ASSIMILATION OF LAND SURFACE DATA

PAUL R. HOUSER
NASA Goddard Space Flight Center,
Greenbelt, Maryland 20771 USA

1. Introduction

Charney *et al.* (1969) first suggested combining current and past data in an explicit dynamical model, using the model's prognostic equations to provide time continuity and dynamic coupling amongst the fields (Figure 1). This concept has evolved into a family of techniques known as *four-dimensional data assimilation*. "Assimilation is the process of finding the model representation which is most consistent with the observations" (Lorenc, 1995). In essence, data assimilation merges a range of diverse data fields with a model prediction to provide that model with the best estimate of the current state of the natural environment so that it can then make more accurate predictions. The application of data assimilation in hydrology has been limited to a few one-dimensional, largely theoretical studies (*i.e.* Entekhabi *et al.*, 1994; Milly, 1986), primarily due to the lack of sufficient spatially-distributed hydrologic observations (McLaughlin, 1995). However, the feasibility of synthesizing distributed fields of soil moisture by the novel application of four-dimensional data assimilation applied in a hydrological model was demonstrated by Houser *et al.* (1998). Six Push Broom Microwave Radiometer images gathered over the United States Department of Agriculture, Agricultural Research Service Walnut Gulch Experimental Watershed in southeast Arizona were assimilated into a land surface model using several alternative assimilation procedures. Modification of traditional assimilation methods was required to use these high-density Push Broom Microwave Radiometer observations. The images were found to contain horizontal correlations with length scales of several tens of kilometres, thus allowing information to be advected beyond the area of the image. Information on surface soil moisture was also assimilated into the subsurface using knowledge of the surface-subsurface correlation. Newtonian nudging assimilation procedures were found to be preferable to other techniques because they nearly preserve the observed patterns within the sampled region, but also yield plausible patterns in unmeasured regions, and allow information to be advected in time.

The feasibility of land surface data assimilation methods has been recently tested in research projects conducted at Goddard Space Flight Center, Massachusetts Institute of Technology, and several other institutions. This research focuses on: (1) the use of a one-dimensional Kalman filtering based land assimilation strategy that expands upon the one-dimensional, theoretical assimilation algorithms developed by Entekhabi *et al.* (1994) and Milly, (1986), and (2) the four-dimensional data assimilation strategies developed by Houser *et al.* (1998) and Walker *et al.* (1999).

R. Swinbank et al. (eds.), Data Assimilation for the Earth System, 331–343.

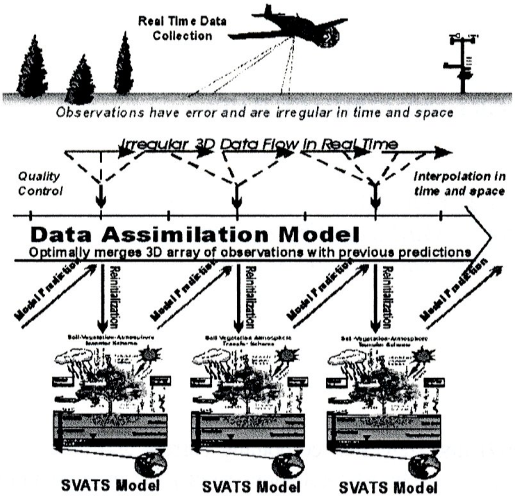


Figure 1: The land surface data assimilation process.

2. The Kalman filter

The Kalman filter attempts to obtain an optimal estimate of the land surface state by combining observations of that state with a land surface model forecast of that state. The Kalman filter has been extensively utilized in data assimilation research (Ghil *et al.*, 1981; Cohn, 1982). The Kalman filter assimilation scheme is a linearized statistical approach that provides a statistically optimal update of the system states, based on the relative magnitudes of the covariances of both the model system state estimate and the observations. The principal advantage of this approach is that the Kalman filter provides a framework within which the entire system is updated with covariances representing the reliability of the observations and model prediction.

The Kalman filter algorithm (Kalman, 1960) tracks the conditional mean of a statistically optimal estimate of a state vector X through a series of forecasting and update steps. To apply the Kalman filter, the equations for evolving the system states must be written in the linear state space formulation of Equation (1). When these equations are non-linear, the Kalman filter is called the extended Kalman filter, and is an approximation of the non-linear system that is based on first-order linearization. Walker and Houser (2001) have implemented a one-dimensional version of the extended Kalman filter with the simplifying assumption that errors in different catchments are uncorrelated.

The ensemble Kalman filter is an alternative to the extended Kalman filter for non-linear problems (Evensen, 1994; Houtekamer and Mitchell, 1998). The ensemble Kalman filter is based on the propagation of an ensemble of states from which the required covariance information is obtained at the time of the update. This approach has successfully been introduced into ocean assimilation at the National Aeronautics and Space Administration Seasonal-to-Interannual Prediction Project (NSIPP) by Keppenne and Rienecker (2000). Reichle *et al.* (2001) applied the ensemble Kalman filter to the soil moisture estimation problem and found it performs well, with the distinct advantage that the error covariance propagation is better behaved in the presence of large model nonlinearities. In general form, the nonlinear land surface model is expressed as:

$$X_k = f_k(x_k) + w_k \tag{1}$$

where X_k is the state vector at time k, and w_k is the model error with covariance Q = $E[w\, w^T]$.

