Modification and Application of the Satellite-Based Land Data Assimilation Scheme for Very Dry Soil Regions Using AMSR-E Images: Model Validation for Mongolia—a CEOP data platform

Pathmathevan MAHADEVAN, Toshio KOIKE, Hideyuki FUJII, Katsunori TAMAGAWA

Department of Civil Engineering, University of Tokyo, Tokyo, Japan

Xin LI

Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, P. R. China

and

Ichirow KAIHOTSU

Dept. of Civil Engineering, Hiroshima University, Higashi-Hiroshima, Japan

(Manuscript received 27 January 2006, in final form 10 August 2006)

Abstract

At climatic time-scales, soil moisture is one of the most important boundary condition controlling fluxes to the atmosphere. Here, we explore the feasibility of synthesizing distributed fields of soil moisture using AMSR-E observations and a novel application of data assimilation within a hydrological model. We modified our existing Land Data Assimilation Scheme (LDAS) by specifically considering: (1) weak constraint assumptions rather than a strong constraint, thus accounting for the existence of model errors; and (2) the effects of volume scattering within the soil medium in the Radiative Transfer Model (RTM). Adopting the “effects of volume scattering within the dry soil medium” in the RTM is a new step for satellite-based data assimilation techniques for the retrieval of soil moisture.

This LDAS can be used to assess model parameters and estimate vertical profiles of soil moisture, especially in very dry regions (volumetric soil moisture is equal to or less than 5–15%) as well as soil-surface and canopy temperatures by comparing passive microwave observations using a unique minimization technique termed Very Fast Simulated Annealing (VFSA). To validate our new LDAS, AMSR-E observations, gathered in Mongolia, were assimilated into the Land Surface Model (LSM) via the modified RTM. Observed micrometeorology boundary conditions for Mongolia were drawn from the CEOP database. In studies that simulate 2-week dry periods, the results of the revised LDAS are in better agreement with observational data than the results of open-loop simulations.

1. Introduction

A quantitative understanding of the spatial distributions and temporal variations of land-surface variables, especially soil moisture, is essential to understand the physics of land atmosphere interaction of water, energy, and
carbon. In addition, soil moisture must be accurately represented in hydrologic and land-surface models because it contributes to regulating the rate of runoff generated during a rain event (Dirmeyer et al. 1999). Extensive amounts of hydrologically relevant remote sensing and ground observations are becoming increasingly available; these data may provide the additional information necessary to constrain the spatial and temporal distribution of land-surface states. However, these observations require special capabilities to enable analysis and interpretation (Walker et al. 2003).

The development of assimilation techniques for land hydrologic data that emphasize soil moisture and temperature, together with innovative remote sensing and surface-based observations, promises a viable mechanism for the synthesis of regional fields of these vital hydrological variables. Data assimilation is the process of incorporating observations into a model (e.g., the land atmosphere interaction system) and filling the region of missing data with model predictions, as well as providing a suite of data-constrained estimates of unobserved quantities (Daley 1991; Talagrand 1997; Kalnay 2003). The development of assimilation techniques for hydrologic data is still in its infancy, although significant progress has been made in advancing hydrologically relevant remote sensing and assimilation techniques via both ground-based and airborne field studies (Reichle et al. 2001; Houser 1998; Walker 2001).

To address these issues, in earlier studies we developed a more satisfactory One-Dimensional Variational (1DVAR) Land Data Assimilation Scheme (LDAS) that assimilates passive microwave observations into a hydrological model (Pathmathevan et al. 2003a). We validated the scheme for different vegetations types (Pathmathevan et al. 2003b) and different bare soils (Pathmathevan et al. 2005) using Ground-Based Microwave Radiometer—6 channels (GBMR-6CH) observations because microwave radiometry provides a remote-sensing tool that is well suited to the determination of soil moisture. The tool is well suited because the thermal emission of soils is related to the appreciable moisture content in the microwave frequency range (Dobson et al. 1985; Jackson and Le Vine 1996; Njoku and Entekhabi 1996), and the potentials of TRMM/TMI, AQUA/AMSR-E and other impending satellite sensors (see Fig. 1) are expected to play important roles in diurnal-cycle studies of land atmosphere interaction.

Despite its obvious potential, 1DVAR-LDAS has a number of limitations: (1) Present LDAS is used with the strong constraint assumption (i.e., the model is assumed to be perfect), yet it is well known that forecast models are imperfect. The model error needs to be taken into account if a sophisticated data assimilation method is to be developed (Zupanski 1997). (2) As well as making the problem computationally simpler and tractable we use a very simple RTM, in addition to strong constraint assumption, to generate brightness temperatures to be assimilated with satellite observations.

The observation operator should be physically based and more accurate. Current radiative transfer models assume that the emission from deep layers (approximately > 5 cm) is negligible; however, this assumption is applicable only for wet soil. As the presence of liquid water strongly increases the dielectric losses, wet soil is opaque for microwaves and the signals emitted signals from a very thin layer at the top of the soil medium. In dry soil, scattering effects are more pronounced at shorter wavelengths, as with dry snow (Chang et al. 1987; Rott 1987), and depend on soil-particle size and shape as well as the complex dielectric constant of the scatterers. The effect of volume scattering means that the microwave emissivity of dry soil does not show a linear relationship with soil moisture. Note that if soil moisture is greater than 10–15%, emissivity decreases linearly with increasing soil moisture (Dobson et al. 1985; Ulaby et al. 1986).

