Land Data Assimilation

Paul R. Houser

National Aeronautics and Space Administration's Goddard Space Flight Center
Greenbelt, Maryland 20771, USA
Paul.Houser@nasa.gov

1. Introduction

Accurate initialization of land surface water and energy stores is critical in environmental prediction because of their regulation of land-atmosphere fluxes over a variety of time scales. Errors in land surface forcing and parameterization accumulate in these integrated land stores leading to incorrect surface water and energy partitioning. However, many new land surface observations are becoming available that may be used to constrain the dynamics of these states. These constraints can be imposed by (1) forcing the land surface primarily by observations, thereby avoiding the often severe numerical weather prediction biases, and (2) using data assimilation techniques to constrain unrealistic storage dynamics. This is the goal underlying the Land Data Assimilation Systems conceptual framework which aims to develop the best estimation of the current state of land surfaces through an best possible integration of land surface observation and simulation.

Significant progress has been made in land-surface observation and modeling at a wide range of scales. Projects such as the International Satellite Land Surface Climatology Project, the Global Soil Wetness Project, and the Global Energy and Water Cycle Experiment, among others have paved the way for the development of an operational Land Data Assimilation System. Several Land Data Assimilation Systems have been implemented in near real time and at high spatial resolution for North American, European, and global domains. These Land Data Assimilation Systems are forced with real time output from numerical prediction models, satellite data, and radar precipitation measurements, and can incorporate land state observations as a constraint to the model dynamics using hydrologic data assimilation methods. Results of Land Data Assimilation System assimilation of land surface temperature, moisture, and snow are showing great promise to improve predictability and understanding of model realism.

2. Land Surface Modeling

Recent advances in understanding soil-water dynamics, plant physiology, micrometeorology, and hydrology, all of which control biosphere-atmosphere interactions, have spurred the development of land surface models (Figure 1). The primary goal of a land surface model is to represent in a simple, yet realistic way, the transfer of mass, energy, and momentum between a vegetated surface and the atmosphere [Dickinson et al., 1993; Sellers et al., 1986]. Land surface model predictions are regular in time and space, but these predictions are influenced by model structure, errors in input variables and model parameters, and inadequate treatment of sub-grid scale spatial variability. Consequently, land surface model predictions of land surface hydrology and land surface states are much improved by the assimilation of land surface observations.

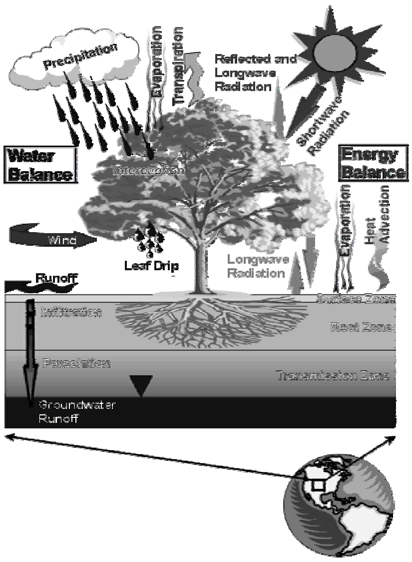


Fig 1: Land surface modeled processes.

3. Remote sensing of the land surface

The emphasis of land surface data assimilation research is to assimilate remotely-sensed observations of the land surface that previous research suggests will provide memory to the land-atmosphere interaction. Remote observations of interest include: (1) temperature, (2) soil moisture (surface moisture content, surface saturation, total water storage), (3) other surface water bodies (lakes, wetlands, and large rivers) and (4) snow (areal extent, snow water equivalent). The remote sensing potential and availability of each of these quantities is described in more detail below.

Remote sensing of surface temperature is a relatively mature technology. The land surface emits thermal infrared radiation at an intensity directly related to its emissivity and temperature. The absorption of this radiation by atmospheric constituents is smallest in the 3 to 5 and 8 to 14 micrometer wavelength ranges, making them the best atmospheric windows for sensing land surface temperature. Some errors due to atmospheric absorption and improperly specified surface emissivity are possible, and the presence of clouds can obscure the signal. Generally, surface temperature remote sensing can be considered an operational technology, with many spaceborne sensors making regular observations (for example, the Landsat Thematic Mapper, Advance Very High Resolution Radiometer, the Moderate Resolution Imaging Spectroradiometer, and the Advanced Spaceborne Thermal Emission and Reflection Radiometer [Lillesand and Kiefer, 1994]). The evolution of land surface temperature is linked to all other land surface processes through physical relationships. These land surface process interconnections can be exploited in a data assimilation framework to constrain all of the predicted land surface states.

