

Article

Improving Soil Moisture Estimation with a Dual Ensemble Kalman Smoother by Jointly Assimilating AMSR-E Brightness Temperature and MODIS LST

Weijing Chen ¹, Huanfeng Shen ^{1,*}, Chunlin Huang ^{2,*} and Xin Li ^{2,3}

¹ School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China; chenweijinghn@whu.edu.cn

² Key Laboratory of Remote Sensing of Gansu Province, Heihe Remote Sensing Experimental Research Station, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; lixin@lzb.ac.cn

³ CAS Center for Excellence in Tibetan Plateau Earth Sciences, Chinese Academy of Sciences, Beijing 100101, China

* Correspondence: shenhf@whu.edu.cn (H.S.); huangcl@lzb.ac.cn (C.H.); Tel.: +86-27-6877-8375 (H.S.); +86-931-496-7975 (C.H.)

Academic Editors: George P. Petropoulos, Xiaofeng Li and Prasad S. Thenkabail

Received: 24 January 2017; Accepted: 13 March 2017; Published: 15 March 2017

Abstract: Uncertainties in model parameters can easily result in systematic differences between model states and observations, which significantly affect the accuracy of soil moisture estimation in data assimilation systems. In this research, a soil moisture assimilation scheme is developed to jointly assimilate AMSR-E (Advanced Microwave Scanning Radiometer-Earth Observing System) brightness temperature (TB) and MODIS (Moderate Resolution Imaging Spectroradiometer) Land Surface Temperature (LST) products, which also corrects model bias by simultaneously updating model states and parameters with a dual ensemble Kalman filter (DEnKS). Common Land Model (CoLM) and a Radiative Transfer Model (RTM) are adopted as model and observation operator, respectively. The assimilation experiment was conducted in Naqu on the Tibet Plateau from 31 May to 27 September 2011. The updated soil temperature at surface obtained by assimilating MODIS LST serving as inputs of RTM is to reduce the differences between the simulated and observed TB, then AMSR-E TB is assimilated to update soil moisture and model parameters. Compared with in situ measurements, the accuracy of soil moisture estimation derived from the assimilation experiment has been tremendously improved at a variety of scales. The updated parameters effectively reduce the states bias of CoLM. The results demonstrate the potential of assimilating AMSR-E TB and MODIS LST to improve the estimation of soil moisture and related parameters. Furthermore, this study indicates that the developed scheme is an effective way to retrieve downscaled soil moisture when assimilating the coarse-scale microwave TB.

Keywords: data assimilation; soil moisture; state-parameter estimation; AMSR-E; MODIS; Common Land Model

1. Introduction

Accurate soil moisture estimation at the land surface plays a very important role in studies of land surface conditions, natural resource management, and the interactions in the land-atmosphere system [1–3]. Soil moisture not only governs the partition of rainfall water into runoff and infiltration but also affects the transfer of heat between soil layers and the division of latent heat flux and sensible heat flux at the land surface. Observation and modeling are two typical methods that are used to acquire soil moisture information, while the former one can be divided into ground station

measurements and remote sensing measurements according to the acquisition way [4–6]. However, the unsatisfactory accuracy of modeling, limited coverage area of ground station measurements and coarse spatio-temporal resolution of remote sensing measurements confine their practical applications. Data assimilation effectively controls such shortages by merging the observations into a dynamic model to acquire spatio-temporal continuous state variables [7–10].

Sequential data assimilation provides near real time integration of observations to correct the states of a system. Two of the most common nonlinear filters in sequential data assimilation are ensemble Kalman filter (EnKF) and particle filtering (PF). The EnKF, originated from Kalman filter to adapt to the application of non-linear systems, was first introduced by Evensen [11] in the context of ocean modeling. The EnKF shows its strength by resolving the recursive issue of the forecast covariance matrix using the propagation of a random ensemble. Nowadays, EnKF is widely used in different fields of application, mainly because of its ease of implementation and computational efficiency [12–17]. Notably, EnKF hold the assumption of a linear Gaussian state-space model which can be relaxed by the use of sequential Monte Carlo method in the form PF [18–21]. PF use a number of particles to represent the probability density function of the system states. But, it is well known that the weights degenerate in high-dimensional problems [22].

Brightness temperature (TB) from passive microwave observations have been widely used to estimate soil moisture for the following three reasons: (1) the relaxation frequency of water lies in the microwave region (0.3–300 GHz), resulting in a significant difference in the dielectric constant between dry and wet soils [23]; (2) the features of microwave remote sensing techniques such as penetrability advantage, imaging capability at night and free from atmospheric disturbances (cloud liquid water or integrated water vapor); and (3) the current development of satellite-based microwave missions provides microwave data sets at the global scale, such as SSM/I (Special Sensor Microwave/Image), AMSR-E, SMOS (Soil Moisture and Ocean Salinity), SMAP (Soil Moisture Active Passive), etc. Although the assimilation of soil moisture retrievals from TB has been applied in land surface and hydrological model [24–27], the performances of data assimilation framework involve risk due to the uncertainties brought about by the inversion algorithm. Retrieved soil moisture data may be inconsistent with model simulations because they may utilize different land surface parameters and auxiliary information such as vegetation, soil texture, and surface temperature. Furthermore, errors in retrieved data will be correlated to errors in the auxiliary information, which in turn can be expected to be correlated with the background information used in the data assimilation system [28]. Thus, the direct assimilation of microwave observations into dynamic models is typically considered and has been demonstrated with several studies in improving soil moisture simulations [29–32]. In the framework of directly assimilating TB, the simulated TB data from the Radiative Transfer Model (RTM) are affected by the hydrothermal condition at the surface and the vegetation cover [33,34]. As the foundation of surface radiation property, incorrect soil temperature at the first layer can directly result in the deviations of simulated TB which will eventually adversely impact the soil moisture estimation. Therefore, it is necessary to correct temperature information before it is input into the RTM. Han et al. [35] designed a joint assimilation experiment of land surface temperature and brightness temperature into Community Land Model. The results showed that the joint assimilation experiment resulted in the best characterization of soil moisture and soil temperature profiles under dry conditions. However, such experiment is just synthetic test, thus this research attempts to establish a joint assimilation framework to take full advantage of diverse satellite data.

