A microwave land data assimilation system: Scheme and preliminary evaluation over China

Xiangjun Tian, Zhenghui Xie, Aiguo Dai, Binghao Jia, Chunxiang Shi

Received 15 April 2010; revised 10 August 2010; accepted 18 August 2010; published 6 November 2010.

To make use of satellite microwave observations for estimating soil moisture, a dual-pass land data assimilation system (DLDAS) is developed in this paper by incorporating a dual-pass assimilation framework into the Community Land Model version 3 (CLM3). In the DLDAS, the model state (volumetric soil moisture content) and model parameters are jointly optimized using the gridded Advanced Microwave Scanning Radiometer–EOS (AMSR-E) satellite brightness temperature (Tb) data through a radiative transfer model (RTM), which acts as an observation operator to provide a link between the model states and the observational variable (i.e., Tb). The DLDAS embeds a state assimilation pass and a parameter calibration pass. In the assimilation pass, the whole soil moisture profiles are assimilated from the Tb data using an ensemble-based four-dimensional variational assimilation method (En4DVar). Simultaneously, several key parameters in the RTM are also optimized using the ensemble Proper Orthogonal Decomposition-based parameter calibration approach (EnPOD_P) in the parameter optimization pass to account for their high variability or uncertainty. To quantify the impacts of the Tb assimilation on CLM3-calculated soil moisture, the original CLM3 (Sim) and the DLDAS (Ass) were run separately over China on a 0.5° grid forced with identical, observation-based atmospheric forcing from 2004 to 2008. Soil moisture data from 226 stations over China are averaged over seven different climate divisions and compared with the soil moisture from the Sim and Ass runs. It is found that the assimilation of the AMSR-E Tb data through the DLDAS greatly improves the soil moisture content within the top 10 cm with reduced mean biases and enhanced correlations with the station data in all divisions except for southwest China, where the current satellite sensors may have difficulties in measuring soil moisture due to the dense vegetation and complex terrain over this region. The results suggest that the AMSR-E Tb data can be used to improve soil moisture simulations over many regions and the DLDAS is a promising new tool for estimating soil moisture content from satellite Tb data.


1. Introduction

Soil moisture plays a key role in land-atmosphere interactions. For example, soil moisture strongly influences the partitioning of surface radiative heating into sensible and latent heat fluxes and hence the evolution of the lower atmospheric conditions [Dirmeyer et al., 2000]. Some studies [e.g., U. S. National Research Council, 1994] suggest that soil moisture’s effect on the atmosphere is secondary only to that of sea surface temperature (SST) on a global scale and even exceeds SST’s effect over land. This issue is highlighted by Koster et al. [2004] using some elaborately designed numerical experiments, which indicate that there are several specific regions where soil moisture anomalies have large impacts on local precipitation.

Accurate knowledge of spatial and temporal variations of soil moisture is needed for weather forecasts and climate studies [Dirmeyer, 2000]. While in situ measurements can provide estimates of soil moisture for limited spatial locations and time periods, soil moisture contents simulated by land surface models provide continuous fields over the globe. Large efforts have been devoted to create estimates of soil moisture fields using land surface models forced with realistic precipitation and other atmospheric forcing data, such as the Global Soil Wetness Project (http://grads.iges.org/gswp) [Dirmeyer et al., 1999, 2006], the North America...
Land Data Assimilation System (NLDAS) [Mitchell et al., 2004], the Global Land Data Assimilation System (GLDAS) (http://ldas.gsfc.nasa.gov/) [Rodell et al., 2004], and others [e.g., Nijssen et al., 2001; Qian et al., 2006; Guo et al., 2007; Tian et al., 2007, 2008a; Tian and Xie, 2008; Sheffield and Wood, 2008]. Current state-of-the-art land surface models can capture many of the spatial and temporal variations in soil moisture; but model results often contain mean biases and may deviate from the true soil moisture evolution because of uncertainties in model parameters, model physics, and forcing data, and because efforts to assimilate sparse in situ soil moisture measurements [Robock et al., 2000] into the models have made few progresses.