Remote sensing of soil moisture content is a developing technology, although the theory and methods are well established [Eley, 1992]. Long-wave passive microwave remote sensing is ideal for soil moisture observation, but there are technical challenges involved in correcting for the effects of vegetation and roughness. Soil moisture remote sensing has previously been limited to aircraft campaigns [e.g. Jackson, 1997a], or analysis of the Defense Meteorological Satellite Program Special Sensor Microwave Imager [Engman, 1995; Jackson, 1997b]. The Special Sensor Microwave Imager has also been successfully employed to monitor surface saturation/inundation [Achutuni and Scofield, 1997; Basist and Grody, 1997]. The Earth Observing System Advanced Microwave Sounding Unit will provide additional C-band microwave observations that may be useful for soil moisture determination. The Tropical Rainfall Measuring Mission's Microwave Imager, which is very similar to the Advanced Microwave Sounding Unit, is much better suited to soil moisture measurement (because of its 10 megahertz channels) than Special Sensor Microwave Imager. All of these sensors have adequate spatial resolution for land surface applications, but have a very limited quantitative measurement capability , especially over dense vegetation. However, Sipple et al., [1994] demonstrated that it is possible to determine saturated areas through dense vegetation using the Scanning Multichannel Microwave Radiometer, which can greatly aid land surface predictions. Because of the large error in remotely-sensed microwave observations of soil moisture, there is a real need to maximize its information by using algorithms that can account for its error and that extend its information in time and space.

An important and emerging technology with respect to land surface observation is the potential to monitor variations in total water storage (ground water, soil water, surface waters (lakes, wetlands, rivers), water stored in vegetation, snow and ice) using satellite observations of the time variable gravity field. The Gravity Recovery and Climate Experiment, an Earth System Science Pathfinder mission launched in 2002, will provide highly accurate estimates of changes in terrestrial water storage in large. Wahr et al. [1998] note that the Gravity Recovery and Climate Experiment will provide estimates of variations in water storage to within 5 millimeters on a monthly basis. Rodell and Famiglietti [1999] have demonstrated the potential utility of these data for hydrologic applications, including their application in large (>150,000 km2) watersheds; and they further discuss the potential power of Gravity Recovery and Climate Experiment observations for constraining modeled water storage in land surface models when combined with surface soil moisture and altimetry observations. Birkett [1995, 1998] demonstrated the potential of satellite radar altimeters to monitor height variations over inland waters, including climatically-sensitive lakes and large rivers and wetlands. Such altimeters are currently operational on the European Space Agency Remote Sensing Satellite 2, the Topex Poseidon satellite, the European Space Agency Environment Satellite, and the Jason 1 satellite.

Key snow variables of interest to land surface understanding include area coverage and snow water equivalent. While the estimation of snow water equivalent by satellite is currently in research mode, snow areal extent can be routinely monitored by many operational platforms, including The Advanced Very High Resolution Radiometer, the Geostationary Operational Environmental Satellite and the Special Sensor Microwave Imager. Recent algorithm developments even permit the determination of the fraction of snow cover within Landsat Thematic Mapper pixels [Rosenthal and Dozier, 1996]. Cline et al. [1998], describe an approach to retrieve snow water equivalent from the joint use of remote sensing and energy balance modeling.

Precipitation is the most important forcing for the land surface. Two general categories of satellite-derived precipitation exist, each with severe limitations. The first category of satellite-based precipitation observation is Geostationary Operational Environmental Satellite Precipitation Index estimates [Arkin and Meisner, 1987]. This very simple method uses the cloud infrared brightness to directly estimate precipitation using a lookup table. This method can provide hourly precipitation estimates, but is limited to convective precipitation structures in the 40 degrees North to 40 degrees South latitude band. The second category uses shortwave passive microwave, as available with the Special Sensor Microwave Imager instrument, the Tropical Rainfall Measurement Mission Microwave Imager, and the Advanced Microwave Scanning Radiometer are sensitive to cloud water vapor quantities and raindrops, and can therefore provide better estimates of precipitation. Because these satellites are not geostationary, their temporal coverage is limited. Many research groups have investigated the derivation of precipitation from this data using methods ranging from simple empirical systems to neural network techniques. The quality of precipitation estimates is expected to be highest from microwave sensors, moderate from Geostationary Operational Environmental Satellite Precipitation Index estimates, and lowest from the numerical model predictions. Generally, the best available precipitation observations are used, or the available observations are optimally merged following Houser *et al.* [1999]. Additional corrections for each data type based on rain gauge and climatological information further increase the accuracy of remotely observed precipitation.