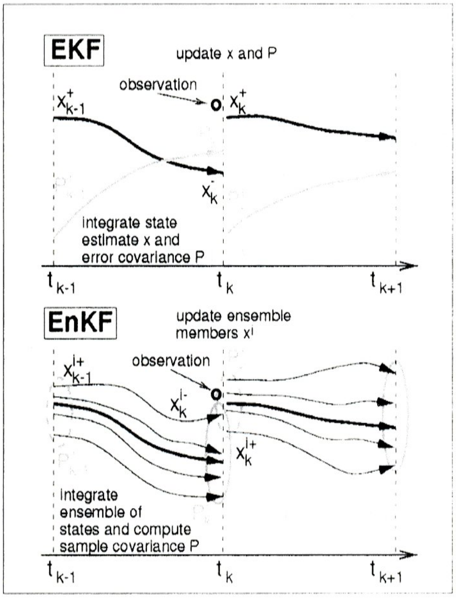


Figure 2: Schematic of the extended Kalman filter (EKF) and the ensemble Kalman filter (EnKF). The extended Kalman filter approximates the error covariance propagation by linearising the land surface model. The ensemble Kalman filter nonlinearly propagates an ensemble from which sample covariances are derived at the update time.

Both the extended Kalman filter and the ensemble Kalman filter work sequentially from one measurement time to the next, applying in turn a forecast step and an update step. Figure 2 illustrates the difference between the ensemble Kalman filter and the extended Kalman filter. During the forecast step, the extended Kalman filter propagates a single estimate of the state vector (from x^+_{k-1} to x^-_k), and integrates the uncertainty (error covariance) (from P^+_{k-1} to P^-_k) of that state, which will be used in the model forecast at the update step. The ensemble Kalman filter simultaneously propagates an ensemble of state vectors each state vector representing a particular realization of the possible model trajectories, and the error covariance is computed from the distribution of the model states in the ensemble.

Using superscripts – and + to refer to the state estimates, individual ensemble members or covariances before and after the update step, respectively, the state variables, covariances during forecast and update steps are expressed as follows:

State forecast:

extended Kalman filter: $X_{k+1}^{-} = f_k(x_k^{+})$ (2)

ensemble Kalman filter: $X_{k+1}^{i-} = f_k(x_k^{i+}) + w_k^{i}$ (3)

Covariance forecast:

extended Kalman filter:

$$P_{k+1}^{-} = F_k P_k^{+} F_k^{T} + Q_k \tag{4}$$

$$[F_k]_{mn} = \frac{\partial f_m}{\partial x_n}\Big|_{x_k^{-}} \tag{5}$$

ensemble Kalman filter

$$P_{k+1}^{-} = \frac{1}{N-1} D_{k+1} D_{k+1}^{T} \tag{6}$$

$$D_{k+1} = [x_{k+1}^{1-} - x_{k+1}^{-}, \ldots\ldots, x_{k+1}^{N-} - x_{k+1}^{-}] \tag{7}$$

$$x_{k+1}^{=} = \frac{1}{N \sum_{i=1}^{N} x_{k+1}^{i-}} \tag{8}$$

During the update step, the observation vector Y is linearly related to the system state vector X and the state independent terms XS_0,

$$y_k = H_k x_k + xs_{0k} + v_k \tag{9}$$

where v_k is the measurement error with covariance $R=E[v\, v^T]$.

During the update step, the extended Kalman filter revises its estimate of the state (from x_k^{-} to x_k^{+}) using the observation and the prognostic state error covariance P_k^{-} and the state error covariance is also updated (from P_k^{-} to P_k^{+}). On the other hand, the ensemble Kalman filter updates each ensemble members separately, using the observations and the diagnosed state error covariance P_k^{-}.

During the update step the Kalman gain weights the observations against the model forecast. Its weighting is determined by the relative magnitude of model uncertainty with respect to the observation covariance. The Kalman gain is given by

$$K_k = P_k^- H_k^T [H_k P_k^- H_k^T + R_k]^{-1} \tag{10}$$

If no observations are available at time k we set K_k=0. Next, the state estimate by the extended Kalman filter or each ensemble member by ensemble Kalman filter is updated using a linear combination of forecast model states and the observations:

Update:

extended Kalman filter: $$x_k^+ = x_k^- + K_k[y_k - H_k x_k^- - xk_k] \tag{11}$$

$$P_K^+ = P_k^- - K_K H_k P_k^- \tag{12}$$

ensemble Kalman filter: $$x_k^{i+} = x_k^{i-} + K_k[y_k - H_k x_k^{i-} - xs_k + v_k^i] \tag{13}$$

It is also important to continuously reevaluate the assumptions and problems that exist in the modelling and assimilation algorithms, so as to not attribute errors to the incorrect error source. A few considerations are outlined below.

As illustrated by Hollingsworth *et al.* (1986), assumptions on the bias and horizontal correlation structure of the model and observations can have a significant impact on error estimation. In practice, data assimilation is often implemented with the assumption that observations and predictions are unbiased and uncorrelated in space. These assumptions work reasonably well for *in situ* observations, but satellite observations are usually biased by use of inaccurate algorithms, and their errors are usually horizontally correlated because the same sensor is making all the observations.

