Assimilation of microwave brightness temperature in a land data assimilation system with multi-observation operators

Binghao Jia,[1] Xiangjun Tian,[1] Zhenghui Xie,[1] Jianguo Liu,[1] and Chunxiang Shi[2]

Received 6 December 2012; revised 5 March 2013; accepted 28 March 2013; published 20 May 2013.

[1] A radiative transfer model (RTM) that provides a link between model states and satellite observations (e.g., brightness temperature) can act as an observation operator in land data assimilation to directly assimilate brightness temperatures. In this study, a microwave Land Data Assimilation System (LDAS) was developed with three RTMs (The radiative transfer model for bare field (QH), land emissivity model (LandEM), and Community Microwave Emission Model (CMEM)) as its multi-observation operators (LDAS-MO). Assimilation experiments using the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) satellite brightness temperature data from July 2005 to December 2008 were then conducted to investigate the impact of the RTMs on the assimilated results over China. It was found that the assimilated volumetric soil-water content using each of the three observation operators improved the estimation of soil moisture content in the top soil layer (0–10 cm), with reduced root mean square errors (RMSEs), and increased correlation coefficients with field observations (OBS) as compared to a control run with no assimilation for the absence of frozen or snow-covered conditions. The assimilated soil moisture for the QH model, which was more sensitive to dry soil than the other models, produced closer correlations with OBS in arid and semi-arid regions while smaller RMSEs were observed for LandEM. CMEM agreed most closely with OBS over the middle and lower reaches of the Yangtze River due to its better simulation of the brightness temperature over densely vegetated areas. To improve assimilation accuracy, a Bayesian model averaging (BMA) scheme for the LDAS-MO was developed. The BMA scheme was found to significantly enhance assimilation capability producing the soil moisture analysis, showing the lowest RMSEs and highest correlations with OBS over all areas. It was demonstrated that the BMA scheme with LDAS-MO has the potential to estimate soil moisture with high accuracy.

Citation: Jia, B., X. Tian, Z. Xie, J. Liu, and C. Shi (2013), Assimilation of microwave brightness temperature in a land data assimilation system with multi-observation operators, *J. Geophys. Res. Atmos.*, *118*, 3972–3985, doi:10.1002/jgrd.50377.

1. Introduction

[2] Soil moisture is one of the most important variables in hydrological, climatological, biological, and ecological processes and plays a crucial role in controlling the exchange of water and heat energy between the land surface and the atmosphere [*Koster and GLACE team*, 2004]. Due to the spatiotemporal limitations of ground observations and the large uncertainties associated with atmospheric forcing and land surface parameterizations, land data assimilation has become an effective way of synthesizing complementary information from measurements and land surface models (LSMs) into a superior estimate of geophysical fields of interest (e.g., soil moisture) [*Reichle et al.*, 2008; *Pan et al.*, 2009; *Draper et al.*, 2012]. The measurements include microwave brightness temperature (T_b) obtained from satellite sensors. Because of its sensitivity to surface soil moisture, low-frequency T_b has the potential advantage of increasing the coverage of soil moisture measurements [*Njoku et al.*, 2003; *Zhan et al.*, 2006]. In order to directly assimilate T_b data into an LSM, a radiative transfer model (RTM) is needed as an observation operator that provides a link between the model states and the satellite-observed T_b data [*Yang et al.*, 2007; *Huang et al.*, 2008; *Jia et al.*, 2009].

[3] *Tian et al.* [2010b] developed a dual-pass microwave land data assimilation system (DLDAS) to assimilate the T_b data from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) using a microwave

[1]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China.
[2]National Meteorological Information Center, China Meteorological Administration, Beijing, China.

Corresponding author: Z. Xie, State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, P.O. Box 9804, Beijing 100029, China, http://www.lasg.ac.cn/staff/xie/xie.htm. (zxie@lasg.iap.ac.cn)

2169-897X/13/10.1002/jgrd.50377

land emissivity model (LandEM) [*Weng et al*., 2001] as the observation operator. A preliminary evaluation over China demonstrated that this system greatly improved the soil moisture content estimation for the top 10 cm compared to the model simulation without assimilation. It was found, however, that a single observation operator may degrade the final analysis due to its large uncertainties.

[4] A multi-model ensemble strategy was seen as a means of exploiting the diversity of skillful predictions produced by different models. In this study, two other RTMs—one from *Yang et al*. [2007] (hereafter QH), and the Community Microwave Emission Model (CMEM) from the European Centre for Medium-Range Weather Forecasts (ECMWF) [*Drusch et al*., 2009]—were incorporated into the DLDAS [*Tian et al*., 2010b] to develop a land data assimilation system with multi-observation operators. Different from previous studies related to land data assimilation [*Yang et al*., 2007; *Jia et al*., 2009; *Tian et al*., 2010b], this paper mainly focuses on the impact of the different observation operators (RTMs) on the assimilated soil-water content, and then on the use of the Bayesian model averaging (BMA) method of combining forecasts from different models [*Raftery et al*., 2005; *Duan et al*., 2007] to enhance the assimilation capability of the new system.

2. A Land Data Assimilation System With Multi-Observation Operators (LDAS-MO)

[5] In this study, a microwave Land Data Assimilation System (LDAS) was developed with the three RTMs (QH, LandEM, and CMEM) as its multi-observation operators (LDAS-MO), which is described in this section. It consists of a forecast operator (the soil-water hydrodynamic model in the National Center for Atmospheric Research (NCAR) Community Land Model version 3 (CLM3) [*Oleson et al*., 2004]) to estimate volumetric soil moisture content, three observation operators (LandEM, CMEM, and QH) to simulate microwave T_b data, and a dual-pass variational assimilation scheme to simultaneously optimize the state variable (e.g., soil moisture) and the parameters in the RTM.

2.1. The Community Land Model CLM3

[6] The CLM3 [*Oleson et al*., 2004] is a global land surface model designed for use as the land component of the Community Climate System Model. In the CLM3, spatial heterogeneity of land surface is represented as a nested subgrid hierarchy in which grid cells are composed of multiple land units, snow/soil columns, and plant functional types (PFTs). Each grid cell can have a different number of land units assigned to it, including glacier, lake, wetland, urban, and vegetated. The second subgrid level, the column, as the component of land unit, is intended to capture potential variability in the soil and snow state variables within a single land unit. Each column can have multiple PFTs. Up to four of the 15 possible PFTs that differ in physiology and structure may coexist on a single column. The biogeophysical processes in the CLM3 are simulated for each subgrid land unit, column, and PFT independently, and each subgrid unit maintains its own prognostic variables. The CLM3 has one vegetation layer, 10 soil layers with different thicknesses, and up to five snow layers (depending on snow depth). The soil temperature and soil-water content are computed by the CLM3 for 10 soil layers to a depth of 3.43 m in each column.

[7] The soil-water model of the CLM3 is taken as the forecast operator, which is expressed as

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} - E - R_{\mathrm{fm}}, \tag{1}$$

where θ is the volumetric soil moisture content (cm^3/cm^3), q is the vertical soil-water flux, E is the root evapotranspiration rate, R_{fm} is the melting (negative) or freezing (positive) rate, and z is the depth from the soil surface. Both q and z are positive downward.

[8] The soil-water flux q is defined by Darcy's law [*Darcy*, 1856]:

$$q = -k\frac{\partial(\psi + z)}{\partial z}, \tag{2}$$

where $k = k_s\left(\frac{\theta}{\theta_s}\right)^{2b+3}$ is the hydraulic conductivity (mm/s), $\psi = \psi_s\left(\frac{\theta}{\theta_s}\right)^{-b}$ is the soil hydraulic potential, and k_s, ψ_s, θ_s, and b are constants. The upper boundary condition is

$$q_0(t) = -k\frac{\partial(\psi + z)}{\partial z}\Big|_0, \tag{3}$$

where $q_0(t)$ is the water flux at the land surface (referred to as infiltration). The volumetric soil moisture content θ is estimated by the CLM3 using equations (1)–(3). In this study, the time step Δt is 0.5 h.

2.2. Radiative Transfer Models

[9] Three radiative transfer models were taken as the observation operators to provide separate links between the model states and observational variables to estimate T_b in LDAS-MO, separately. The first observation operator of LDAS-MO was LandEM [*Weng et al*., 2001] used by *Tian et al*. [2010b], in which the emission and scattering process is characterized by the two-stream approximation. A three-layer medium is considered in the LandEM model, where the top (i.e., air) and bottom (i.e., soil) layers are considered spatially homogeneous and are represented by uniform dielectric constants. By contrast, the middle layer is spatially heterogeneous and contains scatterers such as snow grains, sand particles, and vegetation canopy.

[10] The second operator was the CMEM model developed by ECMWF to simulate low-frequency (1–20 GHz) passive microwave brightness temperatures of the land surface [*Drusch et al*., 2009]. CMEM comprises four modules for computing the contributions from soil, vegetation, snow and atmosphere, with separate parameterizations for each. *Jones et al*. [2004] pointed out that vegetation water content and soil moisture were the main contributors affecting the T_b simulation. An ensemble of 96 simulations for eight land surface models coupled to 12 different configurations of CMEM conducted by *de Rosnay et al*. [2009] suggested that the best CMEM configuration (higher correlation coefficient and smaller centered root mean square error) was the one with the soil dielectric model from *Wang and Schmugge* [1980] and the vegetation opacity model from *Kirdyashev et al.* [1979]. Therefore, the best configuration from *de Rosnay et al.* [2009] and *Jia and Xie* [2011] was used in this study, including the snow emission model from *Pulliainen et al*. [1999] and the atmospheric opacity model from *Pellarin et al*. [2003]

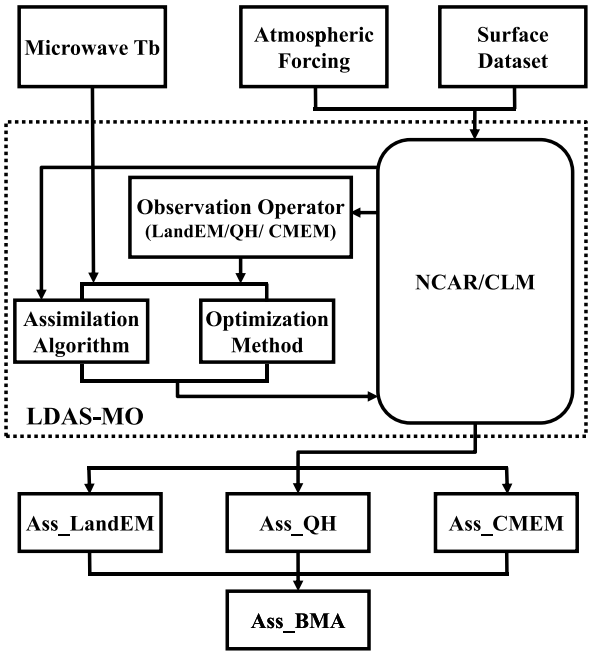

Figure 1. Diagram of the microwave land data assimilation system with multi-observation operators (LDAS-MO) and its input and output.

besides the vegetation opacity model and soil dielectric model mentioned above.

