Validation of a Dual-Pass Microwave Land Data Assimilation System for Estimating Surface Soil Moisture in Semiarid Regions

KUN YANG
Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China

TOSHIKO KOIKE
Department of Civil Engineering, The University of Tokyo, Tokyo, Japan

ICHIROW KAIHOTSU
Department of Natural Environmental Sciences, Hiroshima University, Hiroshima, Japan

JUN QIN
Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China

(Manuscript received 16 June 2008, in final form 1 November 2008)

ABSTRACT

This study examines the capability of a new microwave land data assimilation system (LDAS) for estimating soil moisture in semiarid regions, where soil moisture is very heterogeneous. This system assimilates the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) 6.9- and 18.7-GHz brightness temperatures into a land surface model (LSM), with a radiative transfer model as an observation operator. To reduce errors caused by uncertainties of system parameters, the LDAS uses a dual-pass assimilation algorithm, with a calibration pass to estimate major model parameters from satellite data and an assimilation pass to estimate the near-surface soil moisture. Validation data of soil moisture were collected in a Mongolian semiarid region. Results show that (i) the LDAS-estimated soil moistures are comparable to areal averages of in situ measurements, though the measured soil moistures were highly variable from site to site; (ii) the LSM-simulated soil moistures show less biases when the LSM uses LDAS-calibrated parameter values instead of default parameter values, indicating that the satellite-based calibration does contribute to soil moisture estimations; and (iii) compared to the LSM, the LDAS produces more robust and reliable soil moisture when forcing data become worse. The lower sensitivity of the LDAS output to precipitation is particularly encouraging for applying this system to regions where precipitation data are prone to errors.

1. Introduction

Soil moisture is a highly variable parameter in semiarid regions, where strong coupling between soil moisture and precipitation are suggested by Koster et al. (2004). Soil moisture estimation is imperative for hydro-meteorological studies and water resources management (e.g., Pauwels et al. 2001; Wu and Dickinson 2004; de Goncalves et al. 2006), and how to estimate soil moisture with good accuracy is still a hot topic of land hydrology and satellite remote sensing.

Land hydrological modeling is subject to errors in many factors, including meteorological forcing data, parameterization schemes, and model parameters. Accordingly, individual land models may produce quite diverse results, and uncertainties of modeled soil moisture are significant (Entin et al. 1999; Pitman et al. 1999; Schaeke et al. 2004; Yang et al. 2007b). Remote sensing of passive microwave has been widely used to estimate soil moisture (e.g., Jackson 1993; Njoku and Entekhabi 1996; Shi et al. 1997; Njoku et al. 2003; Wen et al. 2003, 2005), as it is strongly affected by near-surface soil moisture but not much by atmospheric state. However,
microwave signals are also affected by other parameters, such as soil texture and surface roughness (Engman 1991); these parameters must be specified for estimating soil moisture from microwave data (Wang 1983; Wigneron et al. 2003; Njoku and Chan 2006). A land data assimilation system (LDAS), which combines a land surface model (LSM) and satellite signals, may provide a promising means for estimating soil moisture by assimilating microwave data to correct modeling errors (Reichle et al. 2001; Bach and Mauser 2003; Crow and Wood 2003; Ni-Meister et al. 2006; Li et al. 2007). Nonetheless, the assimilation technique does not directly aim at eliminating uncertainties caused by specifying parameter values in land models and satellite schemes. The parameter specification is a common and crucial issue for land modeling, remote sensing, and data assimilation.

Another common issue is how to make a full scale–match comparison between satellite-estimated soil moisture and in situ–observed soil moisture. Soil moisture measured at a single point is often representative merely on a limited scale, depending on heterogeneity of soil properties, land cover, and atmospheric conditions. A satellite microwave pixel may cover tens of kilometers, which can be much larger than the representative scale of single-point measurements of soil moisture. Therefore, a scale–match comparison is essential for evaluating a remote sensing retrieval scheme or data assimilation algorithm for soil moisture. For this purpose, some observational networks have been developed (e.g., Jackson et al. 1999; Cosh et al. 2004; Bosch et al. 2006; Bindlish et al. 2006; Vivoni et al. 2008).

In response to the aforementioned two issues, this study evaluates the output of the land data assimilation system of The University of Tokyo (LDAS-UT) against measurements from a soil moisture network in a semiarid region. LDAS-UT was developed by Yang et al. (2007a) and its applications, developed at a coordinated enhanced observing period (CEOP; Koike 2004) Tibet site, have shown potential for improving the surface energy budget estimation. The soil moisture network of interest in this study was established by Kaihatsu (2005) in Mongolia through the Advanced Earth Observing Satellite II (ADEOS-II) Mongolian Plateau Experiment for ground truth (AMPEX) project to collect data for the development and validation of Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) soil moisture retrieval algorithms. AMPEX data have been archived in CEOP and provide a test bed for remote sensing inversion schemes and data assimilation algorithms to estimate areal soil moisture in semiarid regions.