In cooperation with the National Aeronautics and Space Administration's Seasonal-to-Interannual Prediction Project, we have evaluated the benefits and drawbacks of using a 1-dimensional (vertical) or a 3-dimensional (vertical and horizontal) Kalman filter. At the horizontal scale of the Advanced Microwave Sounding Unit (60 kilometres), soil moisture and snow exhibit little spatial correlation. Further, land surface models have little or no explicit treatment of the physical processes associated with these horizontal scales. 3-dimensional Kalman filtering is also much more computationally intensive than 1-dimensional Kalman filtering. However, there is potential benefit for 3-dimensional assimilation to: (1) extend observations into data-sparse regions, and (2) be used with higher resolution observations that exhibit more spatial correlation.

The potential benefits and drawbacks of using direct radiance assimilation are also being explored. It is possible to design a data assimilation system that assimilates radiances directly, rather than derived quantities such as retrievals, by including a forward model in the assimilation. Such a system is under development in our effort to derive soil moisture from the Tropical Rainfall Measuring Mission's Microwave Imager (TRMM-TMI), and may allow for more precise identification of calibration problems in the Advanced Microwave Sounding Unit.

3. Streamflow assimilation

While altimetry data could be used to provide additional observation of streamflow where the rivers are sufficiently wide (greater than 250 metres) and a stage discharge relationship is available, altimetry data is most likely to be used to update lake/reservoir storage. In this situation we can assimilate the lake/reservoir level and hence storage. This approach requires the existence of volume-elevation relationships depending on the bathymetry of the water body.

The first key to the implementation of a successful streamflow assimilation scheme for correction of the land surface states with the Kalman filter is the adequate specification of the model land surface state error covariance matrix. The important factor is that the model system state error covariance matrix correctly identifies the cross correlation between the soil moisture and snow prognostic variables. Likewise, successful runoff assimilation with the Kalman filter for correction of the runoff forecasts upstream of the streamflow gauging station will depend on adequate specification of the model runoff error covariance matrix. However, the model runoff error covariance matrix is not required for updating the model land surface states.

The variance of the observations can be identified reliably in most cases, since it depends on the characteristics of the measuring device (Georgakakos and Smith, 1990). Observation error estimates are not available with the streamflow data, but with knowledge of the gauging station instrumentation, these error estimates may be readily obtained.

The second key to implementation of a successful streamflow assimilation scheme for the correction of the land surface states with the Kalman filter is the adequate specification of the observation equation. Since runoff is a land surface model flux, the relationship between this flux and the model states, such as soil moisture and snow, must be established. The difficulty with this approach is the time lag between runoff from the source and runoff observed at the gauging station. Hence, this will be the single most difficult component of the runoff assimilation scheme. To ensure that observations are related back to the correct time for each land surface states may require continuous or near-continuous assimilation of runoff data and identification of the time lags associated with the runoff from each catchment.

An effective evaluation of any large-scale modelling endeavour is the most difficult and yet most important aspect. We evaluate the assimilation of runoff data through comparison of model derived surface soil moisture and snow states with remotely sensed surface soil moisture and snow observations. The experimental design involves synthetic studies where model output of soil moisture, snow, and runoff is used for the observation and evaluation data, followed by experiments where these fields are replaced by actual observations. The initial studies are for an individual catchment where runoff at the catchment outlet will be assimilated to correct soil moisture and snow forecasts, evaluating the within-catchment flow routing component. This is followed by the assimilation of runoff for a small number of catchments without any reservoir storage to evaluate the assimilation scheme when the between-catchment routing is included. The final stage is to evaluate the assimilation when there is a reservoir (or reservoirs) included in the runoff network.

4. Soil moisture assimilation

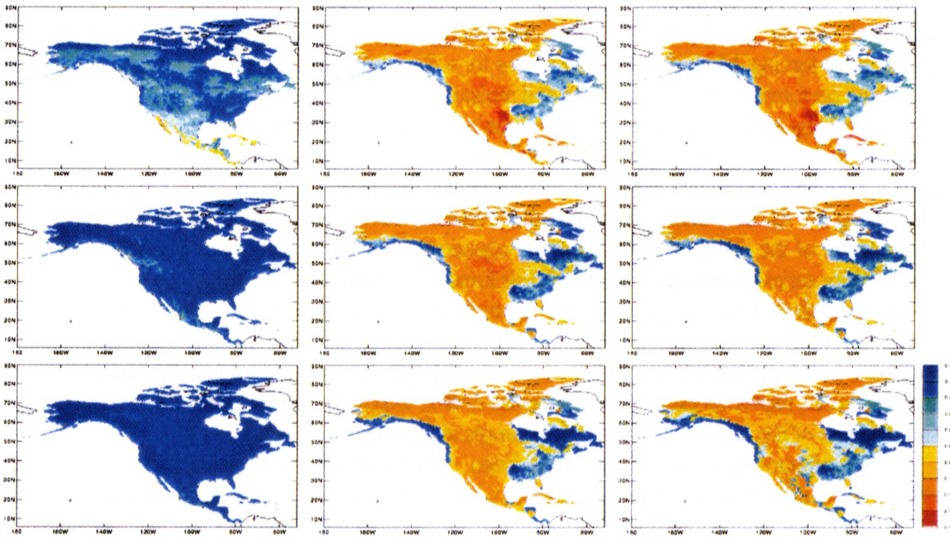
A Kalman filter soil moisture assimilation strategy has been developed (Walker and Houser, 2001). The principal advantage of this approach is that the Kalman filter provides a framework within which the entire system is modified, with covariances representing the reliability of the observations and model prediction. We have used a one-dimensional Kalman filter for updating the soil moisture prognostic variables of the Koster *et al.* (2000) catchment-based land surface model. A one-dimensional Kalman-filter was used because of its computational efficiency and the fact that horizontal correlations between soil moisture prognostic variables of adjacent catchments at the scales of interest to climate modelling are likely only through the large-scale correlation of atmospheric forcing. Moreover, all calculations for soil moisture in the catchment-based land surface model are performed independent of the soil moisture in adjacent catchments.