In addition, the performance of data assimilation largely relies on the effectiveness of the model forecast. However, discrepancies between model descriptions and land surface processes in reality, and the uncertainties in model parameters, forcing data, and initial conditions always cause sustained systematic bias in soil moisture modeling which can deteriorate the performance of data assimilation. Furthermore, the fundamental assumption of the sequential data assimilation algorithm is that no bias exists between the simulations and observations. Thus, it is necessary to eliminate such system deviation before the combination of data and model information in an assimilation

procedure. Parameters uncertainty in land surface model is supposed to be the main source of bias in state variables which calls for a correction procedure. Parameters calibration has been widely used in hydrologic modeling, assimilation and forecast fields with single-objective or multi-objective optimization algorithms [36–38]. However, the optimization process is time consuming and requires abundant historical observational data which is beyond accessibility in most regions. Thus, several approaches have been proposed for simultaneous state and parameter estimation to deal with the issue of parameter correction with state assimilation in land data assimilation [39–43]. Chen et al. [44] compared three ensemble-based state and parameter estimation methods and found that the dual-EnKF can accurately estimate soil moisture and soil texture and has some advantages, including flexibility, simple implementation, and reduced inner storage and operational time requirements. Next, Chu et al. [45] found that dual-EnKS outperformed dual-EnKF for the estimation of soil moisture and that the estimated model parameters (soil texture) were more stable.

The objectives of this study were to explore how to: (1) jointly assimilate coarse-scale AMSR-E TB (0.25 degree) and fine-scale MODIS LST (0.05 degree) to improve soil moisture estimations, and (2) reduce model bias by using a simultaneous state-parameter estimation approach. To achieve this goal, a combination assimilation framework for TB and LST with a parameters estimation procedure is established. LST observations are assimilated to retrieve more accurate soil temperature profile before the calculation of simulated TB in RTM while TB observations are employed to update soil moisture and parameters with dual-EnKS. The performance of the developed soil moisture assimilation framework is evaluated based on observations measured at Naqu, Tibet during the monsoon season. This paper is structurally organized into five Sections. The land surface model and radiative transfer model are briefly described in Section 2, as well as the data assimilation strategy. The study area, data used in this study and experimental design are introduced in Section 3. Section 4 displays the results and discussions for the experiment while some conclusions are drawn in Section 5.

2. Models and Data Assimilation Strategy

2.1. Common Land Model

The CoLM was originally proposed to provide a framework for a truly community developed land component and then effectively implemented as a state of the art model that combined the best features of the LSM (Bonan’s Land Surface Model), BATS (Biosphere-Atmosphere Transfer Scheme), and IAP94 (the 1994 version of the Chinese Academy of Sciences Institute of Atmospheric Physics LSM) codes [46]. CoLM has one vegetation layer, ten unevenly spaced vertical soil layers, and up to five snow layers. Land cover, soil texture and terrain elevation, etc. are necessary for producing the surface conditions. In CoLM, Darcy’s law is used to calculate the soil moisture profile; the temperature variation in a soil layer can be described as a discretization form by the Crank-Nicholson scheme.

The climate conditions contribute to the shallow root systems of meadows in the Tibetan Plateau, which result in high soil organic carbon content (SOC) in the surface soil layer. Remarkably, high SOC may significantly affect soil thermal/hydraulic properties, which brings bias in the estimation of soil hydrothermal variables from CoLM. Therefore, some modifications are necessary for the CoLM parameterization to involve the impacts of SOC on soil condition. The thermal/hydraulic parameters of soil depend on soil composition and are originally calculated based on the soil texture (the percentage of sand and clay) in CoLM. In this study, the modified calculation equations for the related parameters are defined to be the volume-weighted average of the corresponding parameters of soil minerals and peat (the mass of SOC equals 100%) as following [47].