In the present study, we address the major limitations of the existing 1DVAR-LDAS (Pathmathevan et al. 2003a,b, 2005) and its modifications. Firstly, the existing 1DVAR-LDAS is modified into a more physically based assimilation system by incorporating weak constraint assumptions that emphasize the need to account for time-varying model-based errors.

Secondly, the effects of volume scattering within a dry soil medium are introduced in the surface RTM to enable a more accurate estimate of the total emission of radio brightness; this includes scattering effects from deeper soil
layer that are the key when the soil condition is very dry. We validate this approach using microwave observations derived from AMSR-E. This is the first experiment to introduce a RTM with volume scattering capabilities to recent satellite-based data assimilation techniques for the retrievals of soil moisture (Zhan et al. 2004, 2006; Reichle et al. 2001).

Finally, to apply and validate the revised LDAS, AMSR-E observations, gathered over a validation site in Mongolia are assimilated into LSM, which is a Soil Vegetation Atmosphere Transfer (SVAT) scheme, from a selected very dry simulation period. The results show that retrieved land-surface variables can be estimated satisfactorily compared to equivalent estimates derived from an open-loop simulation especially for very dry soil moisture conditions.

2. Model framework

2.1 Revised Land Data Assimilation Scheme

The existing version of LDAS was formulated by recognizing that both model predictions and measurements provide useful information on the actual state of the land surface. Combining these two sources of information in an optimum way, LDAS produces a cost function (Pathmathavan et al. 2003a,b, 2005) under the assumption of strong constraint formulations (Sasaki 1970; Talagrand 1990; Daley 1991; Bouttier and Rabier 1997), as shown in Eq. (2.1):

\[
J(x_0) = \frac{1}{2} \sum_{i=0}^{N} (H_i[M(x_i)] - y_i^o)^T R_i^{-1}(H_i[M(x_i)] - y_i^o) + \frac{1}{2} (x_0 - x_0^b)^T B^{-1}(x_0 - x_0^b)
\]

where \( y^o \) represents all radiometer observations, \( H \) is the observation operator, \( M \) is the model operator, and \( R \) is the observational error covariance matrix. In addition, \( x_0 \) is the initial state of the analysis variable \( x \) or the state vector, \( x^b_0 \) is the background field, and \( B \) is the background error covariance matrix. Thus, the model is used as a strong constraint, i.e., the analysis solution has to satisfy the model equations. In other words, LDAS seeks an initial condition such that the forecast best fits the observations within the assimilation window or interval.

The first important extension of existing LDAS in the revised LDAS is the development of a variational method with a weak constraint.
(Bennett 1992; Bennett et al. 1997) rather than a strong constraint, thus accounting for the existence of model errors in the evolution of the forecast and the forecast error covariance. Accounting for model error in a cost-effective way received more attention because the model errors are expected to be systematic and time-correlated. The forecast errors appear as random forcing on the dynamics, with an a priori forecast error covariance matrix \( Q \). The cost function therefore becomes

\[
J(x_0, \Phi) = \frac{1}{2} \sum_{i=0}^{N} (H_i[M(x_i, u_i)] - y_i^0)^T R_i^{-1} (H_i[M(x_i, u_i)] - y_i^0) \\
+ \frac{1}{2} \sum_{k=0}^{N} (\Phi_k)^T Q_k^{-1} (\Phi_k) \\
+ \frac{1}{2} (x_0 - x_0^0)^T B^{-1} (x_0 - x_0^0)
\]  

(2.2)

where the second term accounts for the model errors. The forecast model \( M \) and observation operator \( H \) are imposed as weak constraints defined by

\[
x_{k+1} = M(x_k, u_k) + \Phi_k
\]  

\[
y_i = H_i(x_i) + e_i
\]  

The indexes \( i \) and \( k \) define observational times and model time-steps, respectively. The nonlinear operator \( M \) includes all deterministic forcing components \( u_k \) at time step \( k \) such as observed micro-meteorologic inputs. The model error correction vector \( \Phi_k \) accounts for the growth in model error from time \( t_k \) to time \( t_{k+1} \). The model errors \( \Phi_k \) are assumed to be stochastic variables that are unbiased and uncorrelated in time, with a known Gaussian distribution and covariance matrices given by \( Q_k \). The observational errors, model errors, and errors in the prior estimates are assumed to be uncorrelated. Figure 2 illustrates model concepts and error assumptions of the revised LDAS.

For an optimal analysis, we aim to find the best estimates \( \langle x_k^a \rangle \) for the expected values of the true states \( \langle x_k \rangle \), given satellite observations \( y \), subject to the model Eq. (2.3) and prior estimates \( \langle x_0^0 \rangle \). Under the statistical assumptions made here, the optimal analysis is given by the weighted least squares (Bayesian) performance function (or cost function) \( J \). The performance function \( J \) minimizes the square error between the model predictions and observed system states, together with the square error in the model equations; all are weighted by the inverse of the covariance matrices over the assimilation window. The model Eq. (2.3) are thus treated as weak constraints on the objective function. The initial states of the system and the model errors at each time step are the control parameters that must be determined by minimizing the cost function.