4. Assimilation of land surface data

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 2). 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 kilometers, 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.

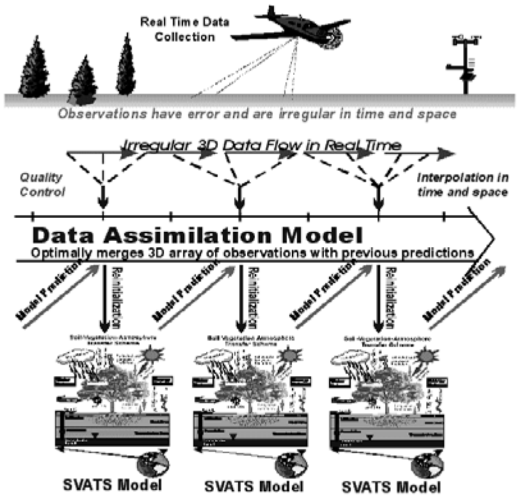


Figure 2: *The land surface data assimilation process.*

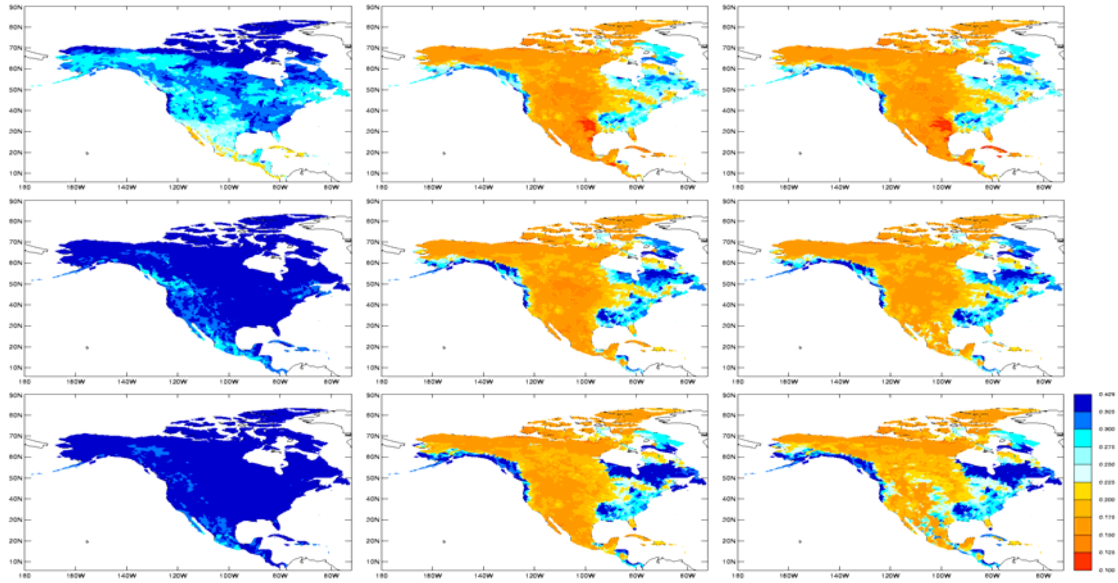
The feasibility of land surface data assimilation methods has been recently tested by the hydrological and numerical weather prediction research communities. This research described here focuses on: (1) the use of a one-dimensional Kalman filtering based land assimilation strategy that expands upon the 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].

5. 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. A one-dimensional Kalman filter was used 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 modeling 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.



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) simulation with degraded initial conditions for soil moisture; (b) simulation with spin-up initial conditions ("truth"); and (c) degraded simulation with assimilation of near-surface soil moisture from the "truth" simulation once every 3 days.

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.

6. 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 and the Special Sensor Microwave/Imager will be attempted..