$$\theta_{m,sat} = 0.489 - 0.00126 \times \%sand \quad (1)$$

$$\theta_{sat} = (1 - V_{soc}) \times (0.489 - 0.00126 \times \%sand) + V_{soc}\theta_{soc,sat} \quad (2)$$

$$\rho_d = 2700 \times (1 - \theta_{m,sat}) \quad (3)$$

$$\lambda_{dry} = (1 - V_{soc}) \times \left(\frac{0.135 \times \rho_d + 64.7}{2700 - 0.947 \times \rho_d} \right) + V_{soc} \lambda_{soc,dry} \quad (4)$$

$$\lambda_s = (1 - V_{soc}) \times \left(\frac{8.8 \times \%sand + 2.92 \times \%clay}{\%sand + \%clay} \right) + V_{soc} \lambda_{soc} \quad (5)$$

$$C_s = (1 - V_{soc}) \times \left(\frac{2.218 \times \%sand + 2.385 \times \%clay}{\%sand + \%clay} \right) \times 10^6 + V_{soc} C_{soc} \quad (6)$$

$$\kappa_{sat} = (1 - V_{soc}) \times \left(0.0070556 \times 10^{-0.884 + 0.0153 \times \%sand} \right) + V_{soc} \kappa_{soc,sat} \quad (7)$$

$$\psi_{sat} = (1 - V_{soc}) \times \left(-10.0 \times 10^{1.88 - 0.0131 \times \%sand} \right) + V_{soc} \psi_{soc,sat} \quad (8)$$

$$b = (1 - V_{soc}) \times (2.91 + 0.159 \times \%clay) + V_{soc} b_{soc} \quad (9)$$

where V_{soc} is the volume fraction of SOC; $\theta_{m,sat}$ is the soil porosity of mineral soil; ρ_d is the bulk density of a dry soil; λ_{dry} and λ_s are the soil thermal conductivity values for dry soil and soil solid fraction, respectively; C_s is the soil heat capacity; κ_{sat} is the saturated hydraulic conductivity; ψ_{sat} is the saturated soil matric potential; and b is the Clapp and Hornberger exponent. The related parameters values for peat (labeled with subscript SOC) used in this study are listed in Table 1.

Table 1. Soil hydraulic and thermal parameters used in CoLM for organic soil.

| | λ_s (W·m ⁻¹ ·K ⁻¹) | λ_{dry} (W·m ⁻¹ ·K ⁻¹) | C_s (J·m ⁻³ ·K ⁻¹) | $\theta_{m,sat}$ | κ_{sat} (m·s ⁻¹) | ψ_{sat} (mm) | b |
|------|--|--|--|------------------|--|----------------------|-----|
| Peat | 0.25 | 0.05 | 2.5×10^6 | 0.9 | 0.1×10^{-3} | -10.3 | 2.7 |

2.2. Microwave Radiative Transfer Model

A RTM was selected as an observation operator to link the soil moisture output from CoLM with the brightness temperature from AMSR-E. In this RTM, vegetation is represented as a single-scattering layer above a rough surface, and the simulated brightness temperature can be expressed as follows:

$$T_{B,H(V)} = T_g(1 - \Gamma_{H(V)}) \exp(-\tau_c) + T_c(1 - \omega)[1 - \exp(-\tau_c)][1 + \Gamma_{H(V)} \times \exp(-\tau_c)] \quad (10)$$

where T_B is the brightness temperature, and the subscript $H(V)$ denotes the vertical (horizontal) polarization. T_g and T_c are the soil temperature and canopy temperature, respectively. τ_c is the vegetation optical thickness relevant to the vegetation water content w_c , the incident angle γ and the wavelength $\lambda[m]$ as proposed by Jackson and Schmugge [48]:

$$\tau_c = b(100\lambda)^\chi w_c / \cos \gamma \quad (11)$$

where b and χ are empirical coefficients, which are set to 3.98 and -1.41 , respectively. These values are suitable for the leaf-dominated field and chosen based on a summary of the relationship between vegetation parameter and wavelength for a variety of vegetation types, as described by Jackson and Schmugge [48]. w_c can be obtained according to Paloscia and Pampaloni [49].

$$w_c = \exp(LAI/3.3) - 1 \quad (12)$$

where LAI (m²/m²) is the leaf area index.

ω is the single-scattering albedo of the vegetation and is calculated using the following empirical formula proposed by Yang [50].

$$\omega = 0.00083/\lambda \quad (13)$$

The Q - h model developed by Wang and Choudhury [51] is used to calculate the soil reflectivity Γ :

$$\Gamma_{H(V)} = [(1 - Q) \cdot R_{H(V)} + Q \cdot R_{V(H)}] \exp(-h) \quad (14)$$

$$Q = 0.35[1 - \exp(-0.6s^2\lambda)] \quad (15)$$

$$h = (2ks \cos \gamma)^2 \quad (16)$$

where Q and h are empirically determined surface roughness parameters, k is the wave number defined as $2\pi/\lambda$, and s is the root-mean-square height (RMS). R is the Fresnel power reflectivity which describes the soil reflectivity of a smooth surface and is calculated by:

$$R_H = \left| \frac{\cos \gamma - \sqrt{\varepsilon_r - \sin^2 \gamma}}{\cos \gamma + \sqrt{\varepsilon_r - \sin^2 \gamma}} \right|^2 \quad (17)$$

$$R_V = \left| \frac{\varepsilon_r \cos \gamma - \sqrt{\varepsilon_r - \sin^2 \gamma}}{\varepsilon_r \cos \gamma + \sqrt{\varepsilon_r - \sin^2 \gamma}} \right|^2 \quad (18)$$

ε_r is angle soil dielectric constant that depends on the soil moisture θ [52]:

$$\varepsilon_r = [1 + (1 - \rho)(\varepsilon_s^\alpha - 1) + \theta^\beta \varepsilon_{fw}^\alpha - \theta]^{1/\alpha} \quad (19)$$

where $\varepsilon_s = 4.7 + 0.0j$ denotes the dielectric constant for a mineral soil in which j is the imaginary part of plural, ρ is the soil porosity, ε_{fw} is the dielectric constant of free water. $\alpha = 0.65$, and β is a soil texture-dependent coefficient [53].