The paper is organized as follows: LDAS-UT is introduced in section 2 [a detailed introduction is presented in Yang et al. (2007a)], and AMPEX data is discussed in section 3. Section 4 presents numerical experiment design, including (i) a control experiment that is driven with Automatic Weather Station (AWS) data, with the results analyzed in section 5; and (ii) several sensitivity studies that are driven with globally available, low-quality datasets, with the results presented in section 6. LDAS-UT estimates are evaluated against observations and its performance is also compared with soil moisture simulations that are not constrained by microwave satellite data. Conclusions and remarks are given in section 7.

2. LDAS-UT

LDAS-UT is a dual-pass land data assimilation system. It assimilates AMSR-E low-frequency brightness temperature ($T_b$) data and its structure is shown in Fig. 1a. The model operator is a LSM, and the observation operator is a radiative transfer model (RTM). The LSM produces near-surface soil moisture ($\theta_{sfc}$), ground temperature ($T_g$), and canopy temperature ($T_c$), which are then fed into the RTM to simulate the brightness temperatures. The difference between simulated $T_b$ ($T_{bp,est}$) and observed $T_b$ ($T_{bp,obs}$) is then minimized by adjusting either system parameters or near-surface soil moisture. The algorithm, cost function, model operator, observation operator, and system parameters are introduced below.

a. Dual-pass assimilation algorithm

Estimating brightness temperatures using the RTM requires the input of several parameters (surface roughness parameters, soil texture, and canopy optical parameters), in addition to surface variables ($\theta_{sfc}, T_g, T_c$). The simulation of these surface variables using the LSM also requires a number of soil and vegetation parameters. The modeled $T_b$ is thus sensitive to some crucial parameters used in the LSM and the RTM, which should be determined before estimating soil moisture. In response, we developed the dual-pass assimilation algorithm. Figure 1b shows the schematic of this algorithm and the detail can be found in Yang et al. (2007a).

Pass 1, the so-called calibration pass, aims at tuning system parameters; pass 2, the so-called assimilation pass, estimates soil moisture. The principle behind this algorithm lies in the system state variables responding to system parameters and to initial near-surface soil moisture at different time scales.

System parameters have a long-term effect on state variables (such as soil moisture); therefore, a long time window (several months or longer) is selected to calibrate
the parameters. It should be noted that the parameter calibration presented in LDAS-UT relies on satellite microwave data instead of in situ observed soil moisture and/or surface temperature; thus, it may have a wide applicability. In principle, pass 1 requires just a single execution because the optimized parameters only include static parameters in the LSM and the RTM. It can be implemented using available high-accuracy data prior to the real-time assimilation of satellite data for soil moisture estimation.

By contrast to the parameters, initial near-surface soil moisture has a short-term effect on the system state variables; therefore, a short time window (~1 day) is selected to estimate their values by minimizing a cost function. This pass is similar to a normal variational data assimilation system.

**b. Cost function**

LDAS-UT assimilates observed brightness temperatures ($T_b$) of the vertical polarization (V) at a lower frequency (6.9 GHz) and at a higher frequency (18.7 GHz). As the lower frequency $T_b$ is much more sensitive to near-surface soil moisture than the higher frequency, their difference is correlated with soil wetness. In pass 1, system parameters are obtained by minimizing a cost function that accounts for the difference between modeled and observed brightness temperatures for a long-term window ($t_{pass1}$; scale of months or longer). The cost function includes an observation error term and a background error term. The observation error term is defined as

$$F_{obs} = \sum_{t=0}^{t_{pass1}} \left[ (T_{b,est} - T_{b,obs})^2 + (T_{b,est} - T_{b,obs}^{18.7V})^2 \right],$$

where the subscript obs denotes the observed value and est is the modeled value.

The background error term in the cost function is related to a soil wetness index (SWI) that is defined by

$$\text{SWI} = \frac{2(T_{b,obg} - T_{b,18.7V})(T_{b,obg} + T_{b,6.9V})}{T_{b,obg} + T_{b,18.7V}},$$

where $T_{b,obg}$ is the observed brightness temperature and $T_{b,18.7V}$ and $T_{b,6.9V}$ are the simulated brightness temperatures at the initial time of each assimilation cycle using the background value of $\theta_{sfc}$ and the renewed value of $\theta_{sfc}$.