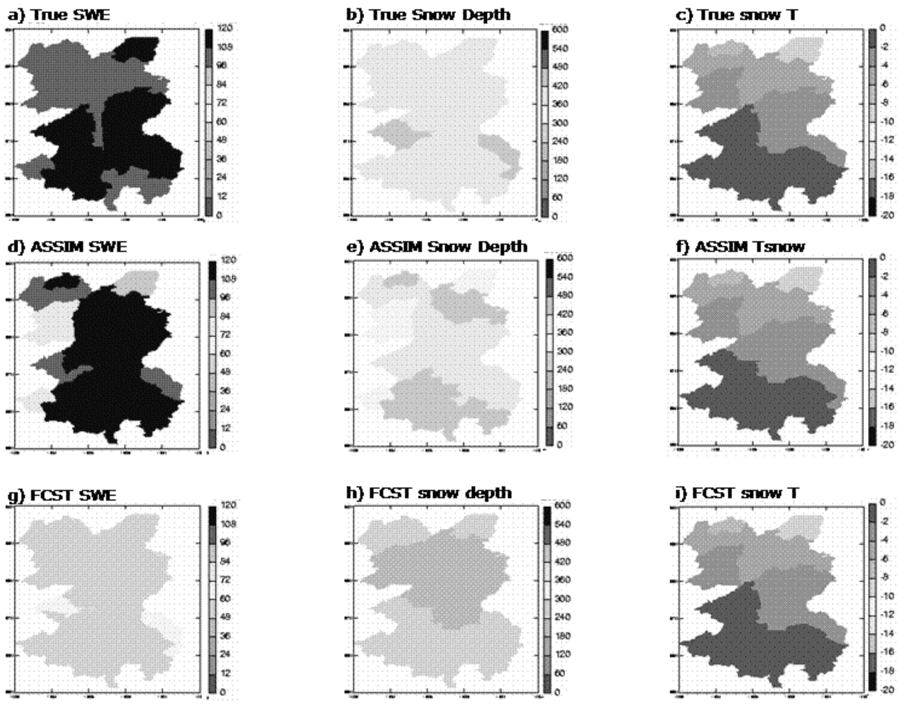


Figure 4: Snapshot of truth, assimilation and forecast (from poor initial condition) on 16 March 1987 from 3-month assimilation starting from 1 January 1987. Here, a), d), g) snow water equivalent (millimeters); b), e), h) snow depth (millimeters), c), f), i) temperature(C).

7. 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 Physical-space Statistical Analysis System [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. 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

$$w^a = w^f + K\left(w^o - Hw^f\right) \tag{1}$$

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 every 3 hours was insufficient. Since w^f is biased, the traditional analysis equation 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,

$$\delta w^a = K\left(w^o - Hw^f + b^f\right) \tag{2}$$

$$w^a = w^f - b^f_{k-1} + \delta a^a \tag{3}$$

$$b^f_k = b^f_{k-1} - \gamma \cdot \delta w^a \tag{4}$$

b^f_k 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 found and the bias correction term is calculated as,

$$f_b = \frac{b^f}{\tau} \tag{5}$$

where τ = 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 modeled the time-dependent bias as,

$$b^f(t) = \sum_j (a_j \cos\omega_j t + b_j \sin\omega_j t) \tag{6}$$

and estimated the Fourier coefficients

$$a_j = a_j - \gamma \delta w^a \cos\omega_j t \tag{7}$$

$$b_j = b_j - \gamma \delta w^a \sin\omega_j t \tag{8}$$

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), with bias correction (Assimilation II), with incremental bias correction (Assimilation III), with diurnal bias correction (Assimilation IV) 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.

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

Table 1: Description of experiments

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-meter 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 meter temperature and specific humidity. Figure 6 allows only for model intercomparison, and we are in the process of obtaining a verification dataset.

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. The Physical-space Statistical Analysis System was used

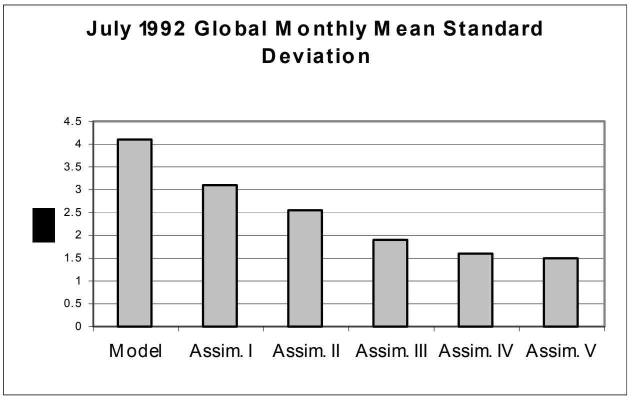


Figure 5: The July 1992 global monthly mean standard deviations of surface skin temperature between the assimilation experiments and the International Satellite Cloud Climatology Project observations (Table 1).