Minimizing \( J \) via minimization techniques yields the optimal solution of the above cost function Eq. (2.2). Here, we employ a heuristic optimization approach termed Very Fast Simulated Annealing (VFSA) (Kruger 1993; Ingber 1989), which is capable of minimizing the variational cost function without using adjoint models, as these are more complicated, especially for the non-linear operates. Pathmathevan et al. (2003a,b) describe the basic concepts.
of the VFSA scheme, its advantages and disadvantages, the procedure of minimizing the LDAS cost function using VFSA, and VFSA performance. The VFSA avoids the pitfalls associated with strong non-linearity and discontinuity, in obtaining the global minimum in the hilly structure of the cost function.

2.2 Land-surface hydrological model for data assimilation

A hydrological model for the assimilation of land-surface data must capture key physical processes while remaining sufficiently efficient that large-scale optimal estimation is computationally feasible. These are the potentially conflicting requirements that need to be traded off when a model is selected. Commonly used physically based land-surface models include BATS (Dickinson et al. 1986) and BATS2 (Dickinson et al. 1998), SiB (Sellers et al. 1986) and SiB2 (Sellers et al. 1996), NOAH-LSM (Mitchell et al. 2000), and VIC-2L (Nijssen et al. 1997).

From the variety of models described in the literature, we selected the one dimensional hydrological processes model SiB2 (Sellers et al. 1996) as the model operator in our LDAS and assumed that lateral moisture and heat fluxes in the unsaturated zone are negligible. This assumption is reasonable for terrain with moderate relief over the spatial scales considered here. SiB2 has been successfully tested by comparing its predictions with field measurements from a soil moisture experiment undertaken in 2002 (SMEX02) (Pathmathevan et al. 2003b) within the USA, an experiment for bare soils (Pathmathevan et al. 2005) undertaken in Japan, and an experiment for regions with frozen soils (Pathmathevan et al. 2003a) undertaken in Tibet.

Additional reasons for our selection of SiB2 as the model operator, the advantages of applying SiB2 to grassland and other land-surface classes, and experiences with some of the earlier global-scale applications (as open-loop simulations) (Sato 1989) have been clearly summarized in previous studies (Pathmathevan et al. 2003a,b, 2005; Sellers et al. 1996; Li and Koike 2001). On the basis of these tests and experiences, we believe that SiB2 is sufficiently accurate and computationally efficient to form the basis of an operational soil moisture data assimilation scheme.

2.3 Radiative transfer model

A Background to the existing RTM

AMSR-E housed on the Aqua satellite features a passive microwave sensor for detecting surface soil moisture. Pathmathevan et al. (2003a,b) used a simplified radiative transfer model that takes into account the main components of microwave emission from moist soil:

\[ T_{BT} = (1 - R)T_S \exp(-\tau_c) + (1 - \omega_c)(1 - \exp(-\tau_c))T_c \]

where \( T_{BT} \) is brightness temperature, \( \omega_c \) is vegetation single-scattering albedo, \( T_S \) is soil temperature, \( T_C \) is vegetation temperature, \( T_S \) and \( T_C \) are outputs of the model operator, \( \tau_c \) is the vegetation opacity, and \( R \) is soil reflectivity (see Fig. 3).

In this model, \( \tau_c \) is a function of the vegetation water-content \( (w_c) \) [kg m\(^{-2}\)] (Jackson and Schmugge 1991):

\[ \tau_c = bw_c / \cos \theta \]

where \( b \) is a coefficient that depends on the canopy structure and frequency \( (\lambda) \) in the form \( b = b' / \lambda^{r} \); \( \theta \) is the incident angle.

Soil reflectivity \( (R) \) can be expressed as a function of the dielectric constant of the soil \( (\varepsilon_r) \) which is then mainly dependent on one of the state variables; the soil moisture \( (m_v) \):

\[ \varepsilon_r^z = 1 + \frac{\rho_b}{\rho_s} (\varepsilon_a^z - 1) + m_v \frac{\rho_a}{\rho_s} - m_v \]

where \( \rho_b \) is the soil bulk density, \( \rho_s \) is the soil
specific density; \( \varepsilon_r \) is the dielectric constant of soil with extremely low moisture content, \( \varepsilon_r \approx (4.7, 0) \), \( \varepsilon_{fw} \) is the dielectric constant of free water, and \( \alpha \) and \( \beta \) are two empirical coefficients that are dependent on the soil texture.

To incorporate the effects of surface roughness, the smooth surface reflectivity \( R_s \) is modified using a polarization mixing parameter \( Q \) and a roughness height parameter \( \eta \) (Wang and Choudhury 1981). \( Q \) is the proportion for which the horizontal and vertical polarizations are mixed. The parameter \( \eta \) is related to the standard deviation of the surface height. The Fresnel power reflectivity from a rough surface can be expressed as

\[
R_r(p, \theta) = [(1 - Q)R_s(p, \theta) + QR_s(q, \theta)] \exp(-\eta) \tag{2.8}
\]

where \( p \) and \( q \) denote the horizontal and vertical polarizations, respectively, and \( r \) and \( s \) denote rough and smooth surfaces, respectively. The roughness height parameter \( \eta \) is given as

\[
\eta = 4(2\pi/\lambda)^2 \sigma^2 \cos^2 \theta \tag{2.9}
\]

where \( \lambda \) is the wavelength and \( \sigma \) is the standard deviation of the surface height.

b Model for volume scattering effects in the soil medium

The products of the RTM are highly sensitive to land-surface skin properties for wet surface conditions; however, for a dry soil, emission from deeper layers makes a significant contribution. In this study, we modified the existing RTM by incorporating volume scattering effects, especially for a dry soil medium. For such a very dry soil medium, we assume that both the soil medium and atmosphere show similar scattering behaviours.