**c. Model operator**

The model operator is the second version of the Simple Biosphere Model (SiB2), developed by Sellers et al. (1996). This LSM includes one canopy layer and three soil layers and describes canopy radiative transfer, aerodynamic canopy transfer, and conductance–photosynthetic processes. The model is a typical dual-source model that

---

**FIG. 1.** (a) LDAS-UT system structure and (b) schematic of dual-pass assimilation technique. Here, $T_g$, $T_c$, and $\theta_{sfc}$ are the ground temperature, canopy temperature, and near-surface soil water content, respectively; $T_b$ is the brightness temperature, $F$ the cost function, and $\Delta t$ the data assimilation window; $\Gamma_p$ is soil reflectivity; and $\tau_c$ is the optical thickness of the vegetation. The polarization, observed value, and estimated value are shown by the subscript $p$, obs, and est, respectively (refer to Yang et al. 2007a for details).
parameterizes heat transfer from both the canopy and the ground. To apply the LSM in arid and semiarid regions, the LSM in LDAS-UT has incorporated an aerodynamic scheme for sparse canopy (Watanabe and Kondo 1990) and a heat transfer scheme for bare soil surfaces (Yang et al. 2008). The LSM is driven with meteorological data (pressure, wind, temperature, humidity, downward radiation, and precipitation) and it produces surface state variables and land fluxes.

d. Observation operator

Microwave brightness temperature is given by

\[
T_{bp} = T_g (1 - \Gamma_p) \exp(-\tau_c) + T_s (1 - \omega) [1 - \exp(-\tau_c)][1 + \Gamma_p \exp(-\tau_c)], \tag{3}
\]

where the subscript \( p \) denotes polarization (vertical or horizontal), \( \Gamma_p \) is soil reflectivity, \( \tau_c \) is the optical thickness of the vegetation, and \( \omega \) is the single-scattering albedo of the vegetation.

The soil reflectivity can be calculated using the \( Q-h \) model developed by Wang and Choudhury (1981):

\[
\Gamma_p = [(1 - Q) R_p + Q R_q] \exp(h), \tag{4}
\]

where the subscripts \( p \) and \( q \) denote vertical and horizontal polarization, respectively; \( Q \) and \( h \) are empirically determined surface roughness parameters; and \( R \) is the Fresnel power reflectivity that describes the soil reflectivity of a smooth surface. The reflectivity depends on soil moisture \( \theta_{sfc} \).

The model parameters in Eqs. (3) and (4) are frequency dependent and are given by

\[
h = (ks)^{0.1\cos\theta}, \tag{5}
\]
\[
Q = Q_0 (ks)^{0.705}, \tag{6}
\]
\[
\tau_c = b'(100\lambda)^{3} w_c / \cos\theta, \quad \text{and} \tag{7}
\]
\[
\omega = 0.00083/\lambda, \tag{8}
\]

where \( \lambda \) (m) is the wavelength; \( k \) is the wavenumber defined as \( 2m/\lambda \); \( \theta \) is the view angle of the satellite sensor; \( s \) (m) is the standard deviation of surface roughness height; \( w_c \) (kg m\(^{-2}\)) is the vegetation water content; and \( Q_0, b' \), and \( \chi \) are empirical coefficients.

e. Optimized parameters

In the LSM and the RTM, some parameters are crucial to the estimation of soil moisture but have high spatial variability. They are soil porosity \( \theta_s \), soil hydraulic and thermal parameters, surface roughness parameters \( (s \) and \( Q_0) \), and the vegetation parameter \( b' \).

Their values are optimized by minimizing the cost function in pass 1 using the shuffled complex evolution method (Duan et al. 1993). Two procedures are adopted to reduce the uncertainties of the parameter estimates.

First, soil thermal and hydraulic parameters are not directly estimated; instead, they are converted from soil texture and soil porosity, with the pedotransfer functions listed below (see references in Yang et al. 2007a):

\[
\rho_s c_s = (0.076 + 0.748 \rho_d / \rho_w) \times 10^5 + (4.195 \times 10^5) \theta_{dc}, \tag{9}
\]
\[
\lambda_s = \lambda_d + (\lambda_m - \lambda_d) \exp[0.36(1 - 1/\theta_{dc})], \tag{10}
\]
\[
\lambda_m = 0.5^{0.5}[7.7(0.01 \times \% \text{sand})^{2.0}(1 - 0.01 \times \% \text{sand})]^{1 - \theta}, \tag{11}
\]
\[
\lambda_d = (0.135 \rho_d + 64.7)/(2700 - 0.947 \rho_d), \tag{12}
\]
\[
K_s = 7.0556 \times 10^{-6.884 + 0.0153 \times \% \text{sand}} \text{ m s}^{-1}, \tag{13}
\]
\[
\psi_s = -0.01 \times 10^{(1.88 - 0.0313 \times \% \text{sand})} \text{ m}, \quad \text{and} \tag{14}
\]
\[
b = 2.91 + 0.159 \times \% \text{clay}, \tag{15}
\]