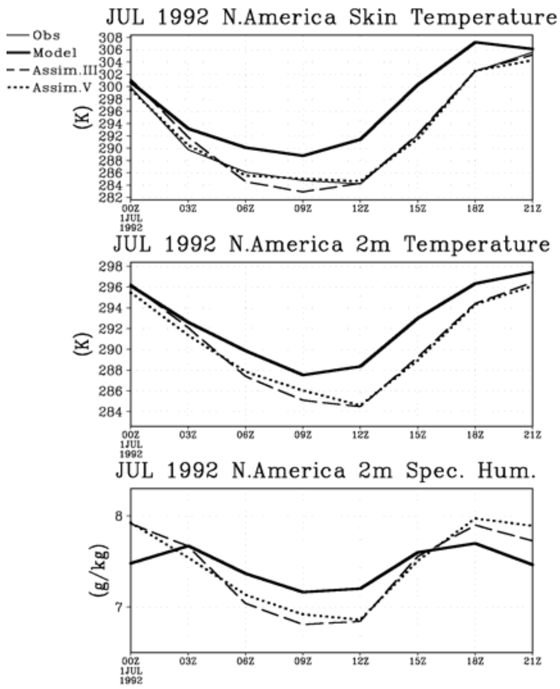


Figure 6. The July 1992 monthly mean diurnal cycle of skin temperature (top), 2 m temperature (middle) and 2 m specific humidity (bottom) over North America for the observations (light solid), Model (heavy solid), Assimilation III (dashed) and Assimilation V (dotted).

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.

8. Land data assimilation systems

The Global Land Data Assimilation System has its basis in the North American Land Data Assimilation System project [Mitchell et al. 1999]. The North American Land Data Assimilation System was initiated in 1998 with the goal of modeling land surface states and fluxes, while relying as much as possible on observation-based parameter and forcing fields in order to avoid biases that are known to exist in forcing fields produced by atmospheric models. The study region for the North American Land Data Assimilation System encompasses the conterminous United States and parts of Mexico and Canada. The land surface models implemented in the North American Land Data Assimilation System are run at 1/8th degree latitude by 1/8th degree longitude resolution. Separate versions of the system have been developed at the National Aeronautics and Space Administration's Goddard Space Flight Center, the National Oceanic and Atmospheric Administration's National Centers for Environmental Prediction, Princeton University, the University of Washington, and National Oceanic and Atmospheric Administration's Office of Hydrology. Each group runs their land surface models both in real time, retrospectively using the same high quality parameter and forcing fields, thus enabling unambiguous intercomparison of the land surface model simulations. Results are being validated by researchers at Rutgers University, using time series of observed variables, including soil moisture and temperature, to validate the strengths and weaknesses of each model. Much of the Global Land Data Assimilation System program code was derived from Goddard Space Flight Center's North American Land Data Assimilation System program code, and many of the project specifications are identical.

One of the primary objectives of the Global Land Data Assimilation System was to develop a system that would allow users to run multiple land surface models without specific knowledge of the models' architectures or physics. Currently, program code for three land surface models has been installed. Designing a Global Land Data Assimilation System simulation only requires modification of a single, simple interface file, which includes switches and variables for many run time options (summarized in Table 2). The Global Land Data Assimilation System program code interprets the forcing data to the individual input requirements of each respective land surface model, so that the same data can be used to force multiple land surface models. Thus, the influence of discrepancies in forcing data can be eliminated when comparing land surface fields simulated by different land surface models.

As a standard, all Global Land Data Assimilation System models run on a common 0.25 degree longitude by 0.25 degree latitude grid which is nearly global, covering all of the land north of latitude 60 degrees South. The Global Land Data Assimilation System also is able to run on 0.5 degree longitude by 0.5 degree latitude, a 1.0 degree longitude by 1.0 degree latitude, and a 2.5 degree longitude by 2.5 degree latutide global grid. Subgrid variability is simulated using a vegetation-based tiling approach, as described in the next section. The model time step is user-defined (15 minutes is standard). Forcing data is typically available on 0.25 degree longitude by 0.25 degree latitude to 1.0 degree longitude by 1.0 degree latitude grids with three or six