[4] Because of its sensitivity to near-surface soil moisture content [Prigent et al., 2005], low-frequency (<20 GHz) microwave brightness temperature (Tb) seen by satellites has been widely used to estimate soil moisture [e.g., Shi et al., 1997; Vinnikov et al., 1999; Njoku et al., 2003; McCabe et al., 2005; Wen et al., 2005; Gao et al., 2006; Paloscia et al., 2006; Sahoo et al., 2008]. Because most current microwave channels are also sensitive to vegetation coverage and other near surface conditions, the retrieved soil moisture content still contains large errors [Sahoo et al., 2008]. There are also efforts to assimilate the satellite Tb data directly into a land surface model to improve simulation of soil moisture and thus the surface energy budget. Reichle et al. [2001] investigated the feasibility of estimating large-scale soil moisture profiles and related land surface variables from 1.4 GHz Tb measurements using variational data assimilation. Crow and Wood [2003] assessed the potential of assimilating surface Tb data into a TOPMODEL-based land surface model using an ensemble Kalman filter (EnKF) method. Seuffert et al. [2004] tested two different soil moisture systems based on a simplified extended Kalman filter (EKF) method and an optimal interpolation (OI) method using screen-level parameters and 1.4 GHz Tbs data. Recently, Yang et al. [2007] proposed a land surface assimilation system to assimilate Advanced Microwave Scanning Radiometer–EOS (AMSR–E) Tb data, and Tian et al. [2009] described an assimilation framework that can improve the soil moisture profiles considerably by assimilating gridded AMSR–E Tb data even though the satellite observations are only for the skin soil layer.

[5] In order to assimilate the Tb data directly, a radiative transfer model (RTM) needs to be incorporated into a land data assimilation system (LDAS) to provide a link between the forecast model states and observational variable(s) (i.e., Tbs in this case). As a result, the performance of a RTM can largely determine the capability of the LDAS in simulating the surface states. Several RTMs [Wang and Choudhury, 1981; Dobson et al., 1985; Fung, 1994; Weng et al., 2001; Chen, 2003; Shi et al., 2005] have been proposed to quantify the land surface emissivity in relation to soil moisture content over various surface conditions. Nevertheless, how to characterize the high surface heterogeneity adequately in a RTM is still an ongoing research, and large improvements are needed in estimating surface emissivity over different landscapes and hydrological conditions such as snow, desert and densely vegetated areas.

[6] How to fully use the microwave Tb data, which are sensitive to soil moisture content only in the top few centimeters, to improve estimates of soil moisture profiles is a difficult inverse problem. We have addressed this issue in a dual-pass land data assimilation framework [Tian et al., 2009] based on an ensemble-based four-dimensional variational method (En4DVar) [Tian et al., 2008b]. It is superior to another explicit 4DVAR method [Qiu et al., 2007], especially in land data assimilation. Several experiments conducted using this framework coupled partially with a land surface model show that surface soil moisture estimates can be significantly improved by assimilating daily Tb data [Tian et al., 2009]. Furthermore, the improvement also propagates to lower layers where no observations are assimilated. Here, the so-called “coupled partially” means we only use an individual one-dimensional (1-D) soil water model forced by the infiltration derived by a land surface model as the forecast model in the assimilation framework.

[7] In this paper, we implement the dual-pass assimilation framework of Tian et al. [2009] into the Community Land Model version 3 (CLM3) [Oleson et al., 2004; Dickinson et al., 2006] to develop a new dual-pass land data assimilation system (DLDAS). To quantify the impacts of the Tb assimilation on soil moisture simulation, we ran the original CLM3 (referred to as Sim) and the DLDAS (referred to as Ass) separately over China on a 0.5° grid forced with observation-based atmospheric forcing from 2004 to 2008, and compared the simulated soil moisture content with in situ observations in seven different climate divisions over China. It is encouraging to find that the assimilated monthly soil moisture covaries closely (r = 0.5 to 0.8) with in situ observations with reduced root mean square (RMS) errors in all divisions except for southwest China.

2. DLDAS Scheme

[8] Tian et al. [2009] described a dual-pass assimilation framework for assimilating soil moisture profiles from AMSR–E low-frequency (6.9 GHz) Tb data. In that study, the CLM3 was forced with observation-based atmospheric forcing data from Qian et al. [2006] to first derive the infiltration, and ground, surface soil and canopy temperatures for use in the dual-pass assimilation framework. That is, the assimilation framework was not really coupled with the CLM3. In this paper, we further incorporate such a dual-pass assimilation framework into the CLM3 to build the DLDAS. The DLDAS consists of a forecast operator (CLM3) to simulate volumetric soil moisture content, a radiative transfer model (RTM) to estimate microwave Tb from the CLM3-simulated surface conditions, and a dual-pass variational assimilation algorithm to simultaneously optimize the state variable (i.e., soil moisture) and the parameters in the RTM using satellite Tb data. The dual-pass variational assimilation algorithm used here originates from Tian et al. [2009], but with the following significant change: the EnPOD_P approach of Tian et al. [2010] is used in the DLDAS to search for the optimal values of the parameters instead of the SCE-UA used in the original assimilation framework described by Tian et al. [2009]. The EnPOD_P considerably outperforms the SCE-UA largely due to the simultaneous optimization of the model states and parameters [Tian et al., 2010].
2.1. CLM3 and Its Soil Water Hydrodynamic Model