[11] The third operator was the QH model from *Yang et al*. [2007], in which the soil dielectric model was from *Dobson et al*. [1985] and the vegetation opacity model from *Jackson and Schmugge* [1991]. As the effect of snow was ignored in *Yang et al*. [2007], we added the homogeneous snow model from *Pulliainen et al*. [1999], which was the same as that in CMEM.

[12] Given the required parameters, the RTM estimates microwave T_b from the input fields provided by the CLM3, such as surface soil moisture content, surface soil temperature, and canopy temperature. However, several parameters in the RTM significantly affect the output, while their values are either highly variable or unavailable. For example, three key parameters in the LandEM model-standard deviation of surface roughness height (σ), soil bulk volume density (ρ_{soil}), and leaf thickness (d, not used by *Tian et al*. [2010b]) are hard to obtain but play an important role on the microwave T_b estimation, as the first two parameters σ and ρ_{soil} mainly influence soil emissivity, and d affects vegetation optical thickness; in CMEM, the vegetation structure parameter a_{geo} [*Kirdyashev et al*., 1979], which affects vegetation optical thickness, was chosen in preference to leaf thickness d, and the other two parameters σ and ρ_{soil} were the same as those in LandEM; for QH, we chose the parameters σ, Q_0, b' used by *Jia et al*. [2009], where Q_0 is an empirical coefficient used to calculate the surface roughness parameter, and b' is an empirical coefficient used to calculate vegetation optical thickness (see equations (10) and (11) in *Yang et al*. [2007] for details). How to obtain accurate values of these parameters is critical for the accuracy of the RTM's outputs and thus the performance of the LDAS-MO. In this study, we used an optimization method in LDAS-MO to search for the optical values of the RTM parameters mentioned above. This issue is presented below.

2.3. A Land Data Assimilation System With Multi-Observation Operators (LDAS-MO)

[13] In this work, we developed a land data assimilation system with multi-observation operators (LDAS-MO) by implementing the three RTMs mentioned above and a dual-pass variational data assimilation scheme into the land surface model CLM3. The diagram of LDAS-MO and its input and output is shown in Figure 1. The dual-pass variational data assimilation scheme was developed based on an ensemble-based four-dimensional variational (En4DVAR) assimilation algorithm [*Tian et al*., 2008] and an Ensemble Proper Orthogonal Decomposition-based Parameter (EnPOD_P) optimization method [*Tian et al*., 2010a] using satellite-observed T_b data. The whole assimilation process includes two phases: the parameter optimization phases and the pure state assimilation phases, as shown in Figure 1 of *Tian et al*. [2010b].

[14] In the state assimilation pass, random perturbation fields for soil moisture were generated by the Monte Carlo method, and each of the perturbation field was then added to the initial background field (soil moisture) to construct the ensemble members. The similar method was applied to the AMSR-E satellite-observed brightness temperatures to produce observational ensemble. In the parameter optimization pass, we first assigned the initial values of the key parameters in the RTMs randomly in their value ranges at the start of each assimilation time window and then used the EnPOD_P method [*Tian et al*., 2010a] to search for their optimal values. It should be noted that the soil bulk volume density (ρ_{soil}) used in both LandEM and CMEM is also a parameter in the CLM3. Once the optimization process ends, the optimized value will be used in the CLM3. In other words, the LDAS-MO could jointly optimize the model state (volumetric soil moisture content) and model parameters (e.g., soil bulk volume density) [*Yang et al*., 2007; *Tian et al*., 2010b]. The initial values and ranges of each parameter will be shown in the discussion section, and the role of the parameter optimization in the assessment of the RTMs will also be provided below.

[15] Generally, the LDAS-MO works as follows: the forecast model (NCAR/CLM) is run first to produce surface-variable forecasts (e.g., soil moisture, ground temperature, canopy temperature, soil temperature, and snow depth). These forecast states are then input into the assimilation pass to conduct the assimilation process by the En4DVAR algorithm, in which the observation operator (QH, LandEM, or CMEM) uses the simulated surface variables from the land surface model and other auxiliary data sources (e.g., soil texture, fractional cover of vegetation types, and leaf area index) from the surface data set of the CLM3 to calculate T_b in order to compare with the satellite-observed values. After the surface variables are improved in the assimilation pass, the chosen key parameters in the RTM are optimized by the EnPOD_P method in the parameter calibration pass using the microwave T_b observations. A more detailed description of the assimilation pass and optimization pass can be found in *Tian et al*. [2010b].

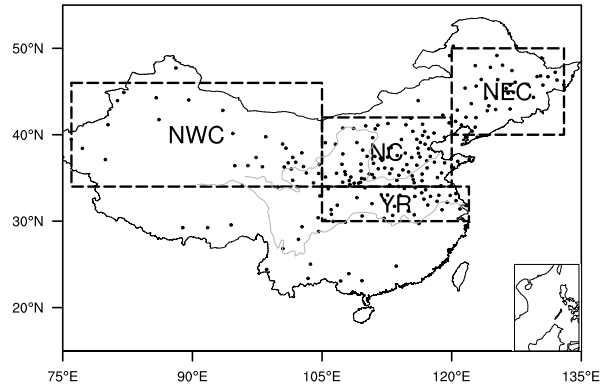

Figure 2. Locations of 226 stations (black dots) with in situ soil moisture observations. Also shown (boxes) are the four sub-regions in China: north-eastern China (NEC, 38), northern China (NC, 88), north-western China (NWC, 36), and middle and lower reaches of the Yangtze River (YR, 31), where the numbers in the parentheses stand for the numbers of observational stations.

3. Assimilation of Microwave Brightness Temperatures in LDAS-MO

3.1. Experiment Design and Verification Data

[16] To investigate the impact of the RTMs on the assimilated results, we conducted one simulation experiment by the original CLM3 (control run, CTL), and three assimilation experiments by the LDAS-MO using the AMSR-E satellite T_b data, with three different observation operators (LandEM, CMEM, and QH). The three assimilation experiments related to LandEM, CMEM, and QH are referred to here as ASS_LandEM, ASS_CMEM, and ASS_QH, respectively.

[17] Here, the AMSR-E gridded C-band (6.9 GHz) microwave brightness temperature data at vertical polarization (twice daily) from 1 July 2005 to 31 December 2008 produced by the National Snow and Ice Data Center (NSIDC), which were downloaded from http://nsidc.org/data/nsidc-0302.html, were used as the satellite observations to be assimilated to the LDAS-MO. Based on the observational and modeling results of *Fujii* [2005], the vertical polarization is relatively insensitive to vegetation coverage and thus is more reliable than the horizontal polarization. This has been used by several studies [*Yang et al*., 2007; *Jia et al*., 2009; *de Rosnay et al*., 2009] and is also supported by the modeling results from *Jia and Xie* [2011]. The NSIDC produced AMSR-E gridded T_b data (0.25°) by interpolating AMSR-E/Aqua L2A swath spatially-resampled brightness temperature data to the output grids using an inverse-distance squared method. To match the spatial resolution of AMSR-E data, all the experiments were run for China (15°N–55°N, 75°E–135°E) in 0.25° increments and half-hour time steps.

[18] The experiments were all forced by the same atmospheric forcing data set with a high temporal-spatial resolution (hourly, 0.1° × 0.1°) [*Shi*, 2008; *Shi et al*., 2011] from 1 July 2005 to 31 December 2008. The precipitation and surface incident solar radiation data in the forcing data set of *Shi* [2008] were derived from the Fengyun-2C satellite retrievals by integrating observed precipitation from about 2000 automated weather stations in the China Meteorological Administration (CMA) operational network; and other atmospheric forcing fields, such as surface air temperature, relative humidity, surface wind speed and air pressure, were interpolated directly from the 6-hourly NCEP-NCAR reanalysis data.

[19] The surface data set (soil and vegetation parameters) required by the CLM3 was chosen the same as that used by CLM3.5 [*Oleson et al*., 2007] as the new one could better reproduce the physical properties of land surface [*Lawrence and Chase*, 2007]. The soil data (e.g., soil color and soil texture) were derived from the International Geosphere-Biosphere Programme soil data set [*Bonan et al*., 2002], with vegetation parameters (e.g., the fractional cover of the plant functional types (PFTs); monthly leaf and stem area index (LAI, SAI); and canopy top and bottom heights for each PFT) derived from Moderate Resolution Imaging Spectroradiometer (MODIS) imagery [*Lawrence and Chase*, 2007].

[20] To perform our experiments from July 2005 to December 2008, we first extended the observation-based atmospheric forcing data from *Qian et al*. [2006] to June 2005 using NCEP-NCAR reanalysis data [*Kalnay et al*., 1996]. The forcing data set from *Qian et al*. [2006] was derived by combining monthly data of temperature and precipitation derived from station records with intra-monthly variations from the NCEP-NCAR reanalysis [*Kalnay et al*., 1996]. Historical records of cloud cover were also used to derive surface incident solar radiation. Second, we conducted a 220 year run by CLM3 forced with recycled 1948–2004 forcing data from *Qian et al*. [2006] in order to spin up the deep soil layers using the method proposed by *Tian et al*. [2010b]. Third, the CLM3 simulation was forced with this extended data set from May 1986 to June 2005 (20 years) to obtain the initial fields for both the assimilation and simulation experiments. From this state at the end of 20 year run, we ran the original CLM3 and the LDAS-MO forced with the atmospheric forcing data created by *Shi* [2008] from July 2005 to December 2008. It should be noted that all of the runs were at the same resolution of 0.25° latitude 0.25° longitude, and the forcing data were also interpolated onto the 0.25° grid.