Following the theoretical basis provided by Tsang and Kong (1977) in their formulation of a 4-stream radiative transfer model for the atmosphere, a similar model for the soil medium has been developed to address volume scattering effects. In developing a volume scattering microwave radiative transfer model, we propose to use the above 4-stream approximation, especially to compute the scattering source term and calculate radiance at any required direction from the formal solution of the RTM (Chandrasekhar 1960). In addition, we adopt the Henyey and Greenstein (1941) scattering phase function, which is similar to Liu’s (1998) 4-stream atmospheric model, using an asymmetry factor calculated from Mie theory rather than calculating the Mie scattering matrix. Liou (1974) showed that the eigenvalue problem can be solved analytically for both 2- and 4-stream approximations that keep these models computationally efficient, and that the 4-stream method substantially improves the accuracy of radiative flux results compared to the 2-stream method.

Implementation of the phase function enables a simple expression when it is expanded in Legendre polynomials and simplifies the numerical calculation dramatically. As no polarization is taken into account in this phase function, the simple model does not simultaneously calculate radiance at both horizontal and vertical polarizations and does not have cross-polarization scattering.

Model formulation

The radiative transfer process for polarized waves in a plane-parallel and azimuthally symmetric soil medium with spherical scattering particles can be expressed by (Tsang and Kong 1977)

\[
\frac{d}{d\tau} \begin{bmatrix} I_V(\tau, \mu) \\ I_H(\tau, \mu) \end{bmatrix} = \begin{bmatrix} I_V(\tau, \mu) \\ I_H(\tau, \mu) \end{bmatrix} - \frac{\omega_0}{2} \begin{bmatrix} P_{VV} & P_{VH} \\ P_{HV} & P_{HH} \end{bmatrix} \begin{bmatrix} I_V(\tau, \mu) \\ I_H(\tau, \mu) \end{bmatrix} d\mu' - (1 - \omega_0)B(\tau) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{2.10}
\]

where \( d\tau = k_i dz \), \( I_p(\tau, \mu) \) is the radiance at optical depth \( \tau \) in the direction \( \mu \) for vertical or horizontal polarization, \( \omega_0 \) is the single scattering albedo, \( B(\tau) \) is the Plank function, and \( P_{ij}(i, j = H \text{ or } V) \) represents the scattering phase functions. The above 4-stream model solves Eq. (2.10) using the discrete ordinate method and introducing approximations for which cross-polarization does not exist. The scattering phase function is expressed by the Henyey Greenstein formulae (Henyey and Greenstein 1941) as follows:

\[
P(\mu\mu') = \sum_l A_l p_l(\mu)p_l(\mu') \tag{2.11}
\]
where $A_l = (2l + 1)g^l$, with $g$ being the asymmetry factor given by Bohren and Huffman (1983) and $p_1(\mu)$ is the first order Legendre polynomial. Using these two adjustments and limiting the stream number to four makes it possible to solve the eigenvalue problem for Eq. (2.10) analytically, which considerably speeds up the computations (Liu 1998). This is an important consideration for our data assimilation application because many model evaluations are necessary.

Following Stamnes and Swanson (1981), a formal solution of Eq. (2.10) without the inclusion of polarization can be expressed as:

$$I(\tau, +\mu) = I(\tau^*, +\mu)e^{-(\tau^* - \tau)/\mu} + \int_{\tau^*}^{\tau} J(t, +\mu)e^{-(t - \tau)/\mu} \frac{dt}{\mu}$$

$$I(\tau, -\mu) = I(0, +\mu)e^{-\tau/\mu} + \int_{0}^{\tau} J(t, -\mu)e^{-(\tau - t)/\mu} \frac{dt}{\mu}$$

(2.12) (2.13)

where $+\mu$ and $-\mu$ represent upward and downward directions, respectively, and $\tau^*$ is the optical depth of the layer. The source term $J(\tau, \mu)$ with the discrete ordinate approximation is:

$$J(\tau, \mu) = \frac{1}{2} \omega_0 \sum_{l=0}^{2n-1} A_l p_l(\mu) \sum_{j=-n}^{n} a_j p_j(\mu) I(\tau, \mu_j) + (1 - \omega_0)B(\tau)$$

(2.14)

where $a_j$ is the quadrature weight for the $j$th quadrature point and $2n$ is equal to 4 in our 4-stream approximation. The solution at the $j$th quadrature point can be expressed as (Liou 1973):

$$I(\tau, \mu_j) = \sum_{j=-n}^{n} L_j W_j(\mu_j)e^{-k_j\tau} + q(\mu_j) + B_1\tau$$

(2.15)

where $k_j$ and $W_j(\mu_j)$ are an eigenvalue and eigenvector, respectively, and $L_j$ and the rest of the calculation can be determined from the continuity of radiance between layers and the boundary condition following the standard procedures described by Liou (1974).