Spatial Resolution Options Degrees Longitude Degrees Latitude	 0.25 0.25	 0.5 0.5	 1.0 1.0	 2.5 2.0
Land Surface Model	Mosaic; Community Land Model; Noah			
Forcing	various model and satellite-derived products			
Initialization	None (constant value); restart file; forcing data			
Subgrid Variability	1 to13 tiles per grid cell (constant or fractional cutoff)			
Elevation Adjustment	temperature; pressure; humidity; longwave radiation			
Data Assimilation	surface temperature; snow cover			
Soil Classification	lookup table; Reynolds et al. [1999]			
Leaf Area Index	lookup table; satellite-derived			
Inland Water Tiles	Community Land Model lakes option			

Table 2 Options available in the Global Land Data Assimilation System user interface.

hourly resolution. The Global Land Data Assimilation System includes spatial and temporal interpolation routines based on commonly accepted algorithms. The Global Land Data Assimilation System uses a static, 1 kilometer resolution, global vegetation classification dataset produced by the University of Maryland [Hansen et al., 2000] from the Advanced Very High Resolution Radiometer data. The Global Land Data Assimilation System also employs a satellite observation-based, 1 kilometer resolution climatology and, when available, a time series of leaf area index. The soil parameter maps used in the Global Land Data Assimilation System were derived from the global soils dataset of Reynolds et al. [1999]. That dataset includes 5 minute resolution global maps of porosity and the percentages of sand, silt, and clay, which are based on the United Nations Food and Agriculture Organization Soil Map of the World (FAO 1990) linked to a global database of over 1300 soil pedons. The Global Land Data Assimilation System uses the Global 30 Arc-Second Elevation Data Set [Verdin and Greenlee, 1996] as its standard. The Global Land Data Assimilation System corrects the modeled temperature, pressure, humidity, and longwave radiation forcing fields based on the difference between the Global Land Data Assimilation System elevation definition and the elevation definition of the model that created the forcing data. Because some land surface models, including the Mosaic land model, ingest surface or bedrock slope as a parameter, geographic information systems software was used to assess the slope at each Global 30 Arc-Second Elevation Data Set pixel, and from those values the mean slope within each Global Land Data Assimilation System grid cell was computed.

9. Results

The Global Land Data Assimilation System runs daily in an operational mode. The Mosaic land model is the current operational model, but parallel simulations with the Community Land Model and the Noah land model are also used. The Goddard Earth Observing System Data Assimilation System is currently the baseline forcing source. The Goddard Earth Observing System Data Assimilation System precipitation and radiation fields are overwritten with the observation-based fields, when and where these are available. The model spatial resolution is 0.25 degrees longitude by 0.25 degrees latitude, and 10% is the minimum tile area allowed. The model time step is 15 minutes, and output is 3-hourly. Typically the daily near real time runs are complete within 36 to 48 hours of real time.

Results are presented from two simulations (Figure 7) : a Control Run and a Derived Forcing Run. Each run started on 1 January 2001. The Mosaic land model was used for both runs with the operational settings defined in the previous paragraph, except that the combination of forcing fields varied. The forcing data initialization option was used so that the Goddard Earth Observing System Data Assimilation System provided the initial surface energy and water storage states. Evidence suggests that this allowed the model to spin up and achieve reasonable stability in about three months. The Control Run relied on the Goddard Earth Observing System Data Assimilation System forcing exclusively. The Derived Forcing Run used the United States Naval Research Laboratory observation-based precipitation fields and the observation-based downward shortwave and longwave radiation fields.

The greater fine scale variability of the observation-based precipitation is reflected in the fine scale patterns of soil moisture in the Derived Forcing Run. Because rainfall tends to be spatially heterogeneous at local to regional scales and soil moisture shows a high degree of variability at all scales (e.g., Famiglietti et al., [1999]), the fine scale soil moisture variability evident in the Derived Forcing Run results may be preferable to the Control Run results. However, the exact locations of the fine scale features are unlikely to be reliable due to the imprecision of precipitation maps derived from satellite infrared observations of cloud top temperatures.