[9] The CLM3 [Oleson et al., 2004; Dickinson et al., 2006] is a land surface model designed for use in coupled climate models. In the CLM3, a nested subgrid hierarchy, in which grid cells are composed of multiple land units, snow/soil columns, and plant functional types (PFTs), is used to represent the spatial heterogeneity of land surface. Each grid cell can have a different number of columns, and each column can have multiple PFTs. The PFTs, which differ in their ecological and hydrological characteristics, are used to represent the vegetation cover. The grid averaged atmospheric forcing is used to force all subgrid units within a grid cell. Biophysical processes simulated by the CLM3 include solar and longwave radiation interactions with vegetation canopy and soil; momentum and turbulent fluxes from canopy and soil; heat transfer in soil and snow; hydrology in the canopy, soil, and snow; and stomatal physiology and photosynthesis. These processes are simulated for each subgrid land unit, column, and PFT independently, and each subgrid land unit maintains its own state variables. The CLM3 has one vegetation layer, 10 unevenly spaced vertical soil layers, and up to 5 snow layers (depending on the total snow depth). It computes soil temperature and soil water content in the 10 soil layers to a depth of 3.43 m in each column.

[10] The volumetric soil moisture (θ) for 1-D vertical water flow in a soil column in the CLM3 is expressed as

\[ \frac{\partial \theta}{\partial t} = - \frac{\partial q}{\partial z} - E - R_{sm}, \]  

(1)

where \( q \) is the vertical soil water flux, \( E \) is the roots' evapotranspiration rate, and \( R_{sm} \) is the melting (negative) or freezing (positive) rate, and \( z \) is the depth from the soil surface. Both \( q \) and \( z \) are positive downward.

[11] The soil water flux \( q \) is described by Darcy's law [Darcy, 1856],

\[ q = -k \frac{\partial (\varphi + z)}{\partial z}, \]  

(2)

where \( k = k_s \left( \frac{\varphi}{\theta_s} \right)^{2b+3} \) is the hydraulic conductivity, and \( \varphi = \varphi_s \left( \frac{\theta}{\theta_s} \right)^{-b} \) is the soil matric potential, \( k_s, \varphi_s, \theta_s \) and \( b \) are constants. The upper boundary condition is

\[ q_0(t) = -k \left. \frac{\partial (\varphi + z)}{\partial z} \right|_{z=0}, \]  

(3)

where \( q_0(t) \) is the water flux at the land surface (referred to as infiltration). The CLM3 computes soil water content in the 10 soil layers through equations (1)–(3). The time step \( \Delta t \) is 30 min.

2.2. Radiative Transfer Model

[12] We used the microwave land emissivity model (LandEm) of Weng et al. [2001] to quantify the land emissivity over various surface conditions. The comprehensive consideration of various surface conditions in the LandEm attracts us to adopt it in our DLDAS. For completeness, this model is briefly described below. A three-layer medium is considered in the LandEm, as shown schematically in Figure 1 of Weng et al. [2001]. The top and bottom layers are considered spatially homogeneous and are represented by uniform dielectric constants. For example, the top layer is the air having a dielectric constant \( \varepsilon_1 \), whereas the bottom layer is characterized by a dielectric constant \( \varepsilon_3 \). Conversely, the middle layer is spatially inhomogeneous and contains scatters such as snow grains, sand particles, and vegetation canopy. Radiative transfer calculations are used to determine the volumetric scattering within the middle layer, while the modified Fresnel equations are used to determine the reflectance at the two interfaces. The radiance (\( I \)) at a given frequency is obtained by solving the following radiative transfer equation:

\[ \mu \frac{dI(\tau, \mu)}{d\tau} = I(\tau, \mu) - \frac{\omega(\tau)}{2} \int_{-1}^{1} P_0(\tau, \mu, \mu') I(\tau, \mu') d\mu' - \left[1 - \omega(\tau)\right] B(T), \]  

(4)

where \( \omega(\tau) \) is the single-scattering albedo, \( P_0(\tau, \mu, \mu') \) is the phase function, \( B(T) \) is the Planck function, \( T \) is the thermal temperature, \( \tau \) is the optical thickness which is parameterized as a function of the leaf area index and transmissivity, \( \mu \) is the cosine of incident zenith angle, and \( \mu' \) is the cosine of scattering zenith angle.