Forecasting of the soil moisture covariance matrix using Kalman filter theory requires a linear forecast model. However, forecasting of the soil moisture prognostic variables (surface excess, root zone excess and catchment deficit) in the catchment-based land surface model is non-linear. Hence, 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.

In order to perform an update of the soil moisture prognostic variables with the Kalman filter, the observations (near-surface soil moisture) must be linearly related to the soil moisture prognostic variables. In the catchment-based land surface model, the soil moisture prognostic variables are the surface excess, root zone excess, and catchment deficit, which are related to the observed volumetric soil moisture of the surface layer through a complicated non-linear function.

A set of numerical experiments have been undertaken for North America to illustrate the effectiveness of the Kalman filter assimilation scheme in providing a more accurate estimate of the soil moisture storage throughout the entire soil profile (Figure 3). Moreover, the corresponding positive influence on the water balance components, namely evapotranspiration and runoff, has been investigated. In this experiment, atmospheric forcing data and soil and vegetation properties from the first International Satellite Land-Surface Climatology Project (Sellers *et al.*, 1996) have been used as model input for the year 1987.

Using the 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 North America was forecast for 1987. The forecasts of near-surface soil moisture were output every 3 days to represent the near-surface soil moisture measurements from remote sensors. In addition to soil moisture, the land surface model provided estimates of evapotranspiration and runoff for each of the catchments. This simulation provided the "true" soil moisture and water balance data for comparison with degraded simulations. Moreover, it allowed evaluation of the effectiveness of assimilating near-surface soil moisture data for improving the land surface model forecast of soil moisture and water budget components, when initialized with poor soil moisture initial conditions. In the degraded simulation, the initial conditions for the soil moisture prognostic variables from the spin-up were set to arbitrarily wet values uniformly across all of North America. The land surface model was then forced with the same atmospheric data as in the "truth" simulation. The wet initial condition causes over-estimation of evapotranspiration and runoff. The final simulation was to assimilate the near-surface "observations" from the "truth" simulation into the degraded simulation every 3 days. The effect of assimilation on the soil moisture forecasts can be seen in Figure 3. These results show that after only 1 month of assimilation, the "true" soil moisture has been retrieved for the majority of North America.



a b c

Figure 3: Comparison of gravimetric soil moisture on 30 January 1987 in near-surface (top row), root zone (middle row) and entire profile (bottom row) from: (a; left column) simulation with degraded initial conditions for soil moisture; (b; middle column) simulation with spin-up initial conditions ("truth"); and (c; right column) degraded simulation with assimilation of near-surface soil moisture from the "truth" simulation once every 3 days.

5. Snow assimilation

Snow plays an important role in governing both the global energy and water budgets, due to its high albedo, thermal properties, and being a medium-term water store. However, the problem of accurately forecasting snow in regional and global atmospheric and hydrologic models is difficult, as a result of subgrid-scale variability of snow, and errors in the model forcing data. Hence, any land surface model snow initialization based on model spin-up will be affected by these errors. By assimilating snow observation products into the land surface model, a best estimate of snow states may be obtained and model bias can be corrected. We are implementing a snow assimilation scheme that optimally merges remotely-sensed snow observations with the catchment-based land surface model forecast. As a first step, identical twin experiments have been performed to test and validate a snow data assimilation scheme. Synthetic observations of snow water equivalent are assimilated and other snow states are subsequently reanalyzed using the updated snow water equivalent. Preliminary results show good agreement between the assimilation and simulated truth. Figure 4 shows snapshots of truth, assimilation and forecast results of 24 continuous catchments in North America on March 16 1987. The assimilation starts from January 1, 1987, with a poor initial condition that assumes no snow is present anywhere. It produces satisfactory estimates of snow water equivalent, snow depth and snow temperature, while the model forecast with the same poor initial condition produces very different states. In the near future, snow water equivalent estimated from the Scanning Multichannel Microwave Radiometer (SMMR) and the Special Sensor Microwave/Imager (SSM/I) will be attempted..