2.3. Data Assimilation Strategy

The assimilation strategy is illustrated in the flowchart shown in Figure 1. Firstly, to diminish the influence of temperature on the simulated TB, MODIS LST observations are assimilated to retrieve more accurate soil temperature profile. EnKF was used to update the soil temperature because the step-by-step feedback can be used to immediately modify the temperature information input into the RTM. Then, the updated soil temperature, the soil moisture and the related parameters are imported into the RTM to calculate the simulated TB. Next, given that parameters uncertainty always causes the long-term bias of soil moisture as well as the simulated TB, a scheme with two-parallel-filters was designed to synchronously estimate the states and parameters and the related parameters using AMSR-E TB observations. In this part, dual-EnKS is used to consider the relationship between states/parameters and observations distributing at the time scale. The first stage was used to estimate the model parameters (they are supposed to be constant between two consecutive observations) while the second stage was employed to update states variables. The two stages evolved in parallel rather than successively, and the data assimilation procedure marched forward without restarting. The updated soil moisture and parameters were only returned to CoLM at the end of every smoother window and acted as the initial conditions for the next smoother window.

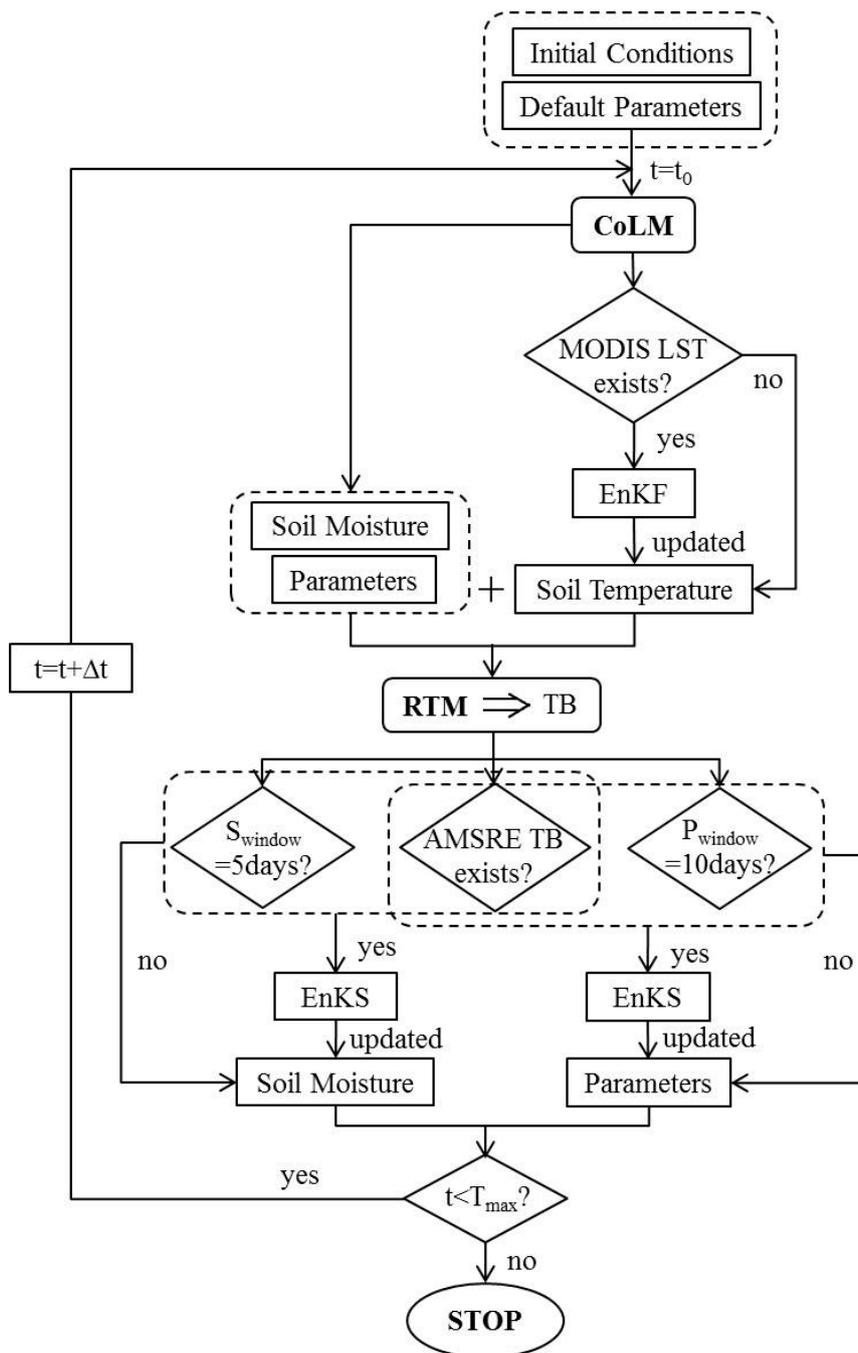


Figure 1. Flowchart of the assimilation Strategy. S_{window} and P_{window} represent the smoother window of dual-EnKS for soil moisture and parameters, respectively.