[13] A solution for (4) was derived at arbitrary viewing angles using a two-stream approximation [Weng and Grody, 2000],

\[ \mu \frac{dI(\tau, \mu)}{d\tau} = [1 - \omega(1 - b)] I(\tau, \mu) - \omega b I(\tau, -\mu) - (1 - \omega)b B, \]

(5)

\[ -\mu \frac{dI(\tau, -\mu)}{d\tau} = [1 - \omega(1 - b)] I(\tau, -\mu) - \omega b I(\tau, \mu) - (1 - \omega)b B, \]

(6)

where \( b \) and \( 1 - b \) are the ratios of the integrated scattering energy in the backward and forward directions, respectively. For an isotropic scattering, \( b = 1/2 \), so the scattered energy is the same in both directions. Since \( b \) is generally less than 1/2, forward scattering is much stronger than backward scattering, and the resulting upwelling radiation is reduced.

[14] Equations (5) and (6) can be combined into decoupled second-order differential equations with constant coefficients, assuming that \( \omega, b, \) and \( B \) are independent of \( \tau \). These equations can be used to analyze the scattering from the atmosphere or surface. The upwelling radiance observed from satellites for an ice cloud layer is derived by neglecting reflections at the cloud top and bottom [Weng and Grody, 2000]. However, for surfaces such as snow, the upwelling radiance is modified by the reflectivity and transmissivity at the upper boundary where a discontinuity in the dielectric constant occurs [see Weng et al., 2001, Figure 1]. As a result, the solutions for the upwelling and downwelling radiance are

\[ I(\tau, \mu) = \frac{B_0 \left[ \gamma_1 e^{\theta(T-n)} - \gamma_2 e^{\theta(T-n)} \right] - B \left[ \beta_1 e^{\theta(T-n)} - \beta_2 e^{\theta(T-n)} \right]}{B_1 \gamma_1 e^{\theta(T-n)} - B_2 \gamma_2 e^{\theta(T-n)}}, \]  

(7)

\[ I(\tau, -\mu) = \frac{B_0 \left[ \gamma_1 e^{\theta(T-n)} - \gamma_2 e^{\theta(T-n)} \right] - B \left[ \beta_1 e^{\theta(T-n)} - \beta_2 e^{\theta(T-n)} \right]}{B_1 \gamma_1 e^{\theta(T-n)} - B_2 \gamma_2 e^{\theta(T-n)}}, \]  

(8)
where $\kappa$ is the eigenvalue in solving the differential equations and related to particle optical parameters as shown in the Notation section of Weng et al. [2001]. Also, $I'_1 = I_1 - B(1 - R_{23}); I'_0 = I_0(1 - R_{12}) - B(1 - R_{23})$, where $I_1$ is the upwelling radiance at $\tau = \tau_0$ from the top layer. The parameter $R_y$ is the reflectivity at the interface between the two layers. The other coefficients and their functions contained in (7) and (8) are listed in the Notation section of Weng et al. [2001].

[15] For an isothermal surface where the temperature is the same for the middle and bottom layers (layers 2 and 3), the upwelling radiance emanating from the second layer is

$$I(\tau, \mu) = \frac{I_0[\gamma_1e^{\frac{\alpha}{\beta}(\tau_1 - \tau)} - \gamma_2e^{-\alpha(\tau_1 - \tau)}]}{\beta_1\gamma_4e^{-\gamma_2(\tau_1 - \tau)} - \beta_2\gamma_3e^{\gamma_2(\tau_1 - \tau)}} + B. \quad (9)$$

[16] The interface between layers 1 and 2 can also cause additional reflection for the incident radiation. Thus the downwelling radiance from the first layer is reflected, and the downwelling radiance from the second layer is internally reflected at the interface, so the total radiance is given by

$$I(\tau_0, \mu) = I_0 R_{12}(\mu) + I(\tau_0, \mu_1)[1 - R_{21}(\mu_1)]. \quad (10)$$

where $\mu_1$ is the cosine of the upwelling angle being related to $\mu$ through Snell’s law.