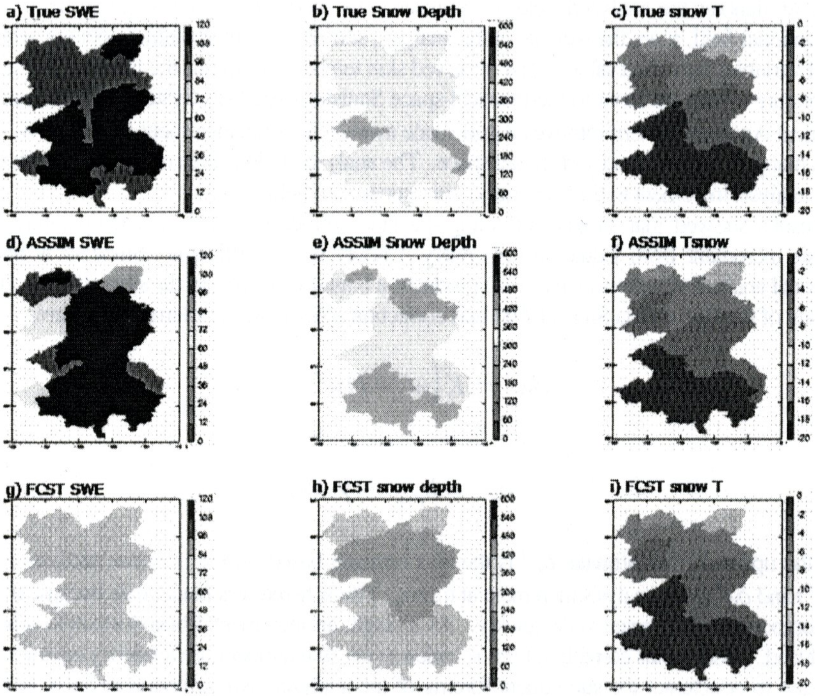


Figure 4: Snapshot of truth, assimilation and forecast (from poor initial condition) on 16 March 1987 from 3-month assimilation starting from 1 January 1987. Top-row: a) to c). Middle row: d) to f). Bottom row: g) to i). Left column: snow water equivalent (millimetres); Middle column: snow depth (millimetres); Right column: temperature (C).

6. Skin temperature assimilation

The land surface skin temperature state is a principle control on land-atmosphere fluxes of water and energy, is closely related to soil water states, and is easily observable from space and aircraft infrared sensors in cloud-free conditions. The usefulness of skin temperature in land data assimilation studies is limited by its very short memory (on the order of minutes) due to the very small heat storage it represents. We used the Physical-space Statistical Analysis System (PSAS, Cohn *et al.*, 1998) in a 2.5 degrees longitude by 2.0 degrees latitude global land surface model to assimilate surface skin temperature observations from International Satellite Cloud Climatology Project (ISCCP). The Physical-space Statistical Analysis System algorithm obtains the best estimate of the state of the system by combining observations with the forecast model first guess. The analysis equation, which encapsulates the Physical-space Statistical Analysis System scheme, is

$$w^a = w^f + K\left(w^o - Hw^f\right) \tag{14}$$

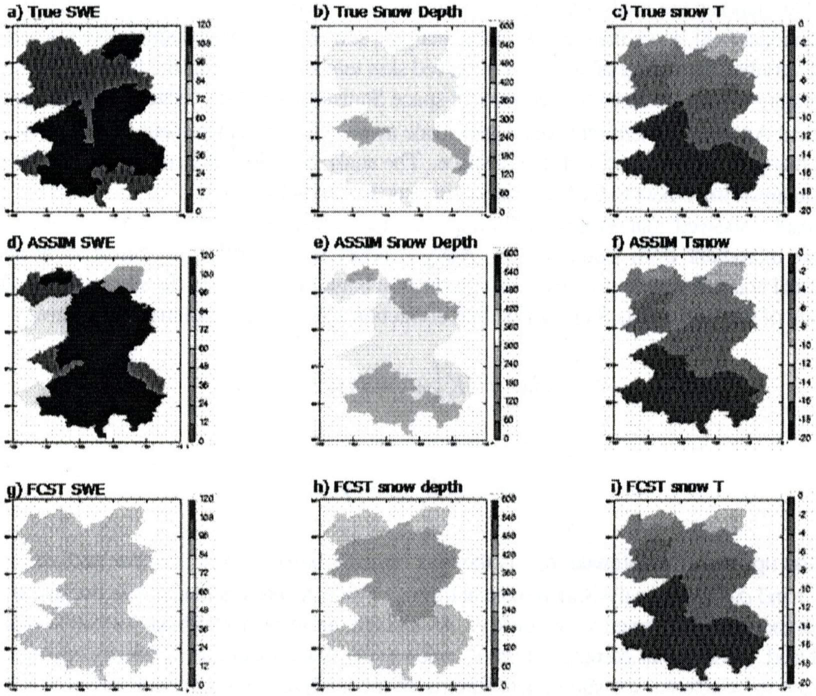


Figure 4: Snapshot of truth, assimilation and forecast (from poor initial condition) on 16 March 1987 from 3-month assimilation starting from 1 January 1987. Top-row: a) to c). Middle row: d) to f). Bottom row: g) to i). Left column: snow water equivalent (millimetres); Middle column: snow depth (millimetres); Right column: temperature (C).