**Numerical setup**

In performing numerical calculations for within the soil medium, we divide the soil medium into many layers and assume that all microphysical properties are uniform within each layer. Apart from the analytical challenges of the solution of the radiative transfer equation, the quality of a radiative transfer model also depends on how well absorbers and scatterers are represented numerically. Therefore, we will briefly consider those formulae that the model uses to represent the major constituents of the soil medium that are considered to have an influence on the radiative transfer.

Firstly, it is possible to estimate the scattering effects of a soil particle by assuming a spherical shape. The average size $(d)$ and total number $(N)$ of particles were fine tuned using available AMSR-E observations, soil density, and the total soil volume at the model application area in Mongolia. The volume scattering coefficient $(k_s)$ is the sum of the scattering cross-section $(\sigma_s(d))$ of all particles:

$$k_s = \sum_{i=1}^{N} \sigma_s(d_i)$$

(2.16)

where $\sigma_s(d)$ was estimated using Mie theory (Bohren and Huffman 1983).

Secondly, the minor nature of changes in soil moisture and temperature profiles within the soil medium indicates that changes in the dielectric properties of each layer can be ignored. Consequently, reflections at each boundary are ignored, except for those at the surface layer that are associated with significant differences in dielectric properties across the air soil interface. Finally, to solve Eq. (2.10), we need to apply the boundary condition for the bottom of the soil medium. The bottom of the soil medium is set at 1.5 m (layer thickness is 5 cm), and the brightness temperature at the bottom is assumed to be the soil temperature at that level, as derived from model output.

**c Revised radiative transfer model**

The total emission of scattering to the top of the soil medium is estimated using the 4-stream volume scattering model, and the total microwave brightness temperature at the surface level is defined as $T_b$. As mentioned above, the minor nature of changes in the soil moisture profile within the soil medium means that changes in the dielectric properties within each layer can be neglected; consequently,
reflections at each boundary are also ignored. However, significant differences in the dielectric properties of air and soil at the surface layer mean that we consider a thin layer at the air-soil interface to account for reflection and refraction.

Figure 4 shows a schematic of 4-stream radiative transfer processes within the soil medium and reflection and refraction processes at the top of the soil medium. Following corrections for Fresnel reflection and surface scattering the final brightness temperature $T_b^+$ at the top of the soil medium is given as:

$$T_b^+(\theta_0, p) = e(\theta_0, p)T_b^-$$  \hspace{1cm} \text{(2.17)}

Note that if the soil medium shows considerable wetness the $T_b^-$ is automatically considered as the physical temperature at the top layer of the soil medium.

Finally, to estimate the brightness temperature observed using AMSR-E, corrections for vegetation, and if necessary for the atmosphere should be carried out as explained in Eq. (2.5) and (2.6); the model parameters can then be determined with significant effort. Table 1 summarises the assigned RTM parameters used in the application of the model in Mongolia.

3. Experimental domain and data

Mongolia has been the focus of extensive hydrological and meteorological research in recent years. To develop and verify algorithms for satellite remote sensing of soil moisture using microwave radiometers, in particular AMSR-E and AMSR, soil moisture conditions and related meteorological/hydrological factors are currently being monitored by automatic stations (Automatic Weather Stations (AWS) and Automatic Stations of Soil Hydrology (ASSH); Fig. 5) distributed throughout the Mongolian Plateau, where high-quality data can be obtained because of relatively uniform ground-surface conditions.

This project was implemented under the framework of JAXA-JRA and named as “Ground Truth for Evaluation of Soil Moisture and Geophysical/Vegetation Parameters Related to Ground Surface Conditions with AMSR and GLI in the Mongolian Plateau”. Under this project, data from automatic stations are available for every year since 2000. A large number of ground-based observations have also been collected in the study area since 2000, and all data are available online for research use via CEOP framework (http://www.ceop.net).

3.1 Atmospheric forcing

Among the six forcing variables required to run SiB2, shortwave downward radiation (in the format of net radiation), air temperature, wind speed, and precipitation were obtained from the Mongolia experiment. Longwave downward radiation was calculated using Brunt’s equation (Moenteith 1973), a standard method in SiB2, and vapor pressure was transformed from humidity and air temperature. From the limited number of observational points, all the data were first checked in terms of their quality and then interpolated using the Inverse Distance Weighted (IDW) interpolation method, with topographical corrections to all simulation grids.

To obtain the relationship between net radiation and shortwave downward radiation we used MOLTS (Model Output of Location Time Series) data for the Mongolia reference sourced from the CEOP data bank. Estimated shortwave downward radiation was compared (see Fig. 6) with available observations undertaken at Mandangobi station (http://atmos.cr.chiba-u.ac.jp/aerosol/skynet/index.html) in Mongolia by the SKYNET data centre. Reasonable results were observed and incorporated into model simulations.

3.2 SiB2 parameterization associated with the Mongolia Plateau

If we assume that the lateral moisture and heat fluxes in the unsaturated zone are negligible, it is then possible to partition the experi-
mental domain into one-dimensional (1-D) grid cells. In 1-D studies (Montaldo et al. 2001; Castelli et al. 1999; Entekhabi et al. 1994; Galantowicz et al. 1999), computational resources are not the limiting factor. Generally, for all grids, the land cover type is classified as short vegetation/C4 grassland, which is the sixth class in Seller’s definition of global land-cover classification (Sellers et al. 1996). Most of the static parameters associated with land-cover type were obtained directly from the above references, while canopy height, vegetation-cover fraction, and root depth were derived from records from the Mongolia experiment.