where \( \theta_s \) is the soil porosity; \( \rho_s c_s \) (J K\(^{-1}\) m\(^{-3}\)) is the soil heat capacity; \( \lambda_s \) (W m\(^{-1}\) K\(^{-1}\)) is the thermal conductivity of the soil; \( K_s \) (m s\(^{-1}\)), \( \psi_s \) (m), and \( b \) are the hydraulic parameters of Clapp and Hornberger (1978); \( \rho_d \) (kg m\(^{-3}\)) is the bulk density of a dry soil calculated using \( \rho_d = \rho_s (1 - \theta_s) \) and \( \rho_s = 2650 \text{ kg m}^{-3} \); \( \rho_w = 1000 \text{ kg m}^{-3} \); and \%sand and \%clay are the percentages of sand content and clay content in the soil, respectively.

This procedure not only reduces the number of optimized parameters and thus reduces the computational cost but it also enhances the physical consistency among the parameter values.

Second, a global dataset is used to provide the default values for soil texture (%sand and %clay) and an uncertainty is added to the default values as the upper and lower bounds of the optimized parameters.

3. AMPEX data

As shown in Fig. 2, the CEOP/Mongolia reference site was located in Mandal Govi of Mongolia. The site covers a flat area of 120 km \( \times \) 160 km in a semiarid grassland, where 12 long-term automatic stations for soil hydrology (ASSH) and six automatic weather stations (AWSs) were deployed. At ASSH, soil temperature and moisture were measured at 3 and 10 cm. AWS measurements included wind, temperature, humidity, pressure, and precipitation (see measuring height in Table 1). Among them, data at two AWSs [Tsagaandelger
(TDS) and Choyr (CRS)] were not archived in CEOP datasets. In this study, near-surface soil moisture data used for validation are measurements at 3-cm depth of 12 ASSH and four AWS [Bayantsagaan (BTS), Delgertsogt site (DGS), Deren site (DRS), and Madalgobi site (MGS)], and meteorological data from the four AWSs are used to drive the LDAS-UT. Table 2 shows the annual mean (1 October 2002 to 30 September 2003) of some meteorological parameters, indicating that the climatological conditions are similar in the experimental area.

AMPEX also collected AMSR-E brightness temperature data for the 12 ASSH stations for the observational period, which were resampled from the native resolutions of 56 km for 6.9 GHz and 21 km for 19.7 GHz to a resolution of 0.1°.

4. Design of numerical experiments

This paper investigates a control case and several sensitivity cases. The control case is used to show how LDAS-UT works to improve the soil moisture estimate, while the sensitivity cases are used to show the effect of forcing data on LDAS-UT output. The experiment design is shown in Table 3.

In the control case, LDAS-UT is applied to the 12 ASSH stations. The forcing data is averaged at the four AWSs, but downward radiations are taken from Global

**Table 1.** Measured variables and sensor heights above the surface at AMPEX AWS.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MGS</th>
<th>DRS</th>
<th>DGS</th>
<th>BTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air temperature</td>
<td>1.6</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Relative humidity</td>
<td>1.6</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Wind speed</td>
<td>2.5</td>
<td>2.4</td>
<td>2.45</td>
<td>2.4</td>
</tr>
<tr>
<td>Precipitation</td>
<td>1.15</td>
<td>1.0</td>
<td>1.05</td>
<td>1.0</td>
</tr>
<tr>
<td>Air pressure</td>
<td>1.1</td>
<td>1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
</tbody>
</table>

**Table 2.** Total precipitation (rain, mm), annual mean air temperature ($T_a$, K), and specific humidity ($q_r$, g kg$^{-1}$) at AMPEX AWS for the October 2002–September 2003 period. Asterisk (*) means data for the monsoon season were missing.

<table>
<thead>
<tr>
<th></th>
<th>MGS</th>
<th>DRS</th>
<th>DGS</th>
<th>BTS</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rain</td>
<td>200.2</td>
<td>149.6</td>
<td>188.0</td>
<td>176.8</td>
<td>178.7</td>
</tr>
<tr>
<td>$T_a$</td>
<td>275.1</td>
<td>274.9</td>
<td>274.3</td>
<td>273.1</td>
<td>274.4</td>
</tr>
<tr>
<td>$q_r$</td>
<td>3.1</td>
<td>1.8$^*$</td>
<td>3.5</td>
<td>3.2</td>
<td>3.2</td>
</tr>
</tbody>
</table>
Land Data Assimilation System (GLDAS; Rodell et al. 2004) 1° × 1° data, because they were not measured through AMPEX. The forcing data are applied to all the stations. The input microwave data are the station-dependent 0.1° gridded brightness temperatures of AMSR-E 6.9- and 18.7-GHz vertical polarization. Results of this case will be analyzed in section 5.