6. Skin temperature assimilation

The land surface skin temperature state is a principle control on land-atmosphere fluxes of water and energy, is closely related to soil water states, and is easily observable from space and aircraft infrared sensors in cloud-free conditions. The usefulness of skin temperature in land data assimilation studies is limited by its very short memory (on the order of minutes) due to the very small heat storage it represents. We used the Physical-space Statistical Analysis System (PSAS, Cohn *et al.*, 1998) in a 2.5 degrees longitude by 2.0 degrees latitude global land surface model to assimilate surface skin temperature observations from International Satellite Cloud Climatology Project (ISCCP). The Physical-space Statistical Analysis System algorithm obtains the best estimate of the state of the system by combining observations with the forecast model first guess. The analysis equation, which encapsulates the Physical-space Statistical Analysis System scheme, is

$$w^a = w^f + K\left(w^o - Hw^f\right) \tag{14}$$

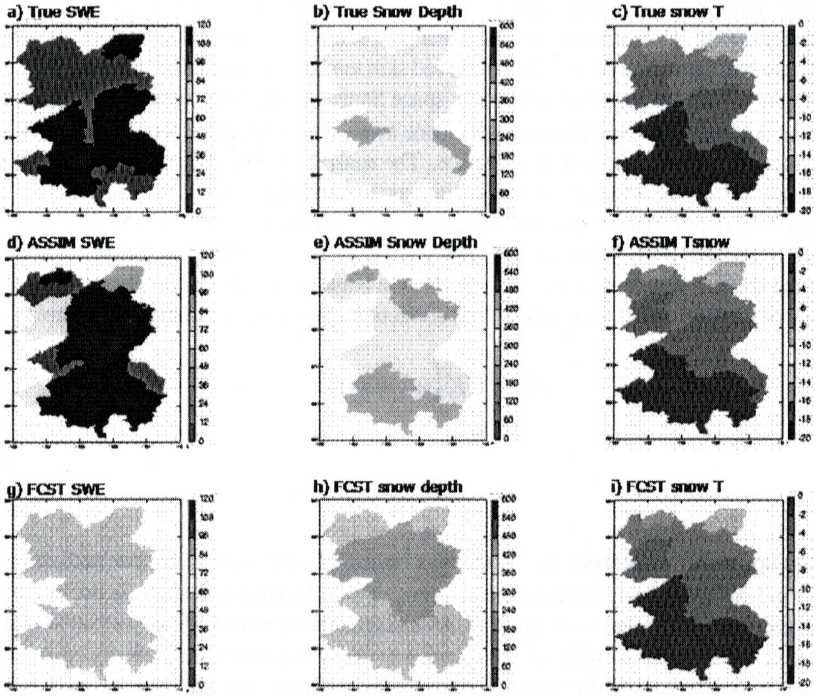


Figure 4: Snapshot of truth, assimilation and forecast (from poor initial condition) on 16 March 1987 from 3-month assimilation starting from 1 January 1987. Top-row: a) to c). Middle row: d) to f). Bottom row: g) to i). Left column: snow water equivalent (millimetres); Middle column: snow depth (millimetres); Right column: temperature (C).

6. Skin temperature assimilation

The land surface skin temperature state is a principle control on land-atmosphere fluxes of water and energy, is closely related to soil water states, and is easily observable from space and aircraft infrared sensors in cloud-free conditions. The usefulness of skin temperature in land data assimilation studies is limited by its very short memory (on the order of minutes) due to the very small heat storage it represents. We used the Physical-space Statistical Analysis System (PSAS, Cohn *et al.*, 1998) in a 2.5 degrees longitude by 2.0 degrees latitude global land surface model to assimilate surface skin temperature observations from International Satellite Cloud Climatology Project (ISCCP). The Physical-space Statistical Analysis System algorithm obtains the best estimate of the state of the system by combining observations with the forecast model first guess. The analysis equation, which encapsulates the Physical-space Statistical Analysis System scheme, is

$$w^a = w^f + K\left(w^o - Hw^f\right) \tag{14}$$

where w^a denotes the analyzed field, w^f represents the model forecast first guess field, w^o is the observational field, K are the weights of the analysis, and H is the interpolation operator which maps model variables into observables. The observed skin temperature minus the forecast first guess skin temperature values are input to the Physical-space Statistical Analysis System. The Physical-space Statistical Analysis System retrieves a grid space average analysis increment ($\delta w^a = K(w^o - H\,w^f)$), that is mapped into the land surface tile space. The analyzed field is then obtained by adding the tile space analysis increment to the first guess skin temperature field.

Results showed that simply correcting the land surface modeled skin temperature with the analysis increment from Equation (14) every 3 hours was insufficient. Since w^f is biased, the traditional analysis equation such as (14) produces a biased w^a (Dee and da Silva, 1998). Therefore, a variant of the Dee and da Silva (1998) bias correction scheme was implemented where,

$$\delta w^a = K\left(w^o - Hw^f + b^f\right) \tag{15}$$

$$w^a = w^f - b_{k-1}{}^f + \delta w^a \tag{16}$$

$$b_k{}^f = b_{k-1}{}^f - \gamma \cdot \delta w^a \tag{17}$$

b_k^f is the updated bias estimate, b_{k-1}^f is the bias estimate based on the previous analysis increment ($\delta w_{k-1}{}^a$) and δw^a is the analysis increment at time t_k. This scheme was inadequate because the surface skin temperature bias acted very quickly. As a result, an incremental bias correction scheme was introduced, where a bias correction term is added to the skin temperature tendency equation at every time step to counteract the subsequent forcing of the analyzed skin temperature back to the initial state. For this scheme, b^f is computed as in Equation (17) and the bias correction term is calculated as,

