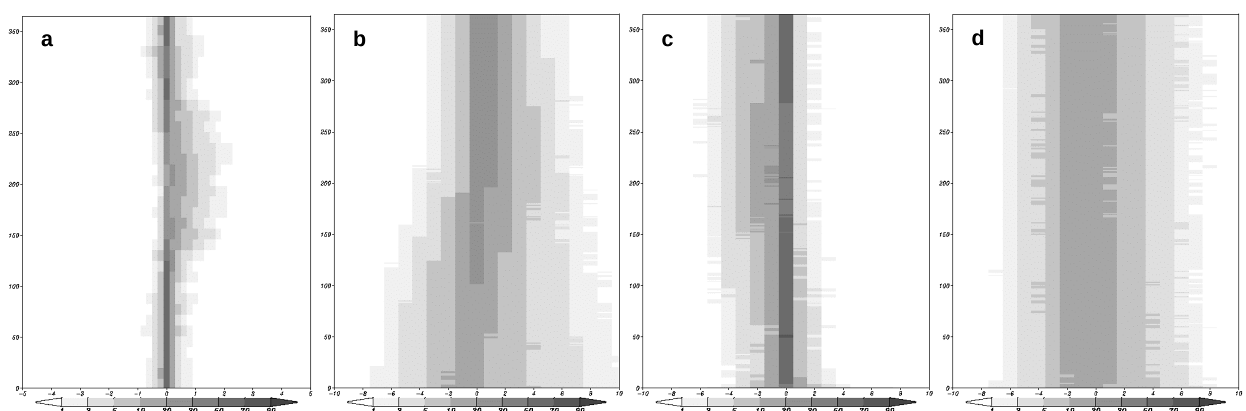

Figure 2: Time series (vertical axis) histogram (% of catchments) of a) errors in precipitation (horizontal axis - mm/day); and errors in soil moisture for the entire soil profile (horizontal axis - % v/v) for b) no assimilation, c) assimilation of perfect observations and d) assimilation of observations with 4% v/v error.

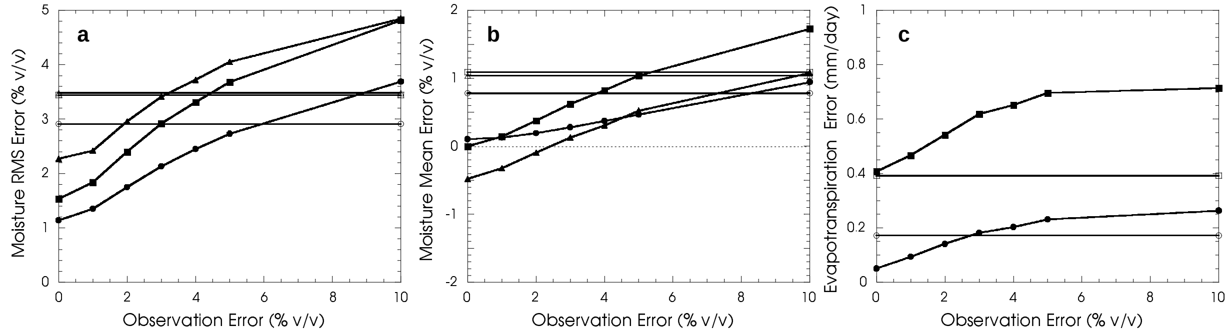

Figure 3: Effect of observation error in surface soil moisture data on: a) surface (circle), root zone (square) and profile (triangle) soil moisture rms error; b) surface (circle), root zone (square) and profile (triangle) soil moisture mean error; and c) evapotranspiration rms (square) and mean (circle) error. Simulations with assimilation (solid symbols) are compared with the simulation without assimilation (open symbols).

meaning that the assimilation could not have as large an impact on the sub-surface soil moisture.