There are three sensitivity case studies and the first two cases each have three runs driven by different forcing data. First, LDAS-UT is driven with three independent datasets, that is, AWS data, GLDAS output for CEOP, and Japan Meteorological Agency (JMA) model output for CEOP (hereafter AWS, GLDAS, and JMA, respectively). AWS is identical to that used in the control case, GLDAS has a resolution of 1° × 1°, and JMA has a resolution of about 60 km. The latter two provide their data at the grid nearest to the AMPEX area. GLDAS precipitation and radiation are satellite-derived or satellite/gauge merged, and other meteorological data are derived by atmospheric assimilation; thus, the quality of the GLDAS data should be generally better than JMA output. GLDAS wind speed is not archived in CEOP, and thus AWS wind speed is inserted into GLDAS data. This case study contributes to the sensitivity analysis of LDAS-UT output to sources of

<table>
<thead>
<tr>
<th>Case</th>
<th>Forcing data</th>
<th>Microwave data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>AWS</td>
<td>Station dependent</td>
</tr>
<tr>
<td>Sensitivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 1</td>
<td>AWS; GLDAS; JMA</td>
<td>Station average</td>
</tr>
<tr>
<td>Case 2</td>
<td>AWS, AWS + GLDAS rain;</td>
<td>Station average</td>
</tr>
<tr>
<td></td>
<td>AWS + JMA rain</td>
<td></td>
</tr>
<tr>
<td>Case 3</td>
<td>AWS + no rain</td>
<td>Station average</td>
</tr>
</tbody>
</table>

TABLE 3. Numerical experiment design. For sensitivity studies, case 1 includes three runs driven by AWS, GLDAS, and JMA, respectively; and case 2 also includes three runs driven by AWS, except precipitation data are from AWS, GLDAS, and JMA, respectively. Microwave data are the station-dependent data (0.1°) in the control case and the station-averaged data in the sensitivity cases.

FIG. 3. Daily and accumulated precipitations of AWS observation, GLDAS, and JMA output in the AMPEX area.
forcing data. Second, LDAS-UT is driven with AWS data, except that precipitation is taken from one of AWS, GLDAS, and JMA. This case contributes to uncertainty analyses of LDAS-UT output in response to precipitation data, whose quality is most questionable in many regions. Third, LDAS-UT is driven by AWS forcing, except that precipitation is set to zero. This case investigates the role of microwave signal in the dual-pass LDAS. These cases will be analyzed in section 6. Note that LDAS-UT in all sensitivity cases assimilates the station-averaged data rather than the 0.1° station-dependent data.

Figure 3 shows daily and accumulated precipitations of AWS, GLDAS, and JMA datasets in the AMPEX area. JMA shows positive biases and GLDAS shows negative biases in both precipitation frequencies and accumulated amount. Figure 4 shows the monthly-mean diurnal variations of other meteorological data. Both GLDAS and JMA show general observed patterns, but biases in temperature and humidity are evident.

Other ancillary data (leaf area index and default soil texture) are directly taken from global datasets. The default values of the soil texture and soil parameters are sourced from 1° × 1° International Satellite Land Surface Climatology Project (ISLSCP) Initiative II soil data (Global Soil Data Task Group 2000), and values of vegetation parameters (classification and coverage) are derived from 1° × 1° ISLSCP Initiative II vegetation data (Loveland et al. 2000). Leaf area index data is sourced from Moderate Resolution Imaging Spectroradiometer (MODIS) 0.25° × 0.25° gridded 8-day leaf area index products (Knyazikhin et al. 1999).

Note that the period of interest in this study is from May to September 2003 to avoid modeling errors due to soil freezing and thawing processes, which have not been parameterized in the system.

5. Control experiment

LDAS-UT is applied to each ASSH station, driven by the averaged AWS data and assimilating AMSR-E 0.1° gridded data. The accuracy of LDAS-UT estimates is analyzed against measurements as follows.

Figure 5 compares the observed and estimated soil moisture for five stations (A3, C2, E4, G6, and H7; see their positions in Fig. 2), where soil moisture data were
continuously recorded. It shows that the estimates and observations agree well for some stations (e.g., A3 and C2) but does not for the other stations (e.g., G6 and H7). This is not surprising, because the footprint of AMSR-E 6.9-GHz data is about 56 km, while observed soil moisture values at 3-cm depth of the 16 stations are quite diverse, as indicated in Fig. 6. The comparisons in Fig. 5, therefore, do not match scale. As soil moisture was spatially heterogeneous to a large extent in this region, a comparison of areal averages would be better than a point-to-footprint comparison. As shown in Fig. 7a, the LDAS-UT estimates are agreeable with the observations after spatial averaging. Some observed dramatic changes in soil moisture are also produced by LDAS-UT, indicating that the dual-pass LDAS has successfully estimated the temporal variations of soil moisture in this region. By contrast, Fig. 7b shows that the LSM simulation with default parameter values overestimates soil moisture in general.