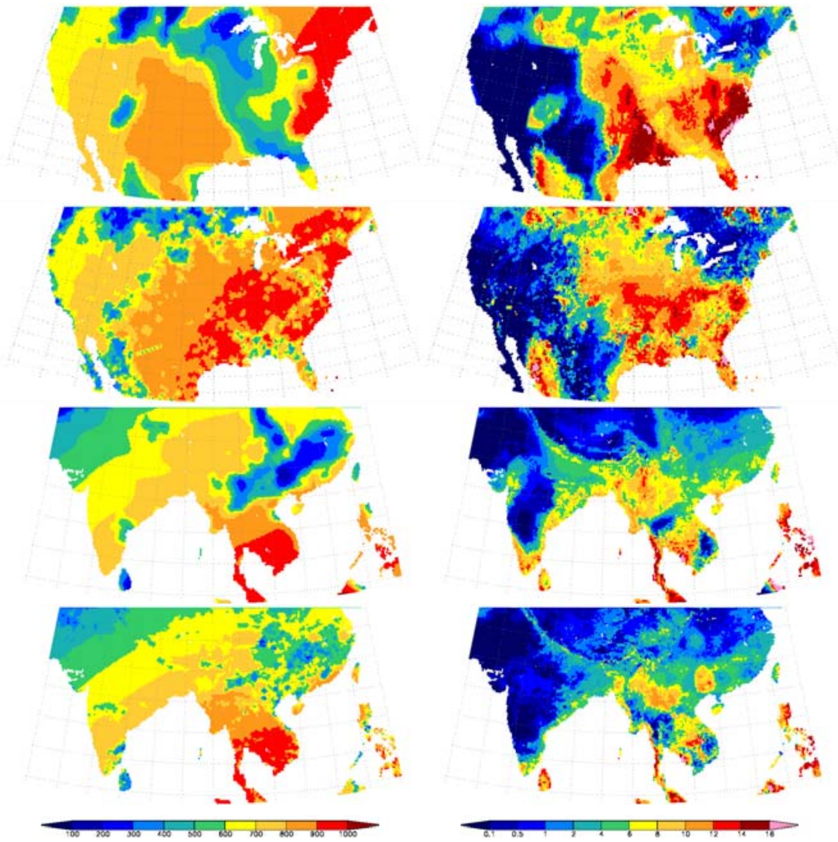


Figure7: Downward shortwave radiation forcing (W/m^2; left) and output total evapotranspiration rate (mm/day; right). From top to bottom: Control Run 18-21Z 31 July 2001; Derived Forcing Run 18-21Z 31 July 2001; Control Run 6-9Z 31 January 2002; Derived Forcing Run 6-9Z 31 January 2002. Top four: central North America; bottom four: southeast Asia.

One of the most ambitious activities of the Global Land Data Assimilation System project has been the assemblage of an archive of global, operational weather forecast model output and observation-based data fields for parameterizing and forcing land surface models. Most of the time series begin around January 2001 and continue up to present. The most recent fields are downloaded daily from forecast centers and groups that process satellite data.

Output fields of land surface states and fluxes from the Global Land Data Assimilation System model simulations are also freely available to the public (http://ldas.gsfc.nasa.gov). The Global Land Data Assimilation System website includes a real time image generator which allows users to view the most recent output fields. Time series are available by request, subject to manpower limitations. It is the intention of the Global Land Data Assimilation System project to encourage broad use of Global Land Data

Assimilation System results: for education, policy making, and social, agricultural, and natural hazards planning, as well as scientific research.

10. Future directions in land surface data assimiation

Land surface data assimilation is in its infancy, with many open areas of research. Development of land surface data assimilation theory and methods is needed to: (i) better quantify and use model and observation errors, (ii) optimize data assimilation computational efficiency for use in large operational applications, (iii) use radiative transfer forward models to enable the assimilation of brightness temperatures directly, (iv) link model calibration and data assimilation to optimally use available observation information, (v) create multivariate land surface assimilation methods to use multiple observations with complementary information, and (vi) quantify the potential of data assimilation downscaling. Further, the regular provision of remotely-sensed land surface variables with improved knowledge of observation errors in time and space are essential to advance land surface data assimilation. Land surface models must also be improved to: (i) provide more observable land model states, parameters, and fluxes, (ii) include advanced processes such as river runoff and routing, vegetation and carbon dynamics, and groundwater interaction to enable the assimilation of emerging observations, (iii) have valid and easily updated adjoint models, and (iv) have knowledge of prediction errors in time and space. The assimilation of new types of land surface observations, such as streamflow, vegetation dynamics, evapotranspiration, and groundwater or total water storage must be developed. Finally, we must understand the impact of land surface assimilation feedbacks on earth system predictions, and optimize the complexity of model, observation, and assimilation for practical real-world land surface problem solving.