To further illustrate the effect of forcing bias on soil moisture retrieval and evapotranspiration forecasts, additional simulations were made without any information on precipitation and with a perturbation standard deviation of 100% on precipitation. The effect of this precipitation bias on soil moisture profile retrieval can be seen in Figure 4. This figure shows that rms errors in soil moisture and evapotranspiration forecasts are improved irrelevant of precipitation bias when perfect observations are assimilated. However, the best results were obtained when the precipitation bias was a minimum. The resulting bias in surface soil moisture and evapotranspiration forecasts with assimilation were largely unaffected by the bias in precipitation, but the soil moisture forecasts for the root zone and entire soil profile were heavily impacted. Moreover, these results have shown that it is better to use poor information on precipitation than to use no information on precipitation.

4. CONCLUSIONS

This study has shown that the observation error in surface soil moisture must be less than the error required in model forecasts of soil moisture, else a slight degradation of the model forecasts may result. Typically, observations of surface soil moisture must have an accuracy better than 5% v/v. This study has also shown the importance of unbiased forcing in an assimilation framework. This is a result of assumptions made by the assimilation algorithm. Bias in the observations would likely have a similar effect on the resulting forecasts. A forthcoming paper will discuss the impact of temporal and spatial resolution.

ACKNOWLEDGEMENTS

This work was supported by funding from the NASA Seasonal-to-Interannual Prediction Project (NSIPP) NRA 98-OES-07.

REFERENCES

Beven, K.J. and M.J. Kirkby, 1979: A physically based, variable contributing area model of basin hydrology, *Hydrological Sciences - Bulletin*, **24**(1), 43-69.

Bras, R.L., and I. Rodriguez-Iturbe, 1985: *Random Functions and Hydrology*, Addison Wesley, Reading, Massachusetts, 559 pp.

Koster, R.D., M.J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a GCM. Part 1: Model structure, *Journal of Geophysical Research - Atmospheres*, *In Press*.

Sellers, P., B.W. Meeson, J. Closs, J. Collatz, F. Corprew, D. Dazlich, F.G. Hall, Y. Kerr, R. Koster, S. Loss, K. Mitchell, J. McManus, D. Myers, K.J. Sun, and P. Try, 1996: The ISLSCP Initiative I global data sets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies. *Bulletin of American Meteorological Society*.

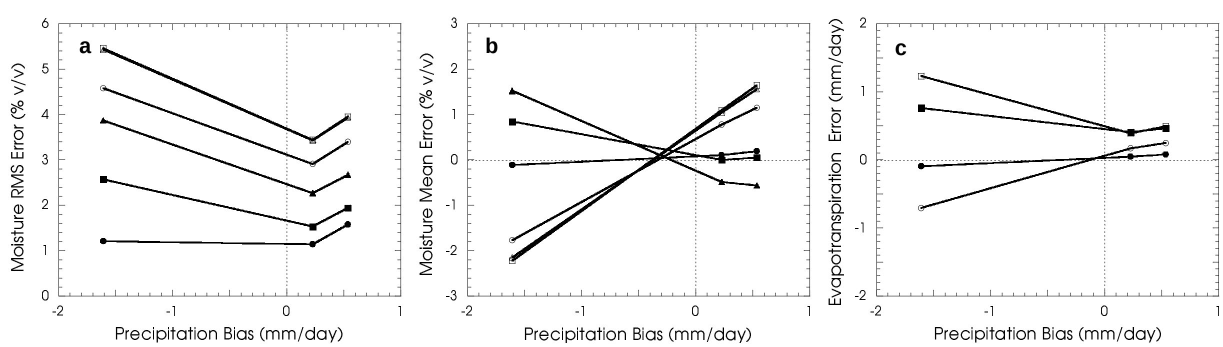

Figure 4: Effect of precipitation bias on: a) surface (circle), root zone (square) and profile (triangle) soil moisture rms error; b) surface (circle), root zone (square) and profile (triangle) soil moisture mean error; and c) evapotranspiration rms (square) and mean (circle) error. Simulations with assimilation (solid symbols) are compared with the simulation without assimilation (open symbols).