To explain how the two passes (i.e., calibration pass and assimilation pass) of LDAS-UT work to improve soil moisture estimations, the LSM was run again with the LDAS-UT–calibrated parameters. Figure 8 shows

![Fig. 5. Comparisons of daily mean near-surface soil volumetric water content between observations and estimations of LDAS-UT for five stations. The day of the year is shown on the x axis.](image)

![Fig. 6. Observed daily mean near-surface soil moisture variations at AMPEX’s 12 ASSH and four AWSs during 30 Apr to 30 Sep 2003.](image)
comparisons between the simulation with default parameters (LSM-def), the one with calibrated parameters (LSM-opt), and LDAS-UT output. As indicated by the error metrics [mean bias error (MBE) and root-mean-square error (RMSE)] in each panel of Fig. 8, LDAS-UT has the best accuracy, and LSM-opt is inferior to LDAS-UT but better than LSM-def. Therefore, the improvement of soil moisture estimations in LDAS-UT is realized through both the parameter calibration and the data assimilation.

6. Sensitivity studies

Areal mean soil moisture can be obtained by assimilating microwave data at each station and then averaging the estimated soil moisture, as in the control case. It can also be obtained by directly assimilating the station-averaged microwave data. Figure 9 shows that these two estimates of areal mean soil moisture are almost identical. This result is consistent with previous studies (Galantowicz et al. 2000; Burke and Simmonds 2003), which indicated that regional mean soil moisture was little affected by subpixel heterogeneity of soil moisture and soil properties. In all sensitivity studies, LDAS-UT assimilates the station-averaged microwave data to obtain areal mean soil moisture, to save the computational time.

a. Case 1: Sensitivity to forcing data sources

LDAS-UT and the LSM are driven with AWS, GLDAS, and JMA data. Figure 10 shows the comparison of daily mean soil water content between the observations, LDAS-UT output, and the LSM output, corresponding to three sets of forcing. It shows a general worse tendency of soil moisture estimates for both LDAS-UT and the LSM, when the forcing data was
changed from AWS and GLDAS to JMA. However, LDAS-UT produces better estimates than the LSM does in any of these three runs, as also indicated by the comparison of the mean biases and root-mean-square errors in Table 4.

b. Case 2: Sensitivity to precipitation

This sensitivity study includes three runs of both LDAS-UT and the LSM, each of which is driven with AWS data, except precipitation data is sourced from one of AWS, GLDAS, and JMA. Unlike what Crosson and Laymon (2002) did (they only changed the precipitation amount), these three sets of data have totally different precipitation amounts and frequency, and thus they might be more reasonable to show the effects of different sources of precipitation data.

Figure 11 shows the temporal variations of near-surface soil moisture estimated from these three runs. The LSM output seems comparably sensitive to positive biases in JMA precipitation and negative biases in GLDAS precipitation (see Fig. 3 for JMA and GLDAS data), while the LDAS-UT output seems more sensitive to positive biases in precipitation than to negative biases in precipitation. This asymmetry of the sensitivity in LDAS-UT output can be related to the nonlinear relationship between soil moisture and brightness temperature. Bright temperature changes more slowly with soil moisture in a drier range than that in a wetter range. Accordingly, the error in the simulated brightness temperature is less, given a negative bias in precipitation data, than that given a positive bias; thus, the assimilation system would take a longer time to reduce the errors in the latter case, and the estimated soil moisture may be more sensitive to positive biases in precipitation. Even so, much lessened is the sensitivity of the LDAS-UT estimated soil water content to both positively and negatively biased precipitation, compared to the LSM estimates. This result demonstrates that LDAS-UT can provide more robust estimates of soil moisture than a land surface model.

c. Case 3: Role of microwave signal

This special case completely removes precipitation from forcing data, and thus soil moisture estimates is
mainly controlled by the satellite data. This case is still different from common microwave remote sensing, because the surface roughness parameters for moisture retrieval is calibrated by the dual-pass assimilation algorithm rather than tuned with in situ observations or specified as default values.