Table 1. Locations of the Four Sub-Regions in China

Identification	Region Name	Location	No.[a]	Characteristics
NEC	north-eastern China	120°E–133°E, 40°N–50°N	38	Semi-humid and middle temperate continental monsoon climate
NC	northern China	105°E–120°E, 34°N–42°N	88	Semi-arid and warm temperate monsoon climate
NWC	north-western China	76°E–105°E, 34°N–46°N	36	Arid and semi-arid and warm temperate monsoon climate
YR	middle and lower reaches of the Yangtze River valley	105°E–122°E, 30°N–34°N	31	Humid and subtropical monsoon climate

[a]No. is the number of observational stations.

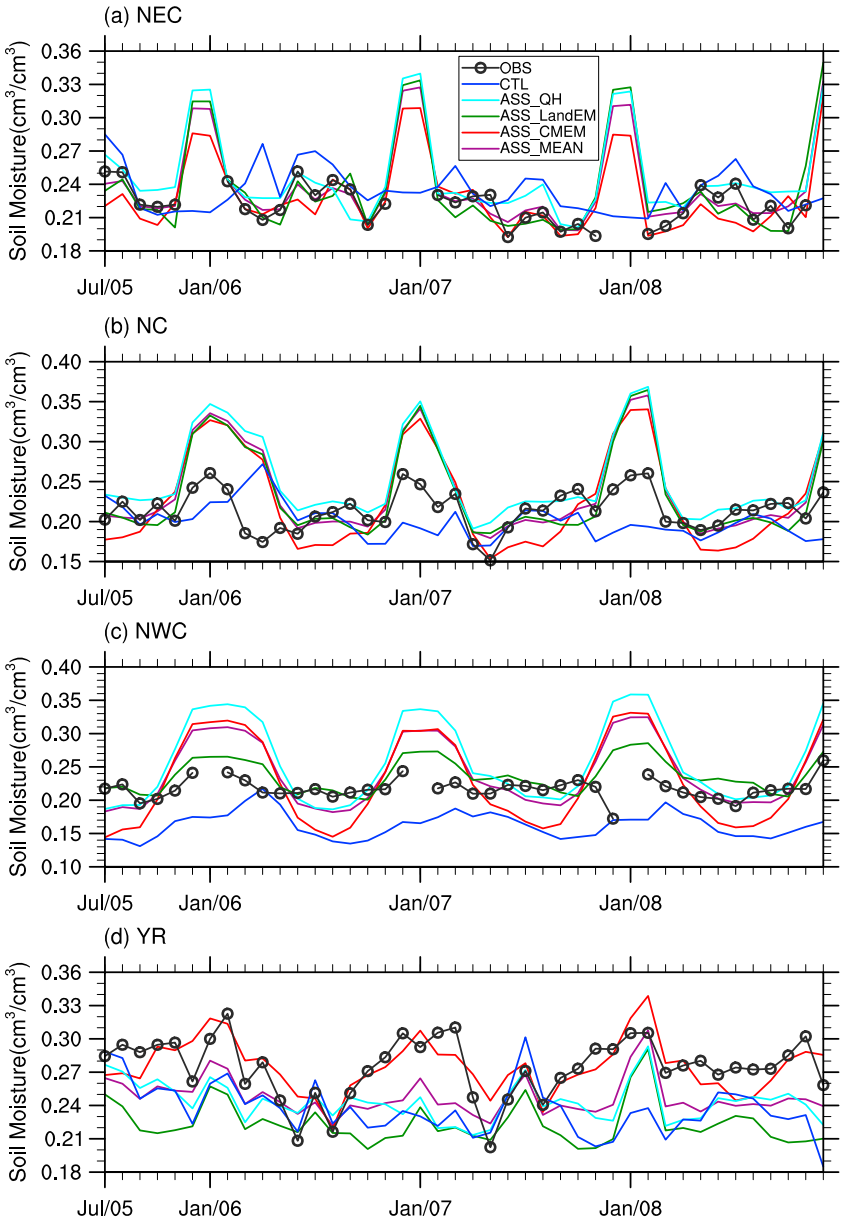

Figure 3. (a–d)Time series of monthly mean volumetric soil moisture content from ground observations (OBS), control run (CTL), and three assimilated cases (ASS_QH, ASS_LandEM, and ASS_CMEM) and their simple arithmetic average (ASS_MEAN) over four sub-regions in China (defined in Table 1).

[21] The LDAS-MO was evaluated using in situ soil moisture data from the CMA. A common issue is how to compare in situ point measurements with assimilated soil moisture averaged over a grid box of satellite observations. Soil moisture measured at a single point is often representative merely on a limited spatial scale, depending on heterogeneity of soil properties, land cover, and atmospheric conditions. To reduce the scale mismatch, we used the arithmetic mean of the observed soil moisture contents from the all available stations as a proxy for the mean state of a reference region, which was compared to the area-weighted average values of the CLM3 simulation over the grids where the observations were available. Figure 2 shows the locations of the 226 field sites, which are grouped into four subdivisions roughly according to the spatial patterns of the dryness and wetness centers in China from *Zhu* [2003] and contain 203 of the field sites. More information about the four subdivisions can be found in Table 1 and Figure 2. Soil moisture observations were collected on the eighth, eighteenth, and 28th days of each month from July 2005 to December 2008. Only observed data for the upper soil layer (0–10 cm) were used in this study because the records for deeper layers were insufficiently complete to be suitable for the evaluation. Although the selected regions in this study were large, we did not consider the highly variable conditions in the following analysis due to limited soil moisture field measurements. Also note that we have not prove that the arithmetic mean is optimal up to now, which should be addressed in the future.

[22] In this study, the ensemble size N for state variables in the En4DVAR assimilation algorithm [*Tian et al.*, 2008]

Table 2. The Mean Bias Errors (MBEs, cm^3/cm^3), Root Mean Square Errors (RMSEs, cm^3/cm^3) and Correlation Coefficients (r) for Monthly Volumetric Soil Water Content From Control Run (CTL) and Three Assimilated Cases (ASS_QH, ASS_LandEM, and ASS_CMEM) Compared to Ground Observations Over Four Sub-Regions From July 2005 to December 2008

	CTL			ASS_QH			ASS_LandEM			ASS_CMEM			ASS_MEAN[a]		
Region	MBE	RMSE	r	MBE	RMSE	r	MBE	RMSE	r	MBE	RMSE	r	MBE	RMSE	r
NEC	0.016	0.023	0.58	0.010	0.017	0.65	−0.002	0.014	0.62	−0.005	0.014	0.64	0.001	0.010	0.81
NC	−0.014	0.035	0.44	0.038	0.054	0.66	0.020	0.048	0.63	0.013	0.049	0.64	0.024	0.049	0.65
NWC	−0.055	0.059	**0.21**[b]	0.035	0.062	0.44	0.020	0.028	0.39	0.006	0.057	0.43	0.019	0.046	0.44
YR	−0.037	0.048	**0.23**[b]	−0.031	0.040	0.37	−0.050	0.057	0.32	0.003	0.019	0.72	−0.027	0.034	0.60

[a]ASS_MEAN is the simple arithmetic average value for the three assimilated cases.
[b]The bold values indicate that the correlation coefficients are not statistically significant ($p < 0.05$).

is chosen to be 60 based on the results of *Tian et al*. [2009]. The background error convariance was estimated by the ensemble samples, which were constructed by adding N perturbed fields to the initial fields using the Monte Carlo method through some sensitivity experiments [*Tian et al*., 2008]. *de Rosnay et al*. [2009] pointed out that compared to AMSR-E observed C-band T_b data, the mean bias for CMEM is in the range of −14 to 11 K over West Africa (depending on the choice of land surface model). We conducted similar experiments to estimate T_b over China for three RTMs (LandEM, CMEM, and QH) coupled to CLM3 without assimilation (shown in the discussion below). It is found that the biases for three RTMs are in the range of −8 to 5 K. Therefore, the absolute biases for three observation operators were all assumed to be 8 K in this study. However, the fixed values for observation biases of three RTMs may have some effects on the assimilated results of the LDAS-MO, which needs more sensitivity experiments and will be addressed in the future. In addition, we did a quality control for the AMSR-E observed T_b data to eliminate the abnormal data, and then they were directly assimilated into the LDAS-MO in the assimilation pass or used to optimize the parameters in the optimization pass.

3.2. Impact of Observation Operators

[23] Figure 3 shows the time series of monthly mean volumetric soil-water content from the assimilation and control run compared to ground observations (OBS) for the four field subdivisions between July 2005 and December 2008. OBS is the arithmetic mean from all available field stations in each subregion; model-simulated outputs were derived using the area-weighted averaging over the grids where the observations were available in the corresponding region. ASS_MEAN stands for the simple arithmetic average for the three assimilated cases ASS_QH, ASS_LandEM, and ASS_CMEM, from which the soil-water content is the assimilated results of the LDAS-MO. It is readily seen in Figure 3 that the volumetric soil-water content analysis by the LDAS-MO using the three observation operators was closer to the OBS than was seen using CTL and captured the time variations of observed soil moisture more closely. Very highly unrealistic simulated values of soil moisture compared to OBS were produced for the boreal winter for all the assimilated cases over most parts of China, however, except for the middle and lower reaches of the Yangtze River (YR). The two main explanations for this result are, first, that none of the three RTMs took frozen soil into consideration; and second, poorer simulation of snow emissivity due to uncertainties in the snow parameters (e.g., snow particle size). However, these observation operators performed differently over different areas.