[17] The emissivity of the three-layer medium is defined as the ratio of the total radiance emanating from the medium to the blackbody radiance calculated using the Planck function; that is, $\varepsilon = I_e/B$. As a result,

$$\varepsilon = \alpha R_{12} + (1 - R_{21}) \left\{ \frac{(1 - \beta)[1 + \gamma e^{-2\alpha(\tau_1 - \tau)}]}{(1 - \beta R_{21}) - (\beta - R_{21})\gamma e^{-2\alpha(\tau_1 - \tau)}} \right\}, \quad (11)$$

where $\alpha = I_0/B$, $\beta = (1 - \alpha)(1 + \alpha)$ (see the Notation section of Weng et al. [2001] for parameter $\alpha$), and $\gamma = (\beta - R_{23})/\beta (1 - \beta R_{23})$. Since the reflectivity at the interface depends on polarization, the emissivity derived from (11) is also a function of polarization. According to (11) the most important parameters affecting the emissivity are the optical thickness $(\tau_1, \tau_0)$. Several techniques have been used to compute the optical parameters for the medium (see Weng et al. [2001] for details).

[18] Based on the above assumptions and with the emissivity being estimated from the LandEm (i.e., equation (11)), the microwave brightness temperature ($T_b$) can be computed using

$$T_b = \varepsilon T_s, \quad (12)$$

where $T_s$ is surface skin temperature. Obviously, equation (12) is a simplification of (31) in Weng et al. [2001].

[19] Given the needed parameters, the RTM first estimates the emissivity from (11) and then $T_b$ from (12) using the inputs of surface soil moisture content, surface soil temperature, and canopy temperature. Several parameters (namely, $W_p = ($sigma, $\text{rhib}$), where $\text{sigma}$ is the standard deviation of surface roughness height formed between medium 2 and 3 and $\text{rhib}$ is bulk volume density of the soil) in the RTM significantly affect the output while their values are either highly variable or unavailable. The parameters $\text{sigma}$ and $\text{rhib}$ can both affect dielectric constant and then influence the soil dielectric constant [Dobson et al., 1985; Choudhury and Chang, 1979], which finally affects the surface emissivity $\varepsilon$ in equations (11) and (12). How to obtain accurate values for these parameters is critical for the accuracy of the RTM outputs and thus the performance of the DLDAS. This issue is addressed further below.

### 2.3. Dual-Pass Variational Assimilation Algorithm

[20] Figure 1 shows the flowchart of the dual-pass assimilation algorithm used in the DLDAS. The algorithm is implemented differently in the parameter calibration phase and the pure assimilation phase. Both the state assimilation pass and the parameter optimization pass are used in the parameter calibration phase in each assimilation time window in order to obtain the optimal parameters for the RTM. However, only the state assimilation pass is used in the pure assimilation phase after the parameters are determined during the parameter calibration phase. Both passes assimilate observed brightness temperature of the vertical polarization at a low (6.9 GHz) frequency. Based on the observational and modeling results of Fujii [2005], the vertical polarization is relatively insensitive to vegetation coverage and thus is more desirable than the horizontal polarization. This is used by Yang et al. [2007] and also supported by Ahmed [1995] and our modeling results [Tian et al., 2009]. However, we realize that there are studies [e.g., Prigent et al., 2001] suggesting a different view regarding the vertical polarization as asensi-

$$J(x_0) = (x_0 - x_0)^T B^{-1} (x_0 - x_0) + \sum_{i=1}^{m} [v_i - H_i(x)]^T R_i^{-1} [v_i - H_i(x)]. \quad (13)$$

with the forecast model $M_{0\rightarrow i}(x_0)$ imposed as strong constraints, defined as

$$x_i = M_{0\rightarrow i}(x_0), \quad (14)$$

where the superscript $T$ stands for a transpose, $x_0$ and $x_0$ are the initial value and a background value, respectively, of the state variable $x$ (i.e., soil moisture content), index $i$ denotes the observational time and subscript 0 denotes the initial value, and the matrices $B$ and $R$ are the background and observational error covariances, respectively. In the DLDAS,
the forecast model $M_{0\rightarrow}$ is the soil water hydrodynamic model coupled with the CLM3, and the observation operator $H_i$ is the RTM. Noticeably, CLM3 already includes soil hydrology model. But in work by Tian et al. [2009], we only used the hydrology model (not coupled with CLM3) as the forecast model. The RTM establishes a mapping between the forecast state space (i.e., soil moisture content $\theta$, calculated by the soil water hydrodynamic model) and the observational variable space (satellite $T_b$) in (14)). Based on the 4DVar framework of equations (13)–(14), the $T_b$ data can be assimilated in each time window to compute soil moisture profiles using the En4DVar method (see Tian et al. [2008b] for more details). Here the assimilation time window is one day and the $T_b$ observation frequency is twice daily (that is, $m$ equals 2 in equation (13)).