Figure 12 shows LDAS-UT–estimated soil moisture compared against observations. Surprisingly, LDAS-UT produces fairly good estimates of soil moisture, though precipitation values have been set to zero. This, in turn, indicates that soil moisture can be estimated with good accuracy from 6.9- and 18.7-GHz $T_b$, given proper parameter values; therefore, the smaller uncertainties of LDAS-UT in sections 6a,b can be attributed to the constraint of the microwave signals.

d. Evaluation of optimized parameter values

Table 5 shows the observed values and the optimized values of soil parameters. All the assimilation cases have produced similar parameter values. The soil porosity value ($\theta_s$) and soil water potential at saturation ($\psi_s$) in all the cases are quite comparable to the observation averaged over the individual sites, but the estimates of the soil hydraulic conductivity ($K_s$) are one order of magnitude lower than the observed one. The estimates of the pore size distribution index ($b$) are also much higher than the observed one. Therefore, the optimized values of $K_s$ and $b$ are tuned values rather than physical ones. The results suggest that the near-surface soil moisture retrieved from the satellite data is not sufficient to physically estimate all the soil parameters, but it is possible to estimate the most sensitive parameters, such as the soil porosity.

Table 6 shows the estimates of the surface roughness parameters $s$, $Q_0$, and vegetation parameter $b'$. The estimates of parameters $Q_0$ and $b'$ are relatively stable in various cases, but the value of $s$, which represents the standard deviation of surface roughness height, nearly decreases with respect to the increase of the precipitation amount in forcing data, except in the case of “AWS + JMA rain.” Brightness temperature is a decreasing function of surface roughness. For a given value of brightness temperature, near-surface soil moisture is thus an increasing function of the surface roughness. In the cases where precipitation data are not used or lower than the observation, the calibration pass would produce an increase in the surface roughness parameter $s$ to make the estimated near-surface soil moisture increase. Therefore, the estimated parameter values in Table 6 are well accordant with the theoretical analysis. The exceptional case (AWS + JMA rain) does not produce a lower value of $s$ (0.252 mm) than other cases, though it uses the maximum precipitation amount. Possibly, other forcing data also exert effects on the parameter estimations, and this needs further investigation in future studies.

7. Conclusions and remarks

Based on data collected at 16 stations in a Mongolian semiarid region, this paper compares the capability of soil moisture estimation by a dual-pass (calibration + assimilation) microwave land data assimilation system (LDAS-UT) and a land surface model as well as their sensitivity to forcing data. Major findings from our study are listed below:

1) LDAS-UT produces good estimates of spatially averaged near-surface soil moisture in this semiarid climate.
region. This performance is attributed to both satellite-based calibration pass and assimilation pass. We also confirm that it is risky to evaluate microwave-derived soil moisture based on measurements at a single site or a sparse network because of the spatial heterogeneity of soil moisture.

2) Compared to the land surface model, LDAS-UT–estimated soil moisture is not sensitive to forcing data. This indicates that the constraint of microwave signal helps lessen uncertainties in land hydrological modeling.

3) Particularly, LDAS-UT produces more robust and reliable soil moisture than a land surface model does when forced by different sources of precipitation data. This result is encouraging and would contribute to hydrometeorological studies for remote regions—such as the Tibet Plateau and the Mongolia Plateau—where globally available datasets of precipitation are usually prone to errors due to the lack of observations.

Previous studies have shown that specifying effective model parameters are crucial for land hydrological models, and many of them have presented sophisticated methods for parameter calibrations; however, they usually have limited applications because their calibrations are based on in situ data (e.g., Merlin et al. 2006; Judge et al. 2008). This paper presents the potential to estimate the effect of parameter values using satellite data rather than in situ data. This is particularly important for the establishment of an operational hydrological model, a remote sensing retrieval scheme, or a land data assimilation system; therefore, this strategy should be pursued in future studies, even though the estimated parameter values might be effective values rather than physical values.

Finally, it should be mentioned that this study mainly addresses how the parameter calibration may improve estimates of soil moisture, but it is possible to further improve the performance of LDAS-UT by introducing advanced assimilation methods, such as an ensemble Kalman filter or particle filters, and this issue should be addressed in future studies.

Acknowledgments. This work was supported by the 100-talent program of Chinese Academy of Sciences and the EU’s Seventh Research Framework Programme “Coordinated Asia–European Long-Term Observing