Values of soil parameters, such as soil porosity, hydraulic conductivity at saturation, soil wetness exponent, and soil tension at saturation, as defined by Clapp and Hornberger (1978), were obtained from Sellers et al. (1996). Table 1 summarizes the additional parameter values used in this summertime study in Mongolia.

### Table 1. Calibrated (ASSH1) and applied vegetation, soil, and RTM parameters for a general grid cell upon the Mongolian experimental plateau.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>Vegetation type</td>
<td>6 in SiB2</td>
</tr>
<tr>
<td>$z_2$</td>
<td>Canopy top height (m)</td>
<td>0.05</td>
</tr>
<tr>
<td>$z_1$</td>
<td>Canopy base height (m)</td>
<td>0.01</td>
</tr>
<tr>
<td>$V$</td>
<td>Vegetation cover fraction</td>
<td>0.3</td>
</tr>
<tr>
<td>$D_r$</td>
<td>Rooting depth (m)</td>
<td>0.15</td>
</tr>
<tr>
<td>$N$</td>
<td>Green leaf fraction</td>
<td>0.50</td>
</tr>
<tr>
<td>$L_T$</td>
<td>Leaf Area Index</td>
<td>0.51</td>
</tr>
<tr>
<td>FPAR</td>
<td>Canopy absorbed fraction of Photosynthetically active radiation (PAR)</td>
<td>0.24</td>
</tr>
</tbody>
</table>

### Static parameters associated with soil type

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>Soil type</td>
<td>Sand loom</td>
</tr>
<tr>
<td>$D_s$</td>
<td>Surface Layer depth (m)</td>
<td>0.05</td>
</tr>
<tr>
<td>$D_T$</td>
<td>Total depth of 3 soil moisture layers (m)</td>
<td>1.5</td>
</tr>
<tr>
<td>$K_s$</td>
<td>Hydraulic conductivity (m/s)</td>
<td>$1.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>$\theta_s$</td>
<td>Soil porosity</td>
<td>0.44</td>
</tr>
<tr>
<td>$\rho_s$</td>
<td>Saturation soil density (Kg m$^{-3}$)</td>
<td>2.65</td>
</tr>
</tbody>
</table>

### Radiative Transfer Model Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q$</td>
<td>Polarization mixing ratio</td>
<td>0.1–0.3</td>
</tr>
<tr>
<td>$h$</td>
<td>Roughness coefficient</td>
<td>0.0–1.0</td>
</tr>
<tr>
<td>$x$</td>
<td>Vegetation canopy parameter</td>
<td>$-1.38$</td>
</tr>
<tr>
<td>$b_1$</td>
<td>Vegetation canopy parameter</td>
<td>9.32</td>
</tr>
<tr>
<td>$d$</td>
<td>Soil particle size (mm)</td>
<td>1.2</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Inclination angle</td>
<td>55.0</td>
</tr>
<tr>
<td>$\omega_v$</td>
<td>Vegetation single scattering coefficient</td>
<td>0–0.1</td>
</tr>
<tr>
<td>$We$</td>
<td>Vegetation water content (Kg m$^{-2}$)</td>
<td>0.0–0.25</td>
</tr>
</tbody>
</table>
3.3 Satellite data

NASA’s Earth Observing System (EOS) Aqua Satellite was launched from Vandenberg AFB, California on May 4, 2002 at 02:54:58 Pacific Daylight Time. The primary goal of Aqua, as the name implies, is to gather information on water in the Earth’s system. Equipped with six state-of-the-art instruments, Aqua will collect data on global precipitation, evaporation, and the cycling of water. The Advanced Microwave Scanning Radiometer—EOS (AMSR-E) is one of the six sensors aboard Aqua. AMSR-E operates on six frequencies (twelve channels) in a total-power passive microwave radiometer system. The AMSR-E sensor measures brightness temperatures at 6.925, 10.65, 18.7, 23.8, 36.5, and 89.0 GHz. Vertically and horizontally polarized measurements are taken at all channels.

In the present study, we used lower frequencies of 6.925, 10.65 and 18.7 GHz where the higher frequencies of 21.3, 37.0, and 85.5 GHz were assumed to be highly sensitive to rainfall effects and were therefore not utilized. Satellite observations of brightness temperatures are available only at a relatively large spatial scale. For our purposes, satellite information is transformed from larger to smaller (grid-size) scales, with interpolation and arithmetic averaging methods used for the downscaling process. It is possible to apply these methods because of the relatively uniform ground-surface conditions.

4. Results and discussions

The developed assimilation scheme was tested using direct brightness temperature observations from the AMSR-E sensor for the summer period (very dry soil conditions) from July 2–19, 2002. We undertook comparisons with the results of open-loop simulations to illustrate the effect of soil moisture initialization errors on model predictions and assess the effectiveness of the data assimilation scheme.

In our applications, the length of the assimilation window is 24 hours. The study area is divided into $26 \times 26$ estimation pixels, and
soil moisture and temperature simulations are performed for these 10 × 10 km² scales of estimation pixels, at a temporal resolution of 1 hour. For this example, remotely sensed microwave radiometer measurements are made available at 10 × 10 km² scale of the observation pixels. Simulation results of soil moisture were then compared with ground observations from the Mongolian Plateau.