2.3.1. Ensemble Kalman Filter

The CoLM implemented with EnKF starts with an initial ensemble (obtained by perturbed the initial state field) and then the simulated states variables propagate through the model, which can be described as follows:

$$X_{i,k+1}^f = M(X_{i,k}^a, \alpha_{k+1}, \beta_{k+1}) + w_i \quad w_i \sim N(0, Q) \tag{20}$$

where $M(\bullet)$ represents a nonlinear model operator (the CoLM in this study) which determines how the states evolve over time. The subscript $k + 1$ represents the time step of model. The superscripts a and f refer to the analysis and forecast states, respectively. $x_{i,k+1}^f$ is the i th member of the forecast

ensemble. α and β stand for atmospheric forcing data and model parameters, which are essential to drive the CoLM. The model error representing the uncertainties related to the model parameters and the model structure is indicated by w_i which conforms to a Gaussian distribution noise field with a mean of zero and the covariance matrix Q .

When observations are available, a linear correction equation is used to update the forecasted state ensemble members as follows:

$$X_{i,k+1}^a = X_{i,k+1}^f + K_{k+1}(Y_{i,k+1} - \hat{Y}_{i,k+1}) \quad (21)$$

where $Y_{i,k+1}$ is generated by adding the stochastic perturbation into the actual observation Y_{k+1} with a mean of zero and the covariance matrix R . The observation error matrix R is assumed to be mutually independent and temporally steady. $\hat{Y}_{i,k+1}$ is the projection of the model state in the observational space based on the conversion of the observation operator $H(\bullet)$, which establishes a relationship between the model states and observations as $\hat{Y}_{i,k+1} = H(X_{i,k+1}^f)$. K_{k+1} is the Kalman gain matrix, which is calculated as follows:

$$K_{k+1} = P_{k+1}^f H^T (H P_{k+1}^f H^T + R)^{-1} \quad (22)$$

P_{k+1}^f is the forecast error covariance matrix. $P_{k+1}^f H^T$ is the cross covariance between the model state forecasts x_{k+1}^f and their projections $H(x_{k+1}^f)$ in the observation space, and $H P_{k+1}^f H^T$ is the error covariance of $H(x_{k+1}^f)$.

Finally, the analysis state variable at the time $k + 1$ is given by the averaged value of the ensemble members. The analyzed ensemble is then integrated forward until the next observation becomes available, and the update process is repeated.

2.3.2. Ensemble Kalman Smoother

To consider the relationships among states or observations that are distributed in time or space and to solve the “sawtooth effect” brought about by the step-by-step update of EnKF, a smoother solution—the EnKS [54–56] is applied to update all the states in a predefined calculation window using the whole observations including in such window.

The implementation of the EnKS requires that the states and observations ensembles must be stored during the evolution of the current smoother window. The CoLM runs forward throughout the current assimilation window at a time to obtain the overall forecasting states in the smoother window, and the augmented state vector X contains the states extending in the temporal space and can be expressed as $X = [x_1 x_2 \cdots x_k]^T$. Correspondingly, the augmented observation vector Y contains all the observations available in the smoother window with $Y = [y_1 y_2 \cdots y_k]^T$. The update of the states stored in the vector X is affected by their relationship with all the observations distributing in the vector Y . The Kalman gain matrix used to update the state vector is calculated by the same formula in EnKF mentioned above.

2.3.3. Inflation of Background Error Covariance

The performance of the assimilation experiment is largely influenced by the presentation of background error covariance, which tends to be underestimated because of the inappropriate presence of model errors and the sampling error caused by the limited ensemble size. This problem can progressively get worse, potentially resulting in a condition called “filter divergence”, in which the ensemble variance becomes vanishingly small and observation information is completely ignored [57]. Therefore, additive inflation was applied to relax the posterior (analysis) perturbations.

$$x_i^a = (1 - \alpha)x_i^a + \alpha x_i^f \quad (23)$$

where x_i^a and x_i^f are the i th member of the deviation ensemble calculated by subtracting the posterior ensemble mean and the prior ensemble mean from the corresponding raw ensemble. α is the inflation factor which is set to 0.5 in this research.

2.3.4. Relaxation Factor for Parameters Updating

The strong update of parameters may cause problems in two aspects [58]: overcorrection for parameters by data assimilation technique because the relationship between TB and parameters is more implicit compared with the relationship between TB and soil moisture; ignorance the observations information at the later stage of experiment period because of the small variance of parameters ensemble. Thus, to prevent non-feasible updates of parameters, a relaxation factor β was used for parameters update.

$$\Psi_{i,k+1}^a = \Psi_{i,k+1}^f + \beta K_{k+1} (Y_{i,k+1} - \hat{Y}_{i,k+1}) \quad (24)$$

where $\Psi_{i,k+1}^a$ and $\Psi_{i,k+1}^f$ are the i th member of the analyzed and forecasted parameters ensemble at the time step $k + 1$, respectively. β is set to 0.45, which was calibrated by minimizing the root-mean-square error between the simulated TBs and AMSR-E TBs.