[22] The assimilated soil moisture content obtained from the state assimilation pass is then passed into the parameter optimization pass for parameter calibration in the same assimilation cycle. In the DLDAS, the parameter calibration is conducted within the following EnPOD_P framework [Tian et al., 2010]: We take the parameter (vector) $W$ as a stochastic variable that propagates according to an identity

Figure 1. Flowchart of the dual-pass variational assimilation system. $T_b$ denotes brightness temperature, and $T_g$ and $T_c$ represent the ground temperature and canopy temperature, respectively.
Figure 2. Locations of the 226 stations (red dots) with in situ soil moisture observations. Also shown (boxes) are the seven subdivisions for regional averaging.

operator. As a result, the parameter $W$ can be calibrated through

$$J(w) = \min_{\|w\| \leq \delta} \sum_{i=1}^{m} [H_i(U_i, W + w) - y_i]^T [H_i(U_i, W + w) - y_i],$$

(15)

where $U_i$ is the vector composed of the input fields provided by the CLM3 such as the surface soil moisture content, soil temperature, etc.; $H_i$ is the observational operator (RTM) and $y_i$ is the observation $(T_b)$ at time $t_i$. The final calibrated parameter $W_0 = W + w_c$.

[23] Generate $N$ random parameter perturbations with the constraint condition $\|w\| \leq \delta$ using the Monte Carlo method and add each perturbation to the background parameter $W$ to produce $N$ parameter fields $W_n$. Compute the observational model (propagator) $y_m(t_i) = H_i(U_i, W_n)$ at time $t_i$ throughout the time window to obtain the state series $Y_n(t_i)$ ($i = 1, \ldots, m$) and then construct the ensembles $X_n$ ($n = 1, 2, \ldots, N$) of the state variable using the perturbed parameters and the perturbed states at the observational times,

$$X_n = (W_n, y_m(t_1), \ldots, y_m(t_m)).$$

(16)

It should be noted that the EnPOD_P used for parameter calibration can optimize both the model states and parameters simultaneously since the ensembles in (16) consist of both the perturbed parameters and states. All the perturbed 4-D fields $X_n$ ($n = 1, 2, \ldots, N$) can expand a finite dimensional space $\Omega(\tilde{X}_1, \ldots, \tilde{X}_N)$. Similarly, the analysis field $\hat{X}_d$ (of the state variable) over the same assimilation time window can also be stored into the following vector:

$$\hat{X}_d = (W_d, y_d(t_1), \ldots, y_d(t_m)).$$

(17)

When the ensemble size $N$ is increased by adding random samples, the ensemble space could cover the analysis vector $\hat{X}_d$; that is, $\hat{X}_d$ is approximately assumed to be embedded in the linear space $\Omega(\tilde{X}_1, \ldots, \tilde{X}_N)$. Subsequently, the optimization problem (15)–(17) can be solved using the Proper Orthogonal Decomposition (POD) technique (see Tian et al., [2009] for more details). The calibration experiments conducted by Tian et al. [2010] demonstrate that our approach is superior to the SCE-UA method [Duan et al., 1993] with improved computation efficiency.

[24] Generally, the DLDAS works as follows: The forecast model (i.e., CLM3) is run first to produce surface variable forecasts (such as soil moisture, ground temperature, soil temperature, and canopy temperature, etc.). These forecast states are then input into the assimilation pass to conduct the assimilation process, in which the LandEM uses the simulated ground temperature and canopy temperature to calculate brightness temperature in order to compare with the observed. After the surface variables are improved in the assimilation pass, some key parameters used in the LandEm are calibrated by the EnPOD_P method in the parameter calibration pass using the brightness temperature observations. However, we have not approved that the DLDAS output is optimal up to now, which should be addressed in the future.

3. Data and Numerical Experiments

[25] In this section, the DLDAS is implemented and evaluated through a regional assimilation experiment over China from 2004 to 2008.

3.1. Data Used for Assimilation and Evaluation

[26] We have obtained data from in situ measurements of soil moisture during 2004–2008 from 226 stations (Figure 2) from the China Meteorological Administration (CMA) operational network through personal communications from J. Zheng (2009). The 226 stations are grouped into seven subdivisions roughly according to the spatial patterns of the dryness and wetness centers in China from Zhu [2003]. The seven subdivisions (see Table 1 and Figure 2 for their definitions) contain 208 of the 226 stations (i.e., 18 of the stations not used in this study). At the 226 stations, soil moisture was measured for three layers from 0 to 10 cm, 10 to 20 cm, and 40 to 50 cm on the 8th, 18th and 28th days of each month and data from January 2004 to December 2008 were used in this study. For the two deeper layers at these stations, the records are too incomplete to be suitable for evaluation.