[24] Table 2 summarizes the information contained in Figure 3. It is found that the CTL from CLM3 underestimates in situ soil moisture measurements over most regions of China except for the north-eastern China (NEC) with a positive bias ($0.016\ cm^3/cm^3$). The result is consistent with the finding of *Qian et al*. [2006] that CLM3-simulated soil moisture forced by their own forcing data set showed a dry bias over eastern and southern China. In addition, the volumetric soil moisture content from CTL shows very low temporal correlations with OBS ($r < 0.25$, Table 2) over north-western China (NWC) and YR. This may be related to the quality of the atmospheric forcing data set from *Shi* [2008] or the performance of CLM3 over these two regions. In contrast, the soil moisture derived from three assimilated cases of the LDAS-MO was improved to agree more closely with ground observations by assimilating the microwave T_b data. Compared to the other two cases, the assimilated soil moisture from ASS_QH was the most closely correlated with OBS for arid and semi-arid regions such as northern China (NC, $r = 0.66$), and NWC ($r = 0.44$), mainly because of the QH model's greater sensitivity to dry soil [*Yang et al*., 2007]. The QH model used in this study not only described surface scattering but also took into account the volumetric scattering effect of dry soil, which makes a substantial contribution to the satellite-received radiance [*Lu et al*., 2006]. ASS_QH also showed a better performance over NEC ($r = 0.65$). This may be due to the fact that the main vegetation functional types over NEC are grass and crop for which the QH model uses a complicated parameterization of vegetation optical thicknesses, choosing different parameters depending on vegetation type (wheat, soybean, or grasses) [*Jackson and Schmugge*, 1991; *Fujii*, 2005]. In addition, ASS_LandEM had the smallest root mean square errors (RMSEs) over these three regions. As *de Rosnay et al*. [2009] had pointed out, the T_b simulation by the Kirdyashev model [*Kirdyashev et al*., 1979] in CMEM agreed most closely with satellite observations over densely vegetated areas. For the same vegetation opacity model, we found that ASS_CMEM agreed best with OBS over YR (higher vegetation fraction), giving the smallest mean bias error (MBE, $0.003\ cm^3/cm^3$) and RMSE ($0.019\ cm^3/cm^3$) and the closest correlations with OBS in this study ($r = 0.72$). Table 2 also shows that ASS_MEAN agrees more closely with OBS than the three assimilated cases over parts of the target areas, especially

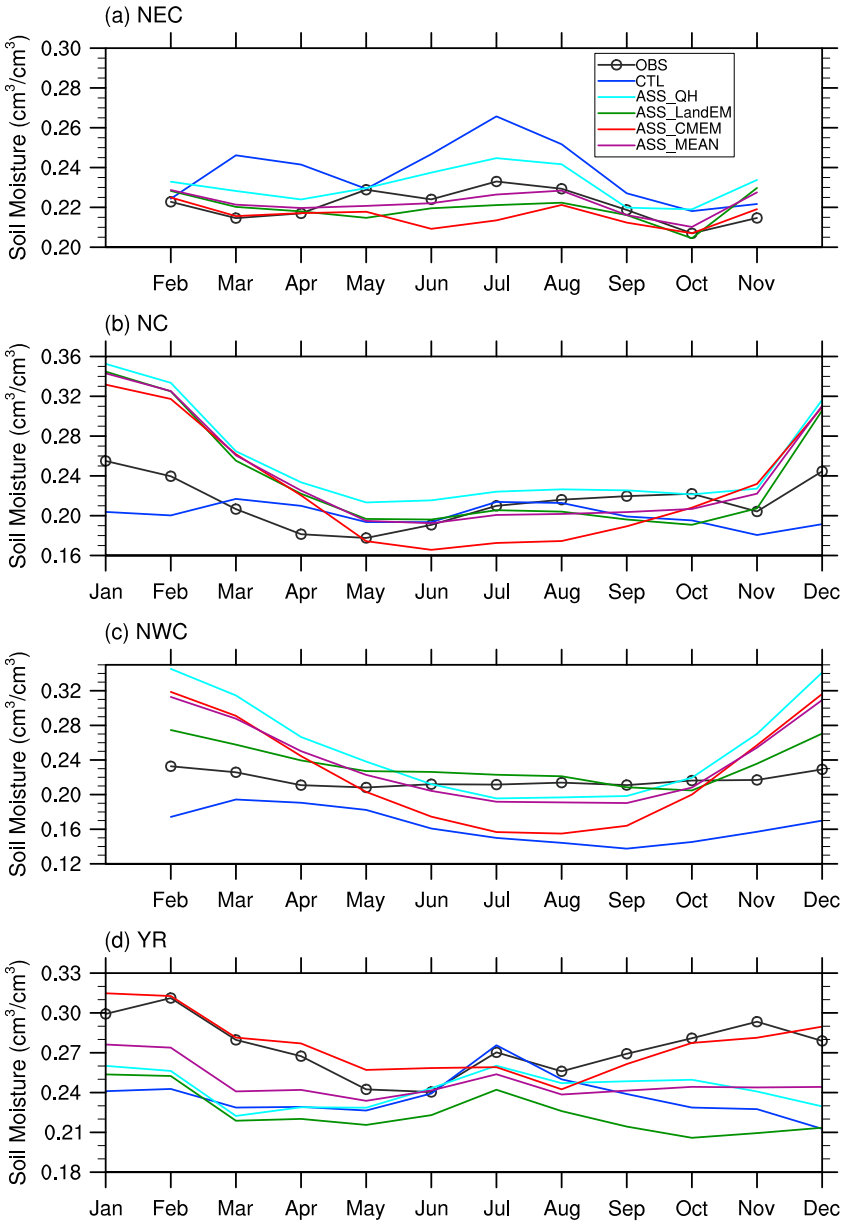

Figure 4. (a–d) Mean monthly volumetric soil moisture (averaged from July 2005 to December 2008) for the four sub-regions in China.

NEC, giving the lowest RMSE ($0.01\,cm^3/cm^3$) and the highest correlation coefficient ($r = 0.81$), yet over WC it produced a greater RMSE than ASS_LandEM relative to OBS and performed worse than ASS_CMEM over YR. All of these results demonstrate that a simple arithmetic ensemble average is not always acceptable. In addition, given the relatively short experiment period of about three and half years and the evaluation at monthly time scales, we conducted an analysis on whether assimilated soil moisture estimations are statistically different from those without assimilation. The statistical significance tests were performed for their mean values and variances using t test ($p < 0.05$) and F test ($p < 0.05$), respectively. It is found that the assimilated soil moisture contents from three assimilated cases in the mean values and variances are statistically different from those from pure CLM3-simulation except over YR for both ASS_QH and ASS_LandEM. This suggests that the improvements through the assimilation are statistically significant over most of the regions.

[25] We also evaluated the mean annual cycle of volumetric soil moisture for the top 10 cm layer from model-based output with and without assimilation against ground observations during July 2005–December 2008 over the four sub-regions (Figure 4). Compared with the CTL case, all the assimilation cases capture observed seasonal variations in soil moisture better for most of the regions. Over the arid region (NWC, Figure 4c), the CLM3 clearly underestimates the observed soil moisture while ASS_LandEM agrees better with OBS. Consistent with the previous results mentioned above (Table 2), ASS_CMEM captures seasonal variations of observed soil moisture best among all cases over YR (Figure 4d).

Table 3. The Bayesian Model Averaging (BMA) Weights for Three Assimilated Cases (ASS_QH, ASS_LandEM, and ASS_CMEM) Over Four Sub-Regions in China

Region	ASS_QH	ASS_LandEM	ASS_CMEM
NEC	0.5822	0.0020	0.4158
NC	0.6759	0.0153	0.3088
NWC	0.8832	0.0048	0.1120
YR	0.0005	0.0013	0.9982

4. Assimilation Skill Enhancement by the BMA Method

[26] Since, as above, the simple arithmetic average method did not explicitly show the best performances over all four sub-regions, we developed a Bayesian model averaging (BMA) scheme with LDAS-MO to enhance its assimilation skill. The BMA algorithm proposed by *Raftery et al.* [2005] is a statistical post-processing method that produces probabilistic forecasts from ensembles, expressed as

$$p(\Delta|f_1,\cdots,f_N)=\sum_{k=1}^{N} w_k g_k(\Delta|f_k), \tag{4}$$

where $f=f_1,\cdots,f_N$ is an ensemble of predictions obtained from N different models (three assimilated cases in this study); w_k is the posterior probability that forecast k is the best; $g_k(\Delta|f_k)$ is the conditional probability density function (PDF) of the forecast variable Δ (soil moisture, in the present case) on f_k. The weight w_k is non-negative, satisfying $\sum_{k=1}^{N} w_k = 1$, and represents each model's contribution to the predictive skill of the ensemble. In this study, the probability distribution of soil moisture was approximated by a gamma distribution [*Tian et al.*, 2011]. This is appropriate when the soil moisture distribution is primarily driven by precipitation. Following *Vrugt et al.* [2008], the BMA weight w_k of the individual competing models was estimated from a calibration data set (i.e., in situ soil moisture observations in this study) using the maximum likelihood technique for which the values were obtained iteratively using the Differential Evolution Adaptive Metropolis Markov Chain Monte Carlo algorithm.