3. Data and Experimental Design

3.1. Data

The Tibetan Plateau is known as the Third Pole with an average elevation of over 4000 m above sea level and covers an area of approximately 2.5×10^6 km². The study area is located in the central region of the Tibetan Plateau and spreads around the town of Naqu over an area of ~ 100 km \times 100 km (91°30'E–92°30'E and 31°N–32°N) (Figure 2). The mean elevation of the study area is about 4650 m and the land cover is primarily alpine meadow with a fraction of water bodies in the western part. This region has a typical semiarid monsoon climate with an annual precipitation amount of approximate 500 mm while most precipitation events take place during May to October. In addition, permafrost is an important part of the natural ecosystems on the Tibetan Plateau with soil thawing-freezing process occurring around May and November in this region. High altitude and high atmospheric transparency contribute to low temperature and strong solar radiation, which are associated with very low biomass on the ground and less water vapor in the atmosphere. The soil is predominantly composed of sand and silt, meanwhile all the climate conditions lead to the shallow root system of vegetation in this region. Thus, the soil organic/carbon content at the top layer turns to be fairly high.

3.1.1. Soil Moisture and Temperature Network

Ground measurements were collected through a mesoscale Central Tibetan Plateau Soil Moisture and Temperature Monitoring Network (CTP-SMTMN) that exactly coincides with the domain of the study area [59]. The SMTMN network was established to study the mechanism of soil-vegetation-atmosphere interactions and to validate satellite soil moisture and temperature products [60]. This network was accomplished in the following three stages: (1) initialization with 30 stations in July 2010 along four branches of the roads gathering at Naqu; (2) enhancement of the observations within a 0.25° quadrature grid by adding 20 stations in July 2011; and (3) the installation of another 5 stations in an ~ 5 km \times 5 km area in June 2012 (Figure 2). At each site, four sensors (ECH2O EC-TM/5TM capacitance probes) were used to monitor soil moisture and temperature at depths of 0–5 cm, 10 cm, 20 cm, and 40 cm, respectively. The time interval of measurement was 30 min, and each record reflects the average state of the soil over the most recent half-hour. The observations were processed to hourly values by arithmetic average to be used in the following results comparison.

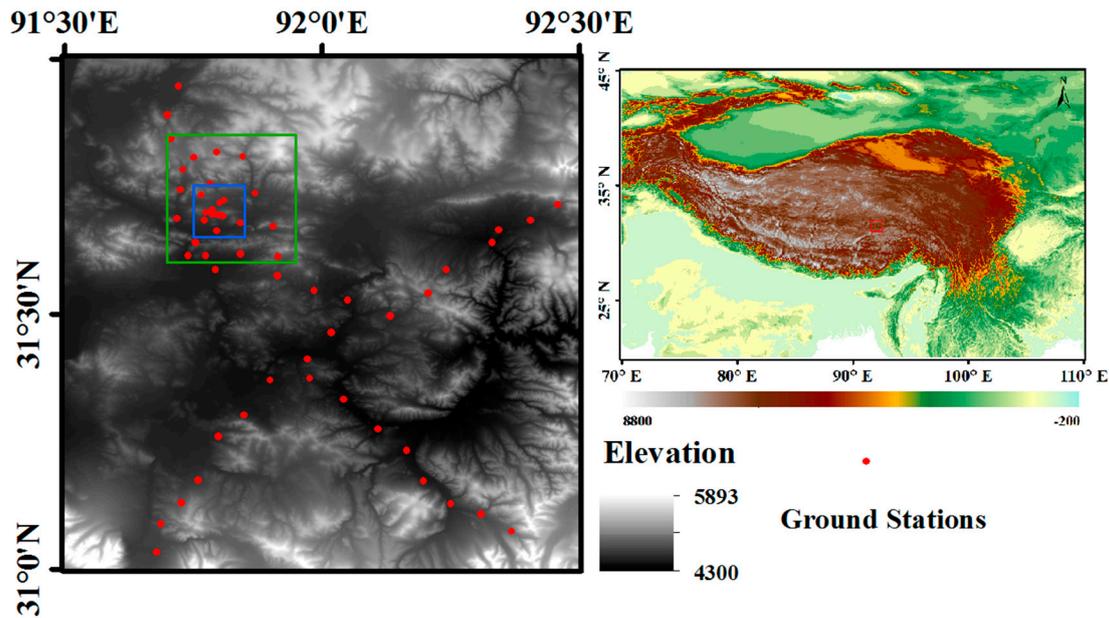


Figure 2. Spatial distribution of the ground stations in the experiment area (the central Tibetan Plateau). The blue and green rectangles indicate the intensive observation grids in CTP-SMTMN.

3.1.2. Forcing Data

The forcing data used to drive CoLM were developed by the hydrometeorological research group at the Institute of Tibetan Plateau Research at the Chinese Academy of Sciences (ITPCAS), which include seven essential meteorological variables (pressure, precipitation, wind speed, specific humidity, near surface air temperature, downward shortwave radiation and downward longwave radiation). This data set were produced by merging the observations collected at 740 operational stations of the CMA (China Meteorological Administration) into the corresponding Princeton meteorological forcing data, the GLDAS (Global Land Data Assimilation Systems) reanalysis forcing data, the TRMM (Tropical Rainfall Measuring Mission) 3B42 precipitation products and the GEWEX-SRB (Global Energy and Water Cycle Experiment—Surface Radiation Budget) radiation data [61]. The ITPCAS forcing data completely cover the geographic scope of China with a spatial resolution of 0.1 degree and a temporal resolution of 3 h [62]. To meet the demand of CoLM, MicroMet (A Meteorological Distribution System for High-Resolution Terrestrial Modeling) [63] is adopted to downscale the atmospheric forcing data, which is mainly based the relationships between meteorological variables and the surrounding landscape (primarily topography). Finally, we obtained a set of forcing data with a higher spatial resolution (0.05 degree) and higher temporal resolution (1 h) for this study area.