Table 5. The observed and estimated soil parameter values [see Eqs. (11) and (13)–(15) for the meaning of the parameters] in the three sensitivity cases introduced in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>$\theta_s$</th>
<th>Sand%</th>
<th>Clay%</th>
<th>$b$</th>
<th>$\psi_s$ (m)</th>
<th>$K_s$ ($10^{-6}$ m s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGS*</td>
<td>0.334</td>
<td>—</td>
<td>—</td>
<td>3.37</td>
<td>—</td>
<td>18.7</td>
</tr>
<tr>
<td>MGS**</td>
<td>0.388</td>
<td>—</td>
<td>—</td>
<td>3.62</td>
<td>−0.12</td>
<td>51.9</td>
</tr>
<tr>
<td>DRS*</td>
<td>0.393</td>
<td>—</td>
<td>—</td>
<td>3.06</td>
<td>—</td>
<td>44.0</td>
</tr>
<tr>
<td>DRS**</td>
<td>0.428</td>
<td>—</td>
<td>—</td>
<td>3.00</td>
<td>−0.11</td>
<td>88.6</td>
</tr>
<tr>
<td>BTS**</td>
<td>0.456</td>
<td>—</td>
<td>—</td>
<td>2.93</td>
<td>—</td>
<td>49.3</td>
</tr>
<tr>
<td>E4*</td>
<td>0.358</td>
<td>—</td>
<td>—</td>
<td>2.70</td>
<td>—</td>
<td>49.3</td>
</tr>
<tr>
<td>G6*</td>
<td>0.341</td>
<td>—</td>
<td>—</td>
<td>3.10</td>
<td>—</td>
<td>19.3</td>
</tr>
<tr>
<td>C2*</td>
<td>0.299</td>
<td>—</td>
<td>—</td>
<td>2.82</td>
<td>—</td>
<td>28.5</td>
</tr>
<tr>
<td>Avg</td>
<td>0.377</td>
<td>—</td>
<td>—</td>
<td>3.08</td>
<td>−0.115</td>
<td>36.7</td>
</tr>
<tr>
<td><strong>LDAS-UT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AWS</td>
<td>0.368</td>
<td>47</td>
<td>28</td>
<td>7.34</td>
<td>−0.18</td>
<td>4.9</td>
</tr>
<tr>
<td>GLDAS</td>
<td>0.355</td>
<td>47</td>
<td>33</td>
<td>8.14</td>
<td>−0.18</td>
<td>4.9</td>
</tr>
<tr>
<td>JMA</td>
<td>0.348</td>
<td>38</td>
<td>30</td>
<td>7.75</td>
<td>−0.24</td>
<td>3.5</td>
</tr>
<tr>
<td>AWS + no rain</td>
<td>0.369</td>
<td>46</td>
<td>32</td>
<td>8.07</td>
<td>−0.19</td>
<td>4.6</td>
</tr>
<tr>
<td>AWS + GLDAS rain</td>
<td>0.361</td>
<td>47</td>
<td>32</td>
<td>7.99</td>
<td>−0.19</td>
<td>4.8</td>
</tr>
<tr>
<td>AWS + JMA rain</td>
<td>0.349</td>
<td>43</td>
<td>32</td>
<td>8.05</td>
<td>−0.21</td>
<td>4.1</td>
</tr>
</tbody>
</table>

* Data from Table 1 of Yamanaka et al. (2005).
** Data from Table 1-2 and Table 1-3 of appendix 1 in Kaihotsu (2005).

Table 6. Estimates of surface roughness parameters and vegetation parameter [see Eqs. (5)–(7) for the meaning of the parameters] in the sensitivity cases, which are arranged in ascending order of total precipitation.

<table>
<thead>
<tr>
<th>Forcing data</th>
<th>Rain (mm)</th>
<th>$s$ (mm)</th>
<th>$Q_0$</th>
<th>$b'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWS + no rain</td>
<td>0</td>
<td>0.315</td>
<td>0.94</td>
<td>3.73</td>
</tr>
<tr>
<td>GLDAS</td>
<td>135</td>
<td>0.214</td>
<td>0.98</td>
<td>3.70</td>
</tr>
<tr>
<td>AWS + GLDAS rain</td>
<td>135</td>
<td>0.209</td>
<td>0.99</td>
<td>3.94</td>
</tr>
<tr>
<td>AWS</td>
<td>173</td>
<td>0.124</td>
<td>0.97</td>
<td>3.98</td>
</tr>
<tr>
<td>JMA</td>
<td>238</td>
<td>0.093</td>
<td>0.80</td>
<td>3.87</td>
</tr>
<tr>
<td>AWS + JMA rain</td>
<td>238</td>
<td>0.252</td>
<td>0.71</td>
<td>3.95</td>
</tr>
</tbody>
</table>
System of Qinghai–Tibetan Plateau Hydrometeorological Processes and the Asian–Monsoon System with Ground Satellite Image Data and Numerical Simulations” (CEOP–AEGIS). AMPEX was implemented in the framework of NASA/JRA’s “Ground Truth for Evaluation of Soil Moisture and Geophysical/Vegetation Parameters Related to Ground Surface Conditions with AMSR and GLI in the Mongolian Plateau” (PI: Prof. Ichiro Kaihotsu of the University of Hiroshima). The Mongolian partnership is with the Institute of Meteorology and Hydrology, and the National Agency for Meteorology, Hydrology and Environment Monitoring of Mongolia. The authors thank the editor and the reviewers for their contributions to the analysis in this manuscript.

REFERENCES


Pitman, A. J., and Coauthors, 1999: Key results and implications from phase 1(c) of the project for intercomparison of land-surface parameterization schemes. Climate Dyn., 15, 673–684.