[27] Here, we chose the time period from July 2005 to December 2006 as the training period, and January 2007 to December 2008 as the evaluation period. The BMA weights of three assimilated cases over the four sub-regions (Figure 2) are presented in Table 3. It was found that the LandEM model had the smallest BMA weights, demonstrating its smaller contribution. This is largely due to its worse performance (lower correlations with OBS as shown in Table 2) compared to the other two observation operators. Figure 5 shows the MBEs, RMSEs, and correlation coefficients of monthly soil moisture from the three assimilated cases, their arithmetic average and BMA-based estimation against OBS between January 2007 and December 2008. It is seen that the BMA-simulated estimation agreed best with OBS for all cases, showing the smallest MBEs and RMSEs and highest correlations with OBS, although it produced wetter biases than ASS_CMEM over parts of northern China. The main reason for this result is that the BMA method evaluated the performance of each individual competing member of the ensemble appropriately during the training period, enabling it to better assign the BMA weights that represented each model's relative forecasting skill. Consequently, there was a significant improvement in the overall forecasting precision of the ensemble. For example, compared to ASS_MEAN, a lower bias ($0.013\ cm^3/cm^3$) and RMSE ($0.018\ cm^3/cm^3$) were found for ASS_BMA over NC. Furthermore, ASS_BMA also showed the best performance over YR, while the largest weight was assigned to ASS_CMEM (Table 3), which had shown better agreement with observed data than had the other two assimilated cases. However, as all ensemble members overestimated soil moisture in winter, the worst performance was also found for ASS_BMA

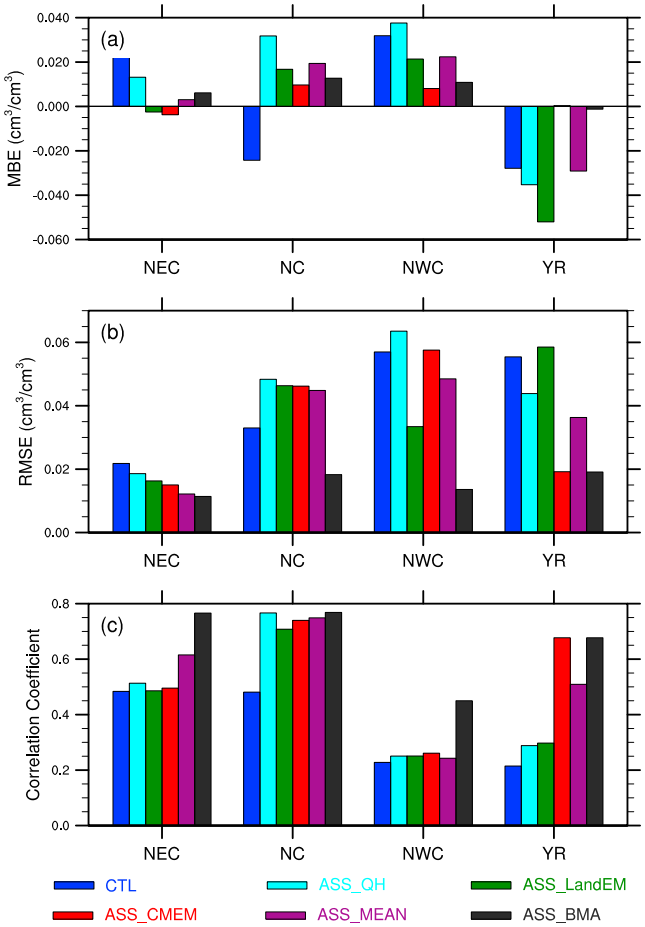

Figure 5. (a) The mean bias errors (MBE), (b) root mean square errors (RMSE), and (c) correlation coefficients between the observed and simulated monthly volumetric soil-water content over four sub-regions in China between January 2007 and December 2008. CTL is the control run without assimilation; ASS_QH, ASS_LandEM, and ASS_CMEM are three assimilated cases with different observation operators; ASS_MEAN is the simple arithmetic average from three assimilated cases; ASS_BMA is the ensemble average of three assimilated cases using the Bayesian model averaging (BMA) method.

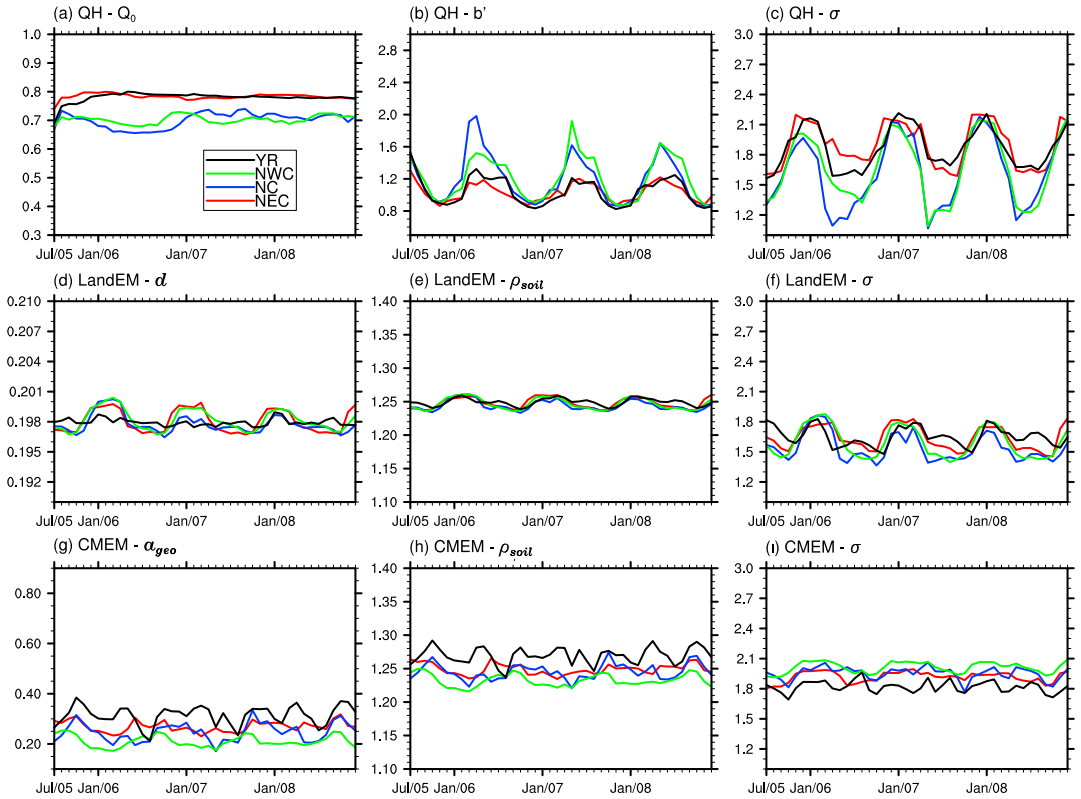

Figure 6. Time series of the monthly mean optimized parameters for three radiative transfer models: (a–c) QH, (d–f) LandEM, and (g–i) CMEM over four sub-regions in China.

(not shown). This demonstrates that the parameterizations of snow and frozen soil in the RTM should be improved to enable a more robust estimate of soil moisture from the assimilation system during the cold seasons.

5. Discussions

5.1. Role of Parameter Estimations

[28] As mentioned in section 2.2, the choice of key parameters to be optimized is critical for the accuracy of the RTM's outputs and thus the performance of the LDAS-MO; however, they may be either highly variable or unavailable, making it difficult or impossible to assign a realistic value to every parameter. For QH, the initial values and ranges (in the brackets) of the three parameter were set $Q_0 = 0.6$ $(0.3 - 1.0)$, $b' = 2.1$ $(0.4 - 3.0)$, $\sigma = 1.6$ $(1.0 - 3.0)$ [*Yang et al.*, 2007]; for LandEM, $d = 0.2$ $(0.19 - 0.21)$, $\rho_{\text{soil}} = 1.18$ $(1.1 - 1.4)$, $\sigma = 1.6$ $(1.0 - 3.0)$ [*Weng et al.*, 2001]; and $a_{\text{geo}} = 0.4$ $(0.1 - 0.9)$, $\rho_{\text{soil}} = 1.18$ $(1.1 - 1.4)$, $\sigma = 1.6$ $(1.0 - 3.0)$ for CMEM [*Holmes et al.*, 2008]. Here, we only show the optimized results of these parameters for three RTMs over the four sub-regions in China (Figure 6). It is found that most of these parameters show clear regional differences except for the soil bulk volume density (ρ_{soil}) in LandEM (Figure 6e). In contrast, there is a larger variation in ρ_{soil} for CMEM which may be related to its subgrid-scale variability (up to seven tiles in a grid cell) [*de Rosnay et al.*, 2009]. In addition, the parameter d (Figure 6d) shows low variability, which may be due to that the leaf thickness d changes little generally. The surface roughness height (σ) in QH shows a larger amplitude than other two RTMs, which suggests that this parameter may be more sensitive to the change of satellite-observed microwave T_b data. However, it is problematic as to which of them is more reasonable due to large discrepancies in the model parameterizations for three RTMs.

[29] To investigate the role of these key parameters in the assessment of the RTMs, we calculated the brightness temperatures using fixed values and optimized parameters (Figure 6), respectively. The fixed values for all parameters were set to be their initial values as mentioned above. Figure 7 shows the comparisons between AMSR-E observed brightness temperatures and those predicted by the three RTMs with and without parameter estimation. It is found that the T_b data from the RTMs with parameter estimation (solid lines) are more close to AMSR-E observations than those using fixed parameters (dashed lines). However, all the RTM-based simulations using fixed or optimized parameter values underestimate T_b about $5 \sim 20$ K in the cold seasons, which may relate to inaccurate parameterizations for snow and frozen soil in the RTMs. It would then affect the soil moisture assimilation results (Figure 3). The MBEs, RMSEs, and correlation coefficients between the T_b data simulated by RTMs and AMSR-E observations are presented in Figure 8. Due to the parameter optimization, the three RTMs showed lower biases (Figure 8a) than those using fixed values. Moreover, the CMEM with parameter estimation agrees best with AMSR-E T_b data over YR among three RTMs (Figure 8c), which may lead to its better performance for soil moisture assimilation than QH and LandEM in the LDAS-MO (Table 2).