3.1.3. Satellite Data

- MODIS LST and LAI data

The MODIS sensors reside aboard both Aqua (1:30 a.m./1:30 p.m.) and Terra (10:30 a.m./10:30 p.m.) platforms, which provide a means for quantifying land surface characteristics (<http://modis-land.gsfc.nasa.gov/>). MOD11C1, MYD11C1 and MCD15A3 are used in this research.

MOD11C1 and MYD11C1 produce daily land surface temperature (LST) deriving from Terra and Aqua respectively in the form of a pair of daytime and nighttime observations configured on a 0.05° geographic CMG (climate modeling grid). The QC fields stored in MOD11C1 and MYD11C1 provide careful scrutiny for the quality of the LST products. LST data from MODIS are often contaminated by clouds, thus we only select the LST data as observations for the assimilation experiment when the quality control flag was zero.

MCD15A3 is a composited product that provides global LAI and FPAR data every 4 days at a resolution of 1 km. To match the spatial resolution of CoLM in this research, the original LAI values were resampled and re-projected to 0.05° from a sinusoidal projection to a UTM projection with a WGS84 datum based on nearest neighbor interpolation. Unreasonable LAI values were filtered out using asymmetric Gaussian model function and then were incorporated into CoLM to substitute the default LAI values.

- AMSR-E brightness temperature data

AMSR-E is a radiometer operating onboard the Sun synchronous and polar orbiting AQUA satellite, which has provided passive microwave measurements at six bands, ranging from 6.9 to 89 GHz. AMSR-E/Aqua Daily Global Quarter-Degree Gridded Brightness Temperatures (NSDIC-0302) were employed as observations in this study. These data are provided in one global cylindrical, equidistant latitude-longitude projection with a resolution of 0.25 degree (quarter-degree), which means the study area includes 16 AMSR-E grids.

3.1.4. The Upscaled Surface Soil Moisture Data

In order to validate the simulated and assimilated surface soil moisture in coarser grid (0.25 and 1.0 degree), an upscaling algorithm, proposed by Qin et al. [64], was adopted to upscale the point-scale surface soil moisture to the grid-scale, in which the upscaled soil moisture was assumed to equal the weighted-average of the measurements from the stations within the grid. This method can be used to address the representative problem of ground stations that results from the heterogeneity of the land surface. First, construct a functional relationship between the station-averaged daily soil moisture and the pixel-averaged daily MODIS-derived apparent thermal inertia (ATI). This relationship was used to calculate the representative soil moisture at a certain spatial scale, which served as the truth to obtain the weight of different stations by Bayesian linear regression. The upscaled ground-based soil moisture observations served as supplemental data are also employed in the results comparison to resolve the mismatch between the ground stations and satellite pixel. A complete description of this upscaled method is illustrated by Qin et al. [64].

3.2. Experimental Design

Given the distinct hydrothermal processes between the frozen stage and unfrozen stage of soil, the experiment of this research was conducted during the soil-thawing period. The experiments for simulation and assimilation were designed to last 120 days commencing on DOY 151, 2011 (31 May 2011). An entire year forcing data before the start date of the experiment (from 31 May 2010 to 30 May 2011) were looped ten times to spin up CoLM to obtain a stable and reasonable distribution of the initial state variables. The initial ensemble size is set to 50 for the simulation/assimilation experiment. To represent the uncertainty of initial condition, 2 K additive noise and 20% multiplicative noise were added to the initial soil temperature profile and soil moisture profile which are independent with each other. According to the sampling depths at the ground stations of CTP-SMTMN, the soil layers of CoLM should be redistricted before the model initialization and the new soil nodes were set to depth of 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4 and 1.6 m. The CoLM operate at resolution of 1 h time step and 0.05 degree.

AMSR-E TBs and MODIS LSTs were introduced as observations in this research. MODIS LST product took charge of providing the temperature information at the fine scale. Due to the fine resolution of forcing data and MODIS LST, the soil moisture and temperature as well as the simulated TBs differed from each other at the fine resolution. Thus, the simulation of soil moisture can be improved at the fine spatial resolution (0.05 degree) despite the assimilation of coarse-scale AMSR-E brightness temperature observations (0.25 degree). It is assumed that each AMSR-E grid has the same impact on the 25 model calculation grids that it contains. Lower microwave frequency corresponds with the higher relative sensitivity of brightness temperature to soil moisture. Therefore, the TBs with

vertical polarization of 6.9 GHz were assimilated into CoLM to improve the estimation of soil moisture and parameters. The observation ensembles were acquired by disturbing the original TBs and LSTs with an additive error of 2 K. The smoother window of the EnKS for parameters estimation was set to 10 days [56] while the window for soil moisture estimation was set to 5 days considering that the model parameters were more stable than soil moisture on a time scale.

The parameter vector to be estimated in this research comprised five parameters, sand content, clay content, soil porosity and SOC fraction from CoLM and RMS from RTM. The initial parameters sets were uniformly sampled in the specific ranges, which are defined according to the model default parameter and the physical range. In order to guarantee the physical meaning of the parameters in the updating process, the sand and clay content were re-checked to ensure that their sum was less than 100%. The extra quantity must be subtracted by adjusting the sand and clay content with $((\%sand + \%clay) - 100)/2.0$ [44].