[27] The AMSR-E $T_b$ data (at 6.9 GHz from vertical (V) polarization, twice daily) from 1 January 2004 to 31 December 2008 used in this study were downloaded from http://wist.echo.nasa.gov/api/ and we then regridded the $T_b$ data onto a 0.25° × 0.25° grid. More information about the AMSR-E data can be found at http://www.ghcc.msfc.nasa.gov/AMSRE.

3.2. Numerical Experiments

[28] To perform our experiments from 2004 to 2008, we first extended the observation-based atmospheric forcing data from Qian et al. [2006] to 2008 using the ERA Interim data (http://data-portal.ecmwf.int/data/d/interim_daily/), which are 6-hourly and have a resolution of 1.5° latitude × 1.5° longitude. The ERA-Interim assimilates surface observations and thus its surface fields (including precipitation and air temperature) are very close to observation-based analyses [Simmons et al., 2010]. We then ran the CLM3 220 years forced with recycled 1948–2004 forcing data from Qian et al. [2006] in order to spin up the deep soil layers. From the state at the
end of this 220 year run, we ran the original CLM3 (Sim) and the DLDAS (Ass) forced with the ERA-Interim forcing data from 2004 to 2008 separately. All of the runs were at the same resolution of 0.5° latitude × 0.5° longitude, and the forcing data were also interpolated onto the 0.5° grid. The assimilation time window in this study is 1 day (48 time steps). The sampling frequency of the satellite $T_b$ data is twice per day.

### 4. Experimental Results

[29] To examine the DLDAS’ performance quantitatively, time series of monthly mean volumetric soil moisture con-

<table>
<thead>
<tr>
<th>Identification</th>
<th>Region Name</th>
<th>Location</th>
<th>Number of Observational Stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>China I</td>
<td>northeast China</td>
<td>120°E–135°E, 40°N–50°N</td>
<td>39</td>
</tr>
<tr>
<td>China II</td>
<td>northern North China</td>
<td>110°E–120°E, 40°N–45°N</td>
<td>13</td>
</tr>
<tr>
<td>China III</td>
<td>southern North China</td>
<td>110°E–120°E, 34°N–40°N</td>
<td>54</td>
</tr>
<tr>
<td>China IV</td>
<td>middle and lower Yangtze River valley</td>
<td>110°E–122°E, 30°N–34°N</td>
<td>28</td>
</tr>
<tr>
<td>China V</td>
<td>eastern northwest China</td>
<td>95°E–110°E, 34°N–42°N</td>
<td>43</td>
</tr>
<tr>
<td>China VI</td>
<td>western northwest China</td>
<td>80°E–95°E, 40°N–50°N</td>
<td>9</td>
</tr>
<tr>
<td>China VII</td>
<td>southwest China</td>
<td>100°E–110°E, 20°N–34°N</td>
<td>22</td>
</tr>
</tbody>
</table>

**Table 1.** Locations of the Seven Subregions in China

**Figure 3.** Time series of monthly volumetric soil moisture (in m$^3$/m$^3$) for the top 10 cm layer from the observations, CLM3 simulation, and LDAS assimilation at regions China I–IV defined in Table 1.
Figure 4. Same as Figure 3 but for regions China V–VII.

Figure 5. (a) Correlation coefficients between the observed and DLDAS-assimilated (black bars) or CLM3-simulated (white bars) monthly soil moisture content shown in Figures 3 and 4. (b) Root-mean-square (RMS) errors (m³/m³) for the assimilated and simulated soil moisture content shown in Figures 3 and 4.
tent extracted from the Ass and Sim runs for the seven subregions in China are compared with their corresponding observations. Area-weighted averaging was used in deriving the Ass and Sim regional time series, whereas simple arithmetic mean of the station series was used for the observational series. Also note that only three reports were used in deriving the monthly mean for the observations while the model monthly mean was derived from complete daily sampling. Figures 3 and 4 compare the time series of assimilated, simulated and observed monthly volumetric soil moisture in the top 10 cm depth from January 2004 to December 2008 for the seven regions. Compared with the Sim case, the assimilated monthly soil moisture is improved considerably and captures the temporal evolution of the observed soil moisture reasonably well with comparable amplitude and seasonal phase for most of the regions, except for China VII. Furthermore, temporal variability is larger in the Ass case than the Sim case. It should be noted that land surface models usually fail to simulate the mean soil moisture content even though they can generally reproduce its anomalies and seasonal variations [Entin et al., 2000; Guo and Dirmeyer, 2006; Qian et al., 2006]. This is also the case for the CLM3-simulated soil moisture: it contains substantial mean biases even though it captures many of the observed variations. The mean bias is greatly reduced in the DLDAS-assimilated soil moisture and the observed temporal evolution is also captured better by the DLDAS in most cases, leading to higher correlations ($r = 0.5$ to $0.8$) and smaller RMS errors (Figure 5). One exception is Region VII (southwest China), where the DLDAS failed to show improvements. This is likely due to the fact that the satellite $T_b$ measurements may be contaminated by the dense vegetation and complex terrain over this region (especially by the complex terrain), which can be inferred from Figure 6. Another possible reason is that the observations are too sparse in this region. Nevertheless, the substantial improvements over most of the regions with different climates and land surface conditions suggest that the DLDAS could provide a promising new tool for land data assimilation.