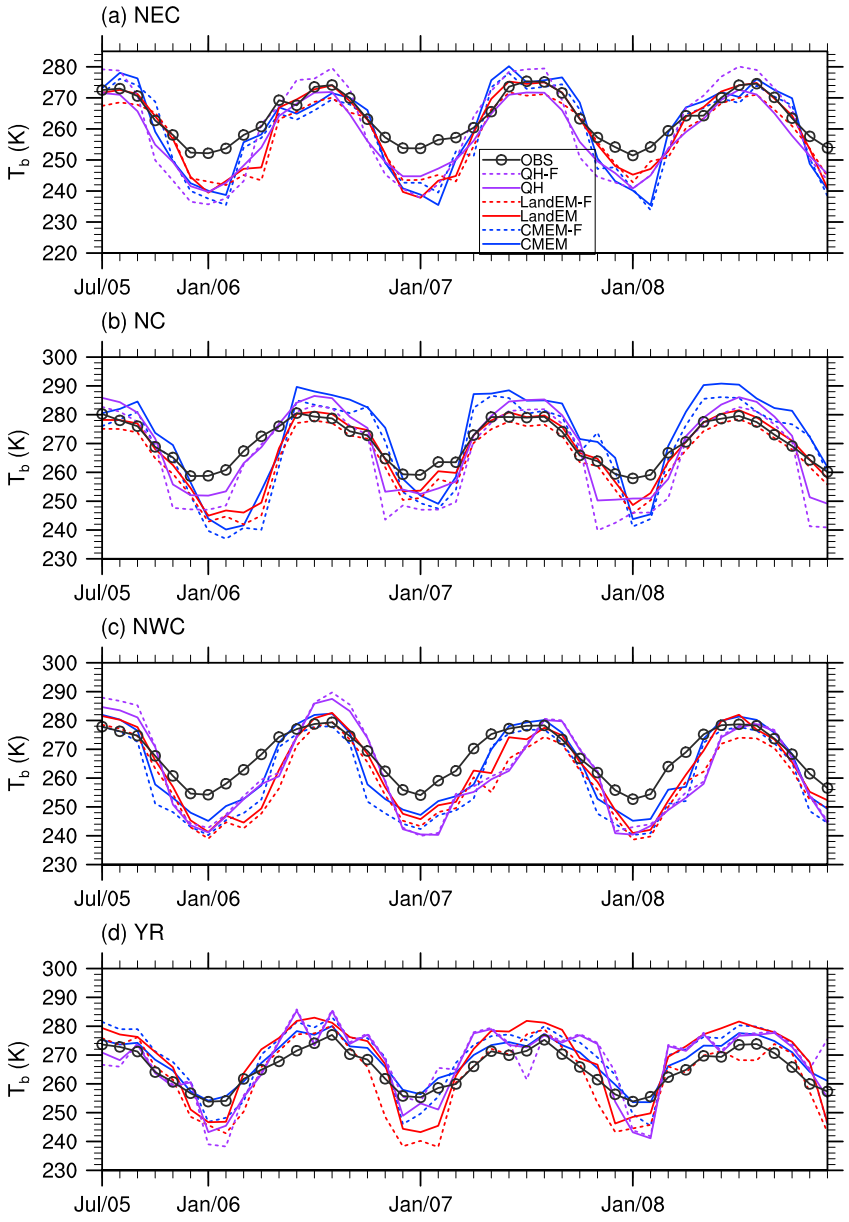

Figure 7. (a–d) Time series of monthly mean brightness temperatures from AMSR-E (OBS) and three RTMs with (QH, LandEM, and CMEM) and without parameter estimation (QH-F, LandEM-F, and CMEM-F) over four sub-regions in China.

5.2. Effects of the Atmospheric Forcing Data Sets

[30] The atmospheric forcing data set of *Shi* [2008] was chosen to run LDAS-MO due to its high spatial-temporal resolutions. However, it is seen from Figure 3 and Table 2 that the correlation coefficients between OBS and CLM3-based soil moisture estimations without assimilation over two sub-regions (NWC and YR) are low ($r < 0.25$). To investigate the role of the atmospheric forcing data set in the CLM3-simulated soil moisture, we chose two other data sets to force the CLM3 to be compared with that from *Shi* [2008] (hereafter FY-2C). One is the ERA-Interim reanalysis [*Dee et al*., 2011], which is 6-hourly and has 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 [*Dee et al*., 2011]. The other one is a global meteorological forcing data set from Princeton University (hereafter Princeton), which is a hybrid of data from the NCEP-NCAR reanalysis [*Kalnay et al*., 1996] and a suite of global observation-based products [*Sheffield et al*., 2006]. The horizontal resolution of the Princeton data set is 1° × 1°, with a temporal resolution of 3 h [*Sheffield et al*., 2006]. To facilitate quantitative comparison with CLM3-simulation results forced by the FY-2C atmospheric data set [*Shi*, 2008], the other two data sets (ERA-Interim and Princeton) were used to force CLM3 with the same spatial-temporal resolutions (0.25° increments and half-hour time steps).

[31] Figure 9 shows the comparison between ground-based monthly volumetric soil moisture and that simulated by CLM3 using three different meteorological forcing data sets

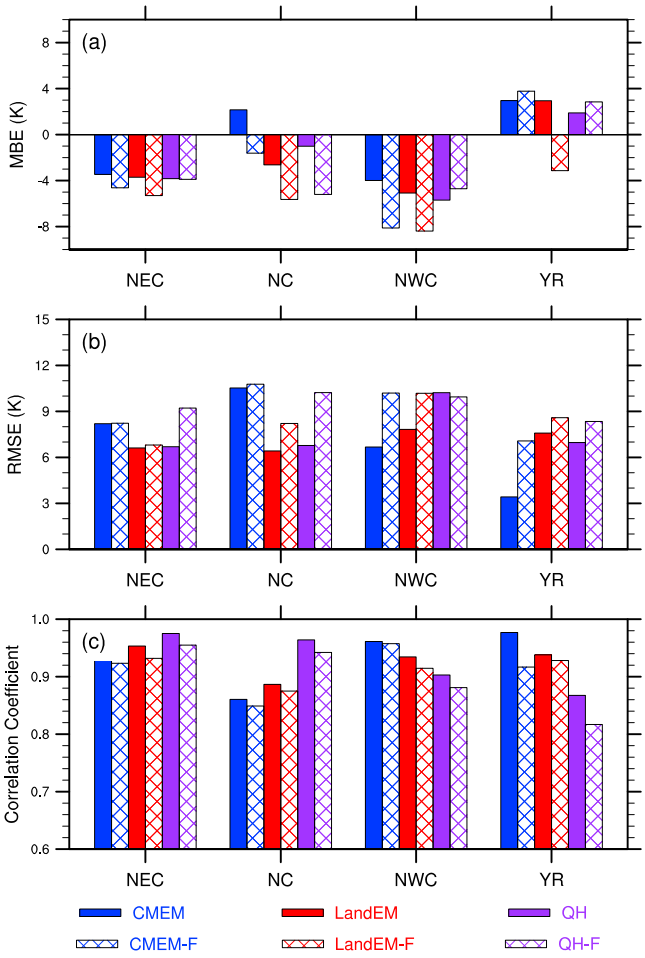

Figure 8. (a) The MBE, (b) RMSE, and (c) correlation coefficients between the AMSR-E observed brightness temperatures and those predicted by the RTMs with (QH, LandEM, and CMEM) and without parameter estimation (QH-F, LandEM-F, and CMEM-F) over four sub-regions in China between July 2005 and December 2008 (shown in Figure 7).

over two sub-regions (NWC and YR) in China. It is clearly observed that the volumetric soil-water content from FY-2C is more close to OBS than that from both the ERA-Interim and Princeton over the two regions, while the later two show larger dry biases (MBE $< -0.087\ \mathrm{cm}^3/\mathrm{cm}^3$ over NWC and MBE $< -0.056\ \mathrm{cm}^3/\mathrm{cm}^3$ over YR). This may be due to that *Shi* [2008] integrated more precipitation gauge data than other two data sets. Over NWC, however, both the soil moisture from ERA-Interim ($r = 0.52$) and Princeton ($r = 0.41$) correlated better with OBS than that from FY-2C ($r = 0.21$), which is not statistically significant ($p < 0.05$). It may be related to sparse precipitation gauge stations over this region and the quality of the precipitation from *Shi* [2008] mainly depends on the FY-2C satellite retrievals, which could not capture the observed precipitation variations well there [*Shi*, 2008]. In addition, the CLM3-simulation results using the three atmospheric forcing data sets all show a low correlation with OBS over YR, which may be related to the inappropriate parameterizations or parameters of the land surface model CLM3 over this region. Moreover, Figure 9 showed clear dry biases in CLM3 over NWC and YR for all three forcing data sets, which were consistent with the findings by several previous studies [*Qian et al.*, 2006; *Oleson et al.*, 2008]. These cases were improved by the LDAS-MO in this study (Table 2).

5.3. Quality Control for the Snow and Frozen Soil Conditions

[32] It is observed from Figure 3 that the LDAS-MO developed in this study produced highly questionable soil moisture analysis over three sub-regions (NEC, NC, and NWC) in the wintertime. Here, we took a simple quality control (QC) step for the LDAS-MO to eliminate the effects of the snow-covered or frozen soil conditions on the assimilated soil moisture analysis. If the snow exists (snow depth > 0) or the surface soil temperature is lower than 273 K, the AMSR-E observed T_b data will not be assimilated into the LDAS-MO. Figure 10 shows the comparison between the assimilated volumetric soil moisture content from LDAS-MO with QC and ground observations over three sub-regions (NEC, NC, and NWC). It is found that

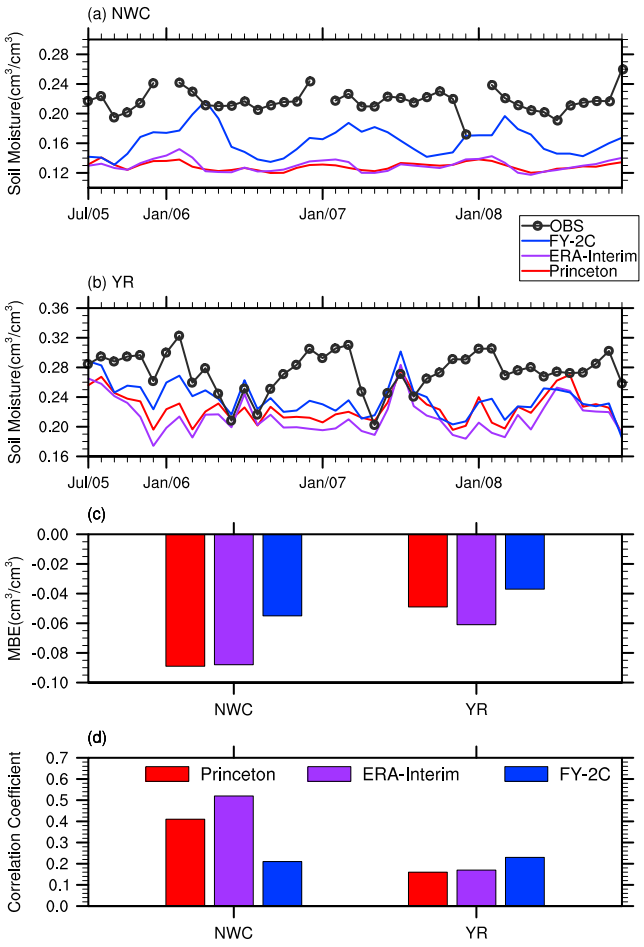

Figure 9. Time series of monthly mean volumetric soil moisture content from ground observations (OBS) and CLM3-simulations using three different atmospheric forcing data sets (Princeton, ERA-Interim, and FY-2C) over two sub-regions (a) NWC and (b) YR in China (defined in Table 1); and the (c) MBE and (d) correlation coefficients between OBS and three simulation results.