$$f_b = \frac{b^f}{\tau} \tag{18}$$

where $\tau = 3$ hours is the frequency of the International Satellite Cloud Climatology Project dataset, *i.e.*, the frequency of assimilation. This scheme effectively removed the time mean bias, but did not remove the bias in the mean diurnal cycle. To account for this deficiency, we modelled the time-dependent bias as,

$$b^f(t) = \sum_j (a_j \cos\omega_j t + b_j \sin\omega_j t) \tag{19}$$

and estimated the Fourier coefficients

$$a_j = a_j - \gamma\delta w^a \cos\omega_j t \tag{20}$$

$$b_j = b_j - \gamma\delta w^a \sin\omega_j t \tag{21}$$

We determined that to adequately account for the diurnal bias changes it is necessary to keep diurnal ($\omega_1 = 2\pi/24$h) and semi-diurnal ($\omega_2 = 2\pi/12$h) harmonics.

The results presented here are based on the evaluation of the discussed techniques for a July 1992 International Satellite Cloud Climatology Project surface skin temperature dataset. The test runs (Table 1) include the simulation without assimilation or bias correction (Model), with Physical-space Statistical Analysis System temperature assimilation (Assimilation I; Equation 14), with bias correction (Assimilation II; Equations 15-17), with incremental bias correction (Assimilation III; Equation 18), with diurnal bias correction (Assimilation IV; Equations 19-21) and with semi-diurnal bias correction (Assimilation V). Figure 5 shows the July 1992 global monthly mean standard deviations of surface skin temperature between the experiments and the observations. The standard deviation decreases gradually with each successive improvement to the methodology, and therefore substantiates the techniques developed. However, the monthly mean standard deviation does not reveal the more visible impact of the diurnal bias correction on the monthly mean diurnal cycle.

Table 1: Description of experiments.

Experiment	Description
Model	No assimilation
Assimilation I	Physical-space Statistical Analysis System assimilation
Assimilation II	Physical-space Statistical Analysis System with bias correction every 3 hours
Assimilation III	Physical-space Statistical Analysis System with incremental bias correction
Assimilation IV	Physical-space Statistical Analysis System with diurnal bias correction
Assimilation V	Physical-space Statistical Analysis System with semi-diurnal bias correction

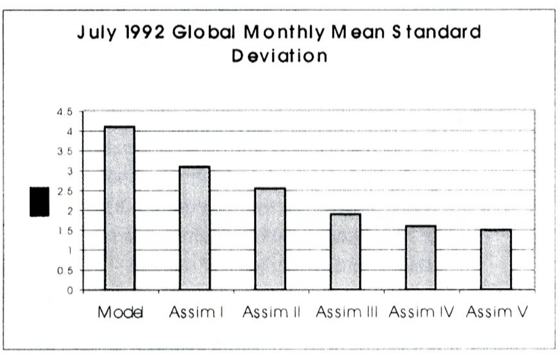


Figure 5: The July 1992 global monthly mean standard deviations of surface skin temperature between the assimilation experiments and the International Satellite Cloud Climatology Project observations (see Table 1).

The July 1992 monthly mean diurnal cycle of surface skin temperature over North America for the International Satellite Cloud Climatology Project observations, Model, Assimilation IV, and Assimilation V, are presented at the top of Figure 6. The effectiveness of implementing semi-diurnal bias correction is shown by how closely Assimilation V matches the observations. Two-metre temperature (middle) and specific humidity (bottom), also displayed in Figure 6, reveal that the inclusion of the bias correction scheme also impacts the surface meteorology fields. Thus, for a decrease in surface skin temperature, due to the bias correction, there is a corresponding decrease in the 2 metre temperature and specific humidity. Figure 6 allows only for model intercomparison, and we are in the process of obtaining a verification dataset.

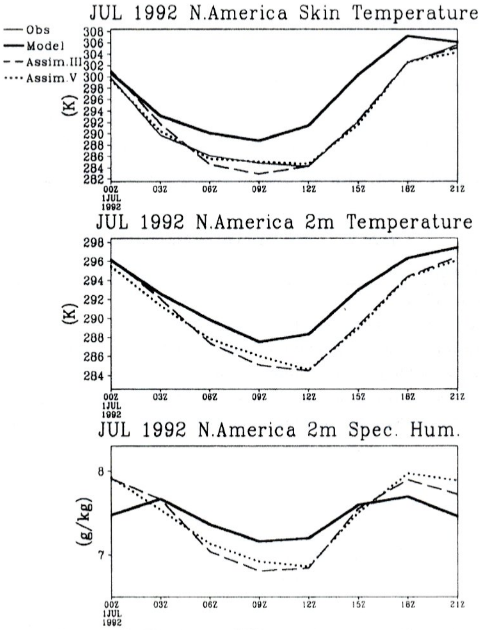


Figure 6. The July 1992 monthly mean diurnal cycle of skin temperature (top), 2 m temperature (middle) and 2 m specific humidity (bottom) over North America for the observations (light solid), Model (heavy solid), Assimilation III (dashed) and Assimilation V (dotted).