For the first run (first assimilation window) of LDAS, the model state was initialized with a reasonable first guess \( x_0(t_0) = x^b + \epsilon \) from a priori knowledge \( x^b \). The first average background values of the volumetric soil moistures were assumed to be 10% for surface and root layers and 22% for deep zones half the porosity, with an error standard deviation of 10%. The initial soil-surface and canopy temperatures were set to equal the air temperature with an error standard deviation of 2 K. Furthermore, we assumed that the observational errors (errors in the measured satellite data) are not correlated with the signal and are temporally invariant with a standard deviation of 3–5 K.

The LDAS was then run, subjected to surface forcing, and updated each day with radio brightness measurements. Prior to the LDAS simulations, SiB2 (open-loop) simulations were carried out continuously for each of the grid cells of the study area for the whole of the simulation period with the same atmospheric forcing and the first guess of the initial conditions. The results from the LDAS were then compared with the open-loop simulations at twelve SSH stations. Sample results for three different stations (ASSH8, ASSH6, and ASSH7) that represent average (~15%), low (~10%), and lowest (~5%) soil moisture conditions, respectively, are discussed in the following sections.

4.1 Retrieval of volumetric soil moisture

Figure 7 summarizes the comparisons of open-loop and assimilated results of daily mean volumetric soil moisture with observations at the surface layer of station ASSH8; simulation results for stations ASSH6 and ASSH7 are summarized in Fig. 8 and Fig. 9, respectively. Mean (AVE) and route mean square errors (RMSE) of temporal variations in surface-layer soil moisture (using observations at 3 cm depth) recorded at the above three stations are summarized in Table 2.

The overall results show that the soil moisture retrieval algorithm using microwave radiometer observations and the variational assimilation scheme quickly bring the state variables on track and produce more satisfactory results than open-loop simulations. Note that soil moisture estimates are also obtained at the two nodes (root and deep layers) below the sur-
face layer, providing a complete soil moisture profile for each pixel. The general patterns at depth are similar to those observed in the surface layer, especially in the root layer. In deep layers, soil moisture is largely invariant at a weekly time-scale; assimilation does not provide a significant improvement at this depth.

The average soil moistures at stations ASSH8 Fig. 7, ASSH6 Fig. 8, and ASSH7 Fig. 9 are 15%, 10%, and 5% respectively, and the open-loop simulations show an initial “drying and recovering” followed by a stable position closer to 10% for the overall simulation period. These simulation trends only show similar results at station ASSH6; the results for this station are markedly different from those of the other two stations (ASSH7 and ASSH8). Variations in soil moisture recorded at ASSH7 and ASSH8 may have been controlled by other physical processes that are impossible to resolve via 1-D simulations; the failure of the 1-D SiB2 (open-loop simulations) application is easily identified where lateral flow effects are most influential.

The assimilation scheme attempted tried to manage this failure because the scheme basically depends on satellite data and the satellite data only reflect the current conditions of the observed field of soil moisture within the assimilation window, which was defined as 24 hours. To fit the low microwave radiation data (higher soil moisture) or vice versa, the assimilation scheme introduces new initial conditions for the model state and produces higher or lower soil moisture integration over the assimilation window. These estimations are more consistent with field observations than those of the open-loop simulation with an a priori (same as the first guesses of the assimilation scheme) initial state.

The assimilation scheme that introduced new initial conditions showed a marked change in soil moisture at the beginning of each assimilation cycle and spurious short-term variability that was not evident in the open loop simulations. This outcome is expected because; (1) defining the B matrix involves a large error, and (2) at present, we only have two satellite images per day (sometimes only one depending on quality control) and these images do not cover all of the diurnal changes in 24-hour assimilations. When the assimilation scheme tries to match the observations by changing the initial conditions, it fails to cover the entirety of the diurnal cycle, and attempts to select the initial conditions to fit only at the observation time zone. This can be improved by increasing the number of microwave observations covering the diurnal changes.
4.2 Spatial distribution of soil moisture and brightness temperature

Figure 10 shows a digital representation of the observed ascending image of AMSR-E brightness temperature (6.925 GHz-H) recorded on July 3, 2002 (Fig. 10a) the estimated brightness temperature (6.925 GHz-H) (Fig. 10b), the volumetric soil moisture of the surface layer (Fig. 10c), and the soil-surface temperature (Fig. 10d) estimated at 14:00 (local time) on July 3, 2002. Compared to the observed brightness temperature, the estimated image of the brightness temperature is slightly different in terms of its spatial distribution.

Although satellite-based radio brightness measurements contain information on land-surface variables, it is clear that they do not tell us much about the detailed spatial structure of the individual grids. As observed brightness temperature is a composite image (e.g., AMSR-E's footprint size is 70 km at 6.9 GHz) of heterogeneous land-surface emission, and the estimated brightness temperature is the emission from each pixel it represents the
actual behavior and emission at the local point. Even the nearby stations ASSH7 and ASSH8 show contrasting soil moisture patterns over the simulation period. These kinds of land surface heterogeneity are currently impossible to identify from the available satellite images because existing spatial resolutions are too coarse. Data assimilation procedures are able to overcome the above critical points by incorporating physically based numerical models and downscaling techniques.