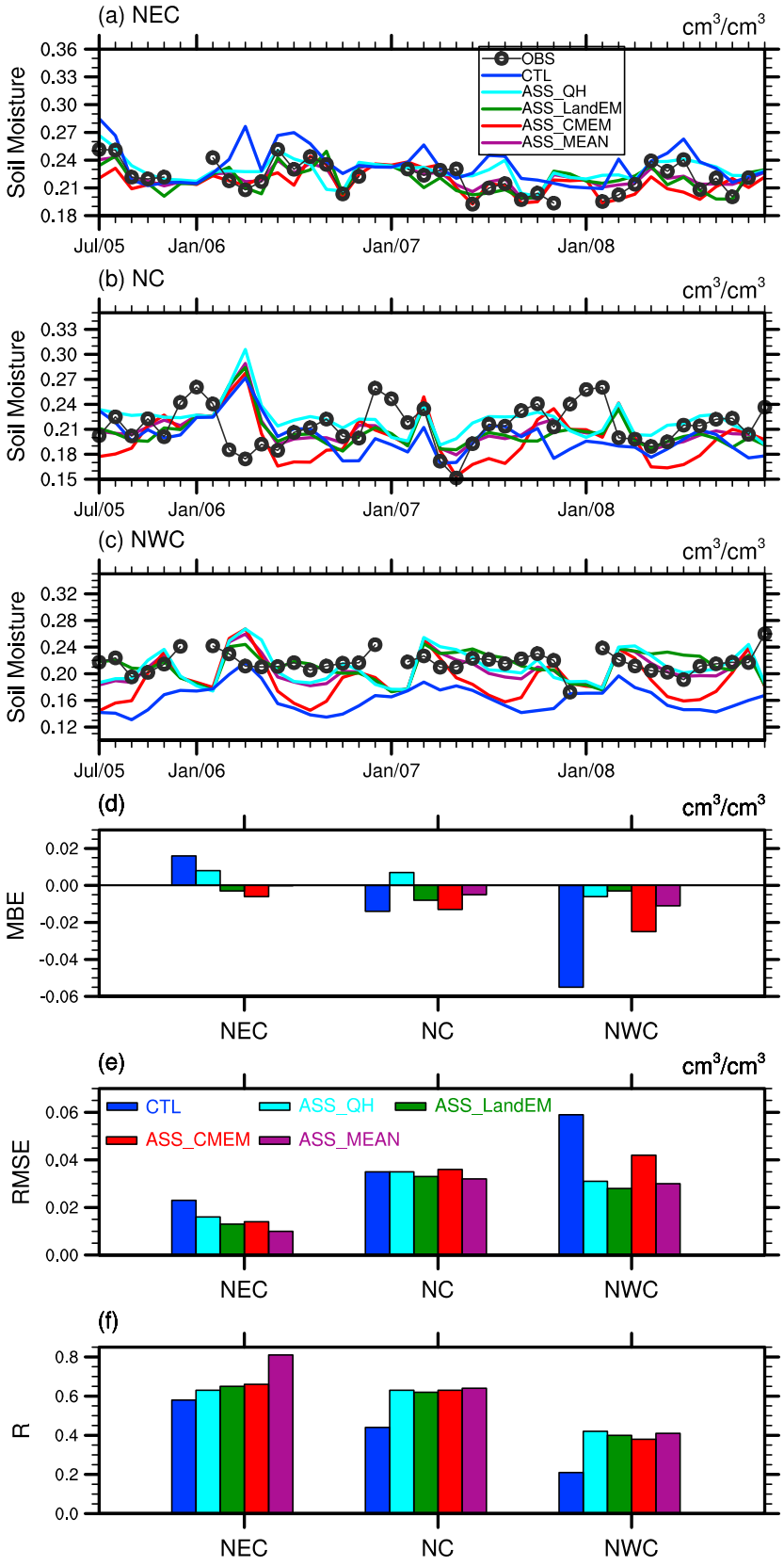

Figure 10. Time series of monthly mean volumetric soil moisture content from ground observations (OBS), control run (CTL), and three assimilated cases (ASS_QH, ASS_LandEM, and ASS_CMEM) and their simple arithmetic average (ASS_MEAN) over three sub-regions (a) NEC, (b) NC, (c) NWC in China; and the (d) MBE, (e) RMSE, (f) correlation coefficients (R) between OBS and simulated soil moisture estimations with or without assimilation.

the assimilated soil moisture estimations with QC are more close to OBS than those without QC (shown in Figure 3). As there is barely assimilation in the cold seasons, the assimilated results with QC are close to pure CLM3-simulations. The MBE, RMSE, correlation coefficients (R) between OBS and simulated soil moisture with or without assimilation are also provided in Figures 10d–10f). The QC method reduced the biases of LDAS-MO, with lower MBEs and RMSEs (Figures 10e–10f) compared to those in Table 2. However, the correlation coefficients with OBS were not improved clearly. In contrary, for example, the correlations (Figure 10f) were lower than those without QC over NC (Table 2) as the CLM3 produced abnormal values compared to OBS in winter over this region (Figure 10b).

6. Conclusions

[33] In this study, a dual-pass microwave land data assimilation system with multi-observation operators (LDAS-MO) was developed, and the impact of the observation operators QH, LandEM, and CMEM on the assimilated soil moisture was then investigated using the assimilation system. Moreover, we also constructed a BMA scheme for the LDAS-MO to improve its assimilation accuracy.

[34] Numerical experiments from the assimilation of AMSR-E microwave brightness temperatures demonstrated that the observation operators each had significant effects on the assimilated volumetric soil-water content. Large discrepancies were observed over different areas between the soil moisture analysis for each of the three observation operators individually, although they all improved the soil moisture estimation in that they all agreed more closely with field observations (OBS) than model simulations without assimilation for the absence of frozen or snow-covered conditions, and were better able to capture the time variations of observed soil moisture. The soil moisture analysis for QH showed higher correlations with OBS over arid and semi-arid regions due to its higher sensitivity to dry soil, while smaller RMSEs were observed for LandEM. CMEM, in which the vegetation opacity model simulated the T_b data in best agreement with satellite observations over densely vegetated areas, agreed most closely with OBS over YR (bias = 0.003 cm^3/cm^3, RMSE = 0.019 cm^3/cm^3, $r = 0.72$). Moreover, CMEM and LandEM captured observed soil moisture seasonality better than CLM3 simulation over YR and NWC, respectively. Results also showed that when the BMA scheme was used, the soil moisture analysis was significantly improved, showing reduced RMSEs and increased correlation coefficients relative to OBS over all regions. The BMA method evaluates the performance of each competing ensemble member during the training period, then assigns the BMA weight that represents the relative forecasting skill of each model. Consequently, there is a significant improvement in the ensemble forecasting precision. This study will be helpful in both better understanding the sensitivity of the observation operators on assimilated results and generating soil moisture data sets with higher accuracy.

[35] However, the least satisfactory performance for all assimilated cases in the coldest seasons was not improved by the BMA method. One approach to partially address this issue is through assimilating satellite-based multi-channel brightness temperatures [*Jin et al*., 2009; *Rees et al*., 2010]. This will provide more information about the snow and frozen soil and then improve near-surface soil moisture estimation in the cold seasons. In addition, parameterizations for snow (e.g., snow grain size, density, and fractional volume) and frozen soil (e.g., freeze-thaw state, volume scatter darkening effect of frozen soil) in the observation operator (RTM) to make the assimilated soil moisture more robust should be discussed in the future. The third strategy is to apply a bias correction method or incorporate multi-model observation operators with BMA algorithm to provide more accurate brightness temperature forecasts and thus improve the assimilated results. It should also be noted that only the area-average biases in soil moisture, which may be lower than the local biases, were provided in this study due to limited soil moisture measurements. The Heihe Watershed Allied Telemetry Experimental Research experiment [*Li et al*., 2013] has been performed since May 2012 over the middle stream of Heihe Basin in the northwest of China, and it can afford soil moisture measurements at 50 field sites covering an area of 5.5×5.5 km^2. These data will be used to evaluate the LDAS-MO at local scale in our future study, which is relevant in the context of the land data assimilation system.

[36] **Acknowledgments.** This research was supported by the National Basic Research Program of China (Grant Nos. 2010CB951101 and 2010CB428403), the National Natural Science Foundation of China (Grant Nos. 91125016 and 41075062), and the Chinese Academy of Sciences Strategic Priority Research Program (Grant No. XDA05110102). The ground observation data set was obtained from the China Meteorological Data Sharing Service System website <http://cdc.cma.gov.cn>. The authors thank NSIDC for providing AMSR-E brightness temperature data. We also thank Dr. Steve Ghan and the three anonymous reviewers for constructive comments and suggestions, which have helped us in improving the paper.