References

Achutuni, R. and R. A. Scofield (1997) The spatial and temporal variability of the DMSP SSM/I global soil wetness index, *AMS Annual Meeting, Proceedings of the13th Conference on Hydrology*, 188-189.

Arkin, P.A., and B.N. Meisner (1987) The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Mon. Weather Rev*. **115**, 51-74.

Basist, A. and N. Grody (1997) Surface wetness and snow cover, *AMS Annual Meeting, Proceedings of the13th Conference on Hydrology*, 190-193.

Birkett, C. M. (1995) The contribution of the TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes, *J. Geophys. Res*. **100**(C12), 25179-25204.

Birkett, C. M. (1998) Contribution of the TOPEX NASA radar altimeter to the global monitoring of large rivers and wetlands, *Wat. Resour. Res*. **34**(5), 1223-1239.

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.

Cline, D. W., R. C. Bales and J. Dozier, Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling, *Wat. Resour. Res*. **34**(5), 1275-1285.

Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz, 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.

Dickinson, R. E., A. Henderson-Sellers, and P. J. Kennedy,(1993) Biosphere-Atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model. *NCAR Technical Note* 387+STR.

Eley, J. (1992) Summary of Workshop, Soil Moisture Modeling. Proceedings of the NHRC Workshop held March 9-10, 1992, *NHRI Symposium Proceedings* 9.

Engman, E. T. (1995) Recent Advances in Remote Sensing in Hydrology. *Reviews of Geophysics*, Supplement, 967-975.

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.

Famiglietti, J. S., J. A. Devereaux, C. A. Laymon, T. Tsegaye, P. R. Houser, T. J. Jackson, S. T. Graham, M. Rodell, and P. J. van Oevelen (1999) Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) Hydrology Experiment, *Wat. Resour. Res*. **35**, 1839-1851.

Hansen, M.C., R.S. DeFries, J. R. G. Townshend, and R. Sohlberg (2000) Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing* **21**, 1331-1364.

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**(12), 3405-3420.

Houser, P., E. Douglass, R. Yang, and A. Silva (1999) Merging Precipitation Observations with Predictions to Develop a Spatially & Temporally Continuous 3-hour Global Product. *GEWEX Conference*, Beijing, China.

Jackson, T. J. (1997a) Southern Great Plains 1997 (SGP97) Hydrology Experiment Plan, http://hydrolab.arsusda.gov/~tjackson .

Jackson, T. J. (1997b) Soil moisture estimation using special satellite microwave/imager satellite data over a grassland region, *Wat. Resour. Res*. **33**(6) 1475-1484.

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

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. Ducharne, M. Stieglitz, 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*, **33** 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.

Mitchell, K., P. Houser, E. Wood, J. Schaake, D. Tarpley, D. Lettenmaier, W. Higgins, C. Marshall, D. Lohmann, M. Ek, B. Cosgrove, J. Entin, Q. Duan, R. Pinker, A. Robock, F. Habets, and K. Vinnikov (1999) GCIP Land Data Assimilation System (LDAS) project now underway, *GEWEX News* 9(4), 3-6.

Reynolds, C. A., T. J. Jackson, and W. J. Rawls (1999) Estimating available water content by linking the FAO Soil Map of the World with global soil profile databases and pedo-transfer functions, *American Geophysical Union, Fall Meeting, Eos Trans*. AGU, 80.

Rodell, M., and J. S. Famiglietti (1999) Detectability of variations in continental water storage from satellite observations of the time dependent gravity field, *Wat. Resour. Res*. **35**, 2705-2723.

Rosenthal, C. W. and Dozier, J. (1996,) Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper, *Wat. Resour. Res*. **32** (1), 115-130.

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. Meteor. Soc.* **77**, 1987-2005.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher (1986) A simple biosphere model (SiB) for use with general circulation models. *J. Atmos. Sci*. **43**, 505-531.

Sipple, S., S. Hamilton, J. Melak, and B. Choudhury (1994) Determination of inundation area in the Amazon river flood plain using SMMR 37 GHz polarization difference. *Remote Sensing Environment* **48**, 70-76.

Verdin, K. L., and S. K. Greenlee (1996) Development of continental scale digital elevation models and extraction of hydrographic features. Proceedings, *Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, Santa Fe, NM, January 21-26, National Center for Geographic Information and Analysis, Santa Barbara, CA.

Wahr, J.; Molenaar, M.; Bryan, F. (1998) Time variablity of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res*. **103** (B12), 30205-30229.

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** (D11), 11761-11774.