Figure 11 shows a comparison of 6.925 GHz vertical and horizontal polarization brightness temperatures recorded at stations ASSH7 and ASSH8. Station ASSH7 is representative of low soil moisture, whereas station ASSH8 is representative of high soil moisture. At both grids where stations ASSH7 and ASSH8 are located, the LDAS input values of AMSR-E observations are the same because of the coarse spatial resolution of the data; however, optimal estimated brightness temperatures estimated by LDAS are different between the two sites because of local differences in soil moisture, roughness, vegetation, and temperature. Of the two stations, estimations at ASSH8 show better results because of the high soil moisture conditions at this site. In addition, volume scattering effects related to very dry soil moisture conditions affect the estimates of brightness temperature at ASSH7. Future studies should therefore consider fine-tuning of the RTM parameters, especially soil particle size, and the adoption of assumptions that are more physically based.

We also note that both the observed and estimated images show that the northern part of the Mongolian Plateau is dryer than the southern part. This is evident from the soil moisture patterns at station ASSH8 (example of the southern part of the plateau where the average soil moisture is \( \sim 15\% \)) and ASSH6 (example of the northern part of the plateau where the average soil moisture is \( \sim 10\% \)). Our own estimations also depend on other sources of information that describe surface conditions, hydraulic properties, soil texture, and vegetation type; these data are generally poorly available for the Mongolian Plateau. More ground-based experiments, both short- and long-term, should be carried out to address land-surface heterogeneity and improve model parameterization.

If we compare input data, model performance, and correlations, it is clear that the results of the assimilation scheme presented in this study are very encouraging. To support our findings, we estimated error behavior at various spatial and temporal scales. Figure 12 shows temporal changes in the spatially averaged error (AVE) and RMSE of surface-layer soil moisture using 12 ASSH stations (Fig. 5) that are within satellite coverage. Here, the LDAS demonstrates the potential to estimate spatial and temporal variations in soil moisture within 2% of RMSE by incorporating satellite radio brightness images.
5. Conclusions and outlook

Our current understanding of hydrological processes and the precipitation ability of water resources is limited in ungagged basins. To obtain information on climate, soils, topography, and vegetation in addition to precipitation, discharge, evaporation, and soil moisture is a challenging problem, and yet a practical problem for many developing countries. Remote sensing is the only approach that can obtain the required measurements. Extensive amounts of hydrologically relevant remote sensing data are becoming increasingly available, requiring special capabilities for analysis and interpretation.

The use of satellite data for hydrological purposes has been limited by the availability of reliable algorithms to compare measured radiance with required/estimated hydrological variables.

In this work, we developed a more efficient 1-D data assimilation algorithm (LDAS) for hydrological applications and tested the approach for very dry soil conditions at the AMSR-E validation site in Mongolia. The LDAS describes a methodology for generating soil moisture initialization states for global climate models that does not rely on spinning-up the land-surface models. Rather, this methodology relies on the assimilation of remotely sensed passive microwave observations using a 1-D variational optimization approach that accounts for both model and measurement errors.

This LDAS is based on a 1-D land-surface model that describes changes in soil moisture, soil temperature, and canopy temperature over time and spatially in a vertical direction. Estimates of the land surface states are derived directly from AMSR-E passive microwave observations that are related to the states using RTM. The RTM products are highly sensitive to land-surface skin properties for wet conditions, however, for a dry surface, the most important issue is obtaining more accurate estimates of brightness temperature that take into account volume scattering effects. Our present research includes a new model that takes volume scattering effects into accounts.

In the present study, we found that the assimilation of AMSR-E images works well for regions with low dry soil moisture, particularly for soil moistures of less than 10%, but not well as regions with high soil moisture (above 10%). It is likely that land surface estimates provided by our assimilation can be improved by adding other types of measurements such as infrared observations of soil temperature and other microwave measurements. Multiplatform and multi-frequency measurements can be naturally included in the assimilation algorithm if the observation equation (RTM) is modified accordingly.

Much more work is necessary before soil moisture data assimilation becomes practical in an operational setting. Obvious future research directions are a test of the algorithm for field data [currently completed for sample agri-
cultural crops (Pathmathevan et al. 2003b) and different bare soil and roughness conditions (Pathmathevan et al. 2005) and the assimilation of additional types of data. Moreover, this scheme can be modified to a 4DVAR algorithm via coupling with a spatially distributed hydrological model rather than using the 1D-SiB2. This will enable measurement of horizontal water flow and coupling with an atmospheric model to independently organize the forcing data. Furthermore, the error matrices can be determined in a systematic and consistent fashion by considering error propagation through the land surface and radiative transfer models, as well as errors associated with the measurements that take into accounts computational expenses.

Finally, we hope that the approach outlined in the present paper is valuable first step towards the comprehensive use of satellite data for large-scale hydrological simulations. Future research and data will enable a better understanding of the underlying physical processes and hence lead to more sophisticated algorithms that at present seem a distant goal.

References


for Hydrological Modelling in the Tibetan Plateau, Proceedings of the Fifth International Study Conference on GEWEX in Asia and Game, Nagoya, Japan, 470–475.


Lakshmi, V., J. Albertson, and J. Schaake, 2001: Land Surface Hydrology, Meteorology, and Climate Observations and Modelling, Published as Water Science and Application vol. 3, by the American Geophysical Union.


Monteith, J.L., 1973: Principles of Environmental; Physics, Edward Arnold Ltd.