References

Bonan, G. B., S. Levis, L. Kergoat, and K. W. Oleson (2002), Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models, *Global Biogeochem. Cycles*, *16*, 1021, doi:10.1029/2000GB001360.

Darcy, H. (1856), Les Fontaines Publiques de la Ville de Dijon, Dalmont, Paris.

de Rosnay, P., M. Drusch, A. Boone, G. Balsamo, B. Decharme, P. Harris, Y. Kerr, T. Pellarin, J. Polcher, and J.-P. Wigneron (2009), AMMA land surface model intercomparison experiment coupled to the community microwave emission model: ALMIP-MEM, *J. Geophys. Res.*, *114*, D05108, doi:10.1029/2008JD010724.

Dee, D. P., et al. (2011), The ERA-interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, *137*, 553–597.

Dobson, M., F. Ulaby, M. Hallikainen, and M. El-Rayes (1985), Microwave dielectric behavior of wet soil-part II: Dielectric mixing models, *IEEE Trans. Geosc. Sci. Remote Sens.*, *38*, 1635–1643.

Draper, C. S., R. H. Reichle, G. J. M. De Lannoy, and Q. Liu (2012), Assimilation of passive and active microwave soil moisture retrievals, *Geophys. Res. Lett.*, *39*, L04401, doi:10.1029/2011GL050655.

Drusch, M., T. Holmes, P. de Rosnay, and G. Balsamo (2009), Comparing ERA-40 based L-band brightness temperatures with Skylab observations: A calibration/validation study using the Community Microwave Emission Model, *J. Hydrometeor.*, *10*, 213–226.

Duan, Q. Y., N. K. Ajami, X. G. Gao, and S. Sorooshian (2007), Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Adv. Water Res.*, *30*, 1371–1386.

Fujii, H. (2005), Development of a microwave radiative transfer model for vegetated land surface based on comprehensive in-situ observations, Ph.D. dissertation, Univ. of Tokyo, Tokyo.

Holmes, T., M. Drusch, J.-P. Wigneron, and R. de Jeu (2008), A global simulation of microwave emission: Error structures based on output from

ECMWF's operational Integrated Forecast System, *IEEE Trans. Geosci. Remote Sens.*, *46*, 846–856.

Huang, C. L., X. Li, L. Lu, and J. Gu (2008), Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter, *Remote Sens. Environ.*, *112*, 888–900.

Jackson, T., and T. Schmugge (1991), Vegetation effects on the microwave emission of soils, *Remote Sens. Environ.*, *36*, 203–212.

Jia, B. H., and Z. H. Xie (2011), Evaluation of the community microwave emission model coupled with the community land model over East Asia, *Atmos. Oceanic Sci. Lett.*, *4*, 209–215.

Jia, B. H., Z. H. Xie, X. J. Tian, and C. X. Shi (2009), A soil moisture assimilation scheme based on the ensemble Kalman filter using microwave brightness temperature, *Sci. China Ser. D – Earth Sci.*, *52*(11), 1835–1848, doi:10.1007/s11430-009-0122-z.

Jin, R., X. Li, and T. Che (2009), A decision tree algorithm for surface soil freeze/thaw classification over China using SSM/I brightness temperature, *Remote Sens. Envrion.*, *113*, 2651–2260.

Jones, A., T. Vukićević, and T. Vonder Haar (2004), A microwave satellite observational operator for variational data assimilation of soil moisture, *J. Hydrometeol.*, *5*, 213–229.

Kalnay, E., et al. (1996), The NCEP-NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc.*, *77*, 437–471.

Kirdyashev, K., A. Chukhlantsev, and A. Shutko (1979), Microwave radiation of the Earth's surface in the presence of vegetation cover, *Radiotekh. Elektr.*, *24*, 256–264.

Koster, R. D., and the GLACE team (2004), Regions of strong coupling between soil moisture and precipitation, *Science*, *305*, 1138–1140.

Lawrence, P. J., and T. N. Chase (2007), Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), *J. Geophys. Res.*, *112*, G01023, doi:10.1029/2006JG000168.

Li, X., et al. (2013), Heihe Watershed Allied Telemetry Experimental Research (HiWATER): Scientific objectives and experimental design, *Bull. Amer. Meteor. Soc.*, doi:10.1175/BAMS-D-12-00154.

Lu, H., T. Koike, N. Hirose, M. Morita, H. Fujii, D. N. Kuria, T. Graf, and H. Tsutsui (2006), A basic study on soil moisture algorithm using ground based observations under dry condition, *Ann. J. Hydraul. Eng., JSCE*, *50*, 7–12.

Njoku, E. G., T. L. Jackson, V. Lakshmi, T. Chan, and S. V. Nghiem (2003), Soil moisture retrieval from AMSR-E, *IEEE Trans. Geosci. Remote Sens.*, *41*, 215–229.

Oleson, K. W., et al. (2004), Technical description of the community land model (CLM), Rep. NCAR/TN-461 + STR, 186 pp., Natl. Cent. For Atmos. Res., Boulder, Colo.

Oleson, K. W., et al. (2007), CLM 3.5 documentation, UCAR, 34 pp, [Available at http://www.cgd.ucar.edu/tss/clm/distribution/clm3.5/CLM3_5_documentation.pdf].

Oleson, et al. (2008), Improvements to the Community Land Model and their impact on the hydrological cycle, *J. Geophys. Res.*, *113*, G01021, doi:10.1029/2007JG000563.

Pan, M., E. F. Wood, D. McLaughlin, D. Eetekhabi, and L. Luo (2009), A multiscale ensemble filtering system for hydrologic data assimilation: Part I, Implementation and synthetic experiment, *J. Hydromet.*, *10*(3), 794–806, doi:10.1175/2009JHM1088.1.

Pellarin, T., J. P. Wigneron, J. C. Calvet, and P. Waldteufel (2003), Two-year global simulation of L-band brightness temperature over land, *IEEE Trans. Geosci. Remote Sens.*, *41*(9), 2135–2139.

Pulliainen, J. T., M. T. Hallikainen, and J. Grandell (1999), HUT snow emission model and its applicability to snow water equivalent retrieval, *IEEE Trans. Geosci. Remote Sens.*, *37*, 1378–1390.

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations, *J. Hydrometeorol.*, *7*, 953–975, doi:10.1175/JHM540.1.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski (2005), Using Bayesian model averaging to calibrate forecast ensembles, *Mon. Weather Rev.*, *133*, 1155–1174, doi:10.1175/MWR2906.1.

Rees, A., J. Lemmetyinen, C. Derksen, J. Pulliainen, and M. English (2010), Observed and modelled effects of ice lens formation on passive microwave brightness temperatures over snow covered tundra, *Remote Sens. Environ.*, *114*, 116–126.

Reichle, R. H., W. T. Crow, R. D. Koster, H. O. Sharif, and S. P. P. Mahanama (2008), Contribution of soil moisture retrievals to land data assimilation products, *Geophys. Res. Lett.*, *35*, L01404, doi:10.1029/2007GL031986.

Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling, *J. Climate*, *19*(13), 3088–3111.

Shi, C. X. (2008), A study on soil moisture remote sensing data assimilation based on ensemble Kalman filter (EnKF) (in Chinese), Ph.D. thesis, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 177 pp.

Shi, C. X., Z. H. Xie, H. Qian, M. L. Liang, and X. C. Yang (2011), China land soil moisture EnKF data assimilation based on satellite remote sensing data, *Sci. China Earth Sci.*, doi:10.1007/s11430-010-4160-3.

Tian, X., Z. Xie, and A. Dai (2008), An ensemble-based explicit four-dimensional variational assimilation method, *J. Geophys. Res.*, *113*, D21124, doi:10.1029/2008JD010358.

Tian, X., Z. Xie, A. Dai, C. Shi, B. Jia, F. Chen, and K. Yang (2009), A dual-pass variational data assimilation framework for estimating soil moisture profiles from AMSR-E microwave brightness temperature, *J. Geophys. Res.*, *114*, D16102, doi:10.1029/2008JD011600.

Tian, X., Z. Xie, and A. Dai (2010a), An ensemble conditional nonlinear optimal perturbation approach: Formulation and applications to parameter calibration, *Water Resour. Res.*, *46*, W09540, doi:10.1029/2009WR008508.

Tian, X., Z. Xie, A. Dai, B. Jia, and C. Shi (2010b), A microwave land data assimilation system: Scheme and preliminary evaluation over China, *J. Geophys. Res.*, *115*, D21113, doi:10.1029/2010JD014370.

Tian, X., Z. Xie, A. Wang, and X. Yang (2011), A new approach for Bayesian model averaging, *Sci. China Earth Sci.*, doi:10.1007/s11430-011-4307-x.

Vrugt, J. A., C. G. H. Diks, and M. P. Clark (2008), Ensemble Bayesian model averaging using Markov Chain Monte Carlo sampling, *Environ. Fluid Mech.*, *8*, 579–595, doi:10.1007/s10652-008-9106-3.

Wang, J. R., and T. Schmugge (1980), An empirical model for the complex dielectric permittivity of soils as a function of water content, *IEEE Trans. Geosci. Remote Sens.*, *18*, 288–295.

Weng, F., B. Yan, and N. C. Grody (2001), A microwave land emissivity model, *J. Geophys. Res.*, *106*(D17), 20115–20123.

Yang, K., T. Watanabe, T. Koike, X. Li, H. Fujii, K. Tamagawa, Y. M. Ma, and H. Ishikawa (2007), An auto-calibration system to assimilate AMSR-E data into a land surface model for estimating soil moisture and surface energy budget, *J. Meteor. Soc. Japan*, *85A*, 229–242.

Zhan, X. W., P. R. Houser, J. P. Walker, and W. Crow (2006), A method for retrieving high resolution soil moisture from Hydros L-Band radiometer and radar observations, *IEEE Trans. Geosci. Remote Sens.*, *44*, 1534–1544.

Zhu, Y. F. (2003), The regional division of dryness/wetness over Eastern China and variations of dryness/wetness in Northern China during the last 530 years (in Chinese), *Acta Geogr. Sin.*, *50*(supplement), 100–107.