Figures 7 and 8 compare the 2004–2008 mean annual cycle of volumetric soil moisture for the top 10 cm layer from the observations, CLM3 simulation and DLDAS assimilation for the seven regions. Consistent with previous studies [e.g., Qian et al., 2006], Figures 7 and 8 show that the CLM3 can reproduce many of the observed seasonal variations in soil moisture but with substantial biases in the annual mean for some of the regions. The DLDAS not only
reduces the mean biases but also captures the seasonal variations better for all but region China VII.

5. Summary and Concluding Remarks

[31] In this study, a land data assimilation system has been developed by incorporating a dual-pass assimilation framework into the Community Land Model version 3 (CLM3). In the DLDAS, a radiative transfer model is taken as the observation operator to simulate brightness temperature in order to compare with the observed. The whole assimilation process is divided into two passes: one is for pure assimilation and the other is for parameter calibration. In the assimilation pass, the whole soil moisture profile is assimilated by the En4DVar method by utilizing the brightness temperature measurements. Several key parameters are also

Figure 7. The 2004–2008 mean of monthly volumetric soil moisture (in m$^3$/m$^3$) for the top 10 cm layer from the observations (black bars), CLM3 simulation (white bars), and LDAS assimilation (gray bars) for regions China I–IV.
calibrated in the calibration pass by the EnPOD_P method also using the brightness temperature observations in the same time step.

To quantify the impacts of the T_b assimilation on the simulated soil moisture, we ran the DLDAS and the original CLM3 over China separately forced with identical atmospheric forcing from 2004 to 2008. In situ measurements of soil moisture content from 226 stations over China during 2004–2008 were averaged over seven different climate divisions and compared with the CLM3-simulated (Sim) and DLDAS-assimilated (Ass) soil moisture. It is found that the assimilation of the twice-daily T_b data by the DLDAS improves the calculated soil moisture content substantially compared with the Sim case, with reduced mean biases and increased correlation with the in situ observations. The soil moisture from the Ass case becomes comparable with the in situ observations over most of the divisions (regions China I–VI), except for southwest China, where dense vegetation and complex terrain may contaminate the AMSR-E T_b data as a measure of surface soil moisture. This suggests that more efforts are needed to improve the estimates of land emissivity over densely vegetated areas. Nevertheless, our results demonstrate that the AMSR-E T_b data can be used to improve soil moisture simulations over many regions and the DLDAS provides a promising new tool for estimating soil moisture content from satellite T_b data.

It should be noted that dense vegetation and complex terrain could greatly degrade the final assimilation results due to the poor ability of current RTMs in characterizing high surface heterogeneity. One approach to partially address this issue is through a multimodel method, in which we can incorporate several RTMs into the DLDAS to build a multimodel observation operator. This strategy can provide more accurate brightness temperature forecasts and thus improve the assimilation results.

Acknowledgments. We thank Banghua Yan for constructive discussions on the microwave land emissivity model and Aihui Wang for help with figure drawing. This work was supported by the National High Technology Research and Development Program of China (grant 2009AA12Z129), the National Basic Research Program (grants 2010CB428403 and 2009CB421407), and the National Natural Science Foundation of China (grants 40705035 and 41075076). The National Center for Atmospheric Research is sponsored by the National Science Foundation.


---

A. Dai, National Center for Atmospheric Center, Boulder, CO 80303, USA.

B. Jia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China.

C. Shi, National Satellite Meteorological Center, China Meteorological Administration, Beijing 100081, China.

X. Tian and Z. Xie, ICCES, LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China. (tianxj@mail.iap.ac.cn)