Similarly, the same corrective effect is visible in the Western Europe surface skin temperature for Assimilation V. The sensible heat flux and latent heat flux also show that the bias correction technique has a substantial impact on the energy budget, where the reduction in skin temperature causes a decrease in the sensible and latent heat flux.

In this study, the Mosaic land model (Koster and Suarez, 1992, 1996) has been forced with near surface atmospheric conditions derived by the Goddard Earth Observing System Data Assimilation System (GEOS-DAS). The Physical-space Statistical Analysis System was used with the Mosaic land model in order to assimilate three hourly International Satellite Cloud Climatology Project surface skin temperature data. Bias correction techniques were developed, since traditional analysis with Physical-space Statistical Analysis System of a biased forecast lead to a biased analysis. The bias correction algorithms that were evaluated included bias correction every 3 hours, incremental bias correction every time step, and bias correction to the mean diurnal and mean semi-diurnal cycle. The results for a July 1992 test case have shown that the semi-diurnal bias correction was most effective. The monthly mean diurnal cycle from the semi-diurnal bias correction experiment closely matched the diurnal cycle from the observations. Also, the semi-diurnal bias correction results show the lowest standard deviation for the global monthly mean between the experiment and the observations.

References

Charney, J. G., M. Halem, and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160-1163.

Cohn, S., 1982: Methods of Sequential Estimation for Determining Initial Data in Numerical Weather Prediction. Ph.D. thesis, Courant Institute of Mathematical Sciences, New York University, New York, NY, 183 pp.

Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich, 1998: Assessing the effects of data selection with the DAO Physical-space Statistical Analysis System. *Mon. Weather Rev.*, **126**, 2913-2926.

Dee, D. P., and A. da Silva, 1998: Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* **124**, 269-295.

Entekhabi, D., H. Nakamura, and E. G. Njoku, 1994: Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Trans. Geosci. Remote Sensing*, **32**, 438-448.

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

Georgakakos, K.P., and Smith, G.F., 1990: On improved hydrologic forecasting - Result from a WMO real time forecasting experiment. *J. Hydrol.*, **114**, 17-45.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: *Application of estimation theory to numerical weather prediction. Dynamic Meteorology: Data Assimilation Methods*, Springer-Verlag, New York, NY, 139-224.

Hollingsworth, A., D. Shaw, P. Lonnberg, L. Illari, K. Arpe, and A. J. Simmons, 1986: Monitoring of Observation and Analysis Quality by a Data Assimilation System. *Mon. Weather Rev.*, **114**, 861-879.

Houser, P. R., W. J. Shuttleworth, H. V. Gupta, J. S. Famiglietti, K. H. Syed, and D. C. Goodrich, 1998: Integration of Soil Moisture Remote Sensing and Hydrologic Modeling using Data Assimilation. *Wat. Resour. Res.*, **34**, 3405-3420.

Houtekamer, P. L. and H. L, Mitchell, 1998: Data assimilation using a Ensemble Kalman filter techniques. *Mon. Weather Rev.*, **126**, 796-811.

Kalman, R.E., 1960: A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D, J. Basic En.g.* **82**, 35-45.

Keppenne, C.L., and M.M. Rienecker, 2001: Development and initial testing of a parallel Ensemble Kalman filter for the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.*, submitted.

Koster, R. D., M. J. Suarez, and M. Heiser, 2000: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. of Hydrometeorology*, **1**, 26-46.

Koster, R. D., and M. J. Suarez, 1992: Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *J. Geophys. Res.*, **97**, 2697-2715.

Koster, R.D., M.J. Suarez, A. Duchame, M. Stiglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a GCM. Part 1: Model structure. *J. Geophys. Res.*, **105**, 24809-24822.

Lorenc, A. C., 1995: Atmospheric Data Assimilation. *Meteorological Office Forecasting Research Div.*, **34**, The Met Office, UK.

McLaughlin, D., 1995: Recent developments in hydrologic data assimilation. *Reviews of Geophysics*, 977-984.

Milly, P. C. D., 1986: Integrated remote sensing modeling of soil moisture: sampling frequency, response time, and accuracy of estimates. *Integrated Design of Hydrological Networks - Proceedings of the Budapest Symposium*, **158**, 201-211.

Reichle, R.H., J.P. Walker, and R.D. Koster, and P.R. Houser, 2002: Extended versus Ensemble Kalman Filtering for Land Data Assimilation. J. *Hydromet.*, **3**, 728-740.

Sellers, P. J., B. W. Meeson, J. Closs, J. Collatz, F. Corprew, D. Dazlich, F. G. Hall, Y. Kerr, R. Koster, S. Los, K. Mitchell, J. McManus, D. Myers, K.-J. Sun, and P. Try, 1996: The ISLSCP Initiative I global datasets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies. *Bull. Amer. Meteorol. Soc.*, **77**, 1987-2005.

Walker, J., 1999: *Estimating Soil Moisture Profile Dynamics From Near-Surface Soil Moisture Measurements and Standard Meteorological Data.* Dissertation Thesis, Dept. of Civil, Surveying and Environmental Engineering, The University of Newcastle.

Walker, J. P., and P. R. Houser, 2001: A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *J. Geophys. Res.*, **106**, 11761-11774.