The variables output from a land surface model largely response to the characteristic of the meteorological field. To account for the uncertainties implying in the meteorological data, a multivariable random field was applied to the forcing data in this study. Generally, the meteorological variables implicitly influence each other, e.g., a positive perturbation of the downward shortwave radiation tends to be associated with negative perturbations to the longwave radiation and the precipitation, and vice versa [65]. Therefore, normally distributed additive perturbations were applied to air temperature and longwave radiation and log-normally distributed multiplicative perturbations were added to precipitation and shortwave radiation. The mean values for the perturbed factors were equal to zero in the additive case and one in the multiplicative case. The standard deviation error of variables and the cross-correlations between variables about the perturbation for the forcing data are listed in Table 2.

Table 2. Summary of the perturbation parameters for the forcing data.

| Variables | Noise Type | Standard Deviation | Cross Correlation |
|---------------------|----------------|---------------------|--------------------|
| Precipitation | Multiplicative | 0.5 | [1.0 -0.8 0.5 0.0, |
| Shortwave radiation | Multiplicative | 0.3 | -0.8 1.0 -0.5 0.4, |
| Longwave radiation | Additive | 30 W/m ² | 0.5 -0.5 1.0 0.4, |
| Air temperature | Additive | 2 K | 0.0 0.4 0.4 1.0] |

3.3. Evaluation Metrics

To assess the performance of the assimilation experiment, we defined several evaluation measures, including the root mean square error (RMSE), the mean bias error (MBE), and the normalized error reduction (NER):

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (X_t - X_{true,t})^2} \quad (25)$$

$$MBE = \frac{1}{T} \sum_{t=1}^T (X_t - X_{true,t}) \quad (26)$$

$$NER = 1 - \frac{RMSE_a}{RMSE_o} \quad (27)$$

where T is the step number; X_t and $X_{true,t}$ represent the results of simulation/assimilation experiment and the true values at step t , respectively. $RMSE_o$ and $RMSE_a$ represent the RMSE of the simulation and assimilation scenarios, respectively.

4. Results and Discussions

4.1. Evaluation of States in 0.05 Degree Model Grids

The spatio-temporal resolution of CoLM in this research was set to 0.05 degree and one hour, as well as the output variables (soil moisture and soil temperature). Thus, the experiment area includes 400 CoLM grids, among which only 49 grids exist ground stations (in situ measurements). The results of the simulation and assimilation experiment were compared with the arithmetic-average of all the in situ measurements included in the corresponding grid.

Thus, the MBE and RMSE for soil moisture and temperature of the 49 grids calculated by the hourly value during the whole experiment period are present by the box-plots in Figure 3. The box-plots reflect the distribution of error in which the quartiles, median and extreme values are marked by solid lines and the abnormal values are denoted by plus signs. The results of the four layers are arranged from left to right with increasing soil depth in every subplot. OLa and DA represent the simulation and assimilation experiment, respectively, and both are forced by same initial conditions (parameters, states) and forcing data. The MBE from OLa in Figure 3a showed a clear underestimation of soil moisture at the first two layers while the deeper layers showed the conflicting situations (overestimation) in most case. Obviously, MBE for soil moisture at the first layer from DA distributed around the 0 value which indicated a large reduction of bias after assimilation. In terms of the MBE for the first two layers, their mean values dropped from -0.138 , -0.075 (OLa) to -0.042 and 0.015 (DA). Meanwhile, the RMSE for soil moisture in Figure 3b also showed a remarkable decrease at the first layer (mean value from 0.149 to 0.085) and slight decrease at the second layer (mean value from 0.094 to 0.076). Noticeably, the spread of the quartiles of MBE and RMSE for soil moisture at the first two layers were narrowed after assimilation which indicated that the improvement of the soil moisture estimation at most single grids surpassed the average level. However, the deeper layers showed dissatisfied results from the assimilation experiment mainly due to the opposite trend of soil moisture simulation from CoLM compared to the first layer. AMSR-E TB improved the accuracy of soil moisture at the first layer by compensating the underestimation phenomenon, which meant the increase of soil moisture at the first layer. Therefore, the infiltration of water from the first layer to deeper layers deteriorated the soil moisture estimation at deeper layers from overestimation to more serious overestimation. Given the penetration depth of microwave, TB can only represent the soil moisture information of a few centimeters under the ground. Thus, the improvement of soil moisture at deeper layers largely depends on the development of model structure especially the water transfer mechanism.

As shown by the MBE from OLa for soil temperature in Figure 3c, the soil temperature simulation at all the four layers appear overestimation. Indicated by the MBE from DA, the overestimation phenomenon of the soil temperature at four layers was obviously alleviated after assimilation. In terms of the mean value of MBE, all the reductions for the four layers reached approximately 1 K with the corresponding value from 1.053 , 1.132 , 1.365 and 1.769 to -0.166 , -0.039 , 0.211 and 0.673 . Meanwhile, the RMSE showed a less significant improvement as seen in Figure 3d, and the reduction of RMSE at deeper layers slightly surpassed that at first layer which mainly reflected by the narrowed spread of the quartiles. Considering that MODIS LST data represent the land surface temperature which should be more relevant to the soil temperature at shallower layers. Thus the improvement of soil temperature profile owed to the coaction of the assimilation of MODIS LST and the improved estimation of soil moisture. Soil moisture influences the heat conductivity at the interfaces of layers and the heat storage in different layers. Moreover, the raise of water content in the deeper layers results in the decline of the soil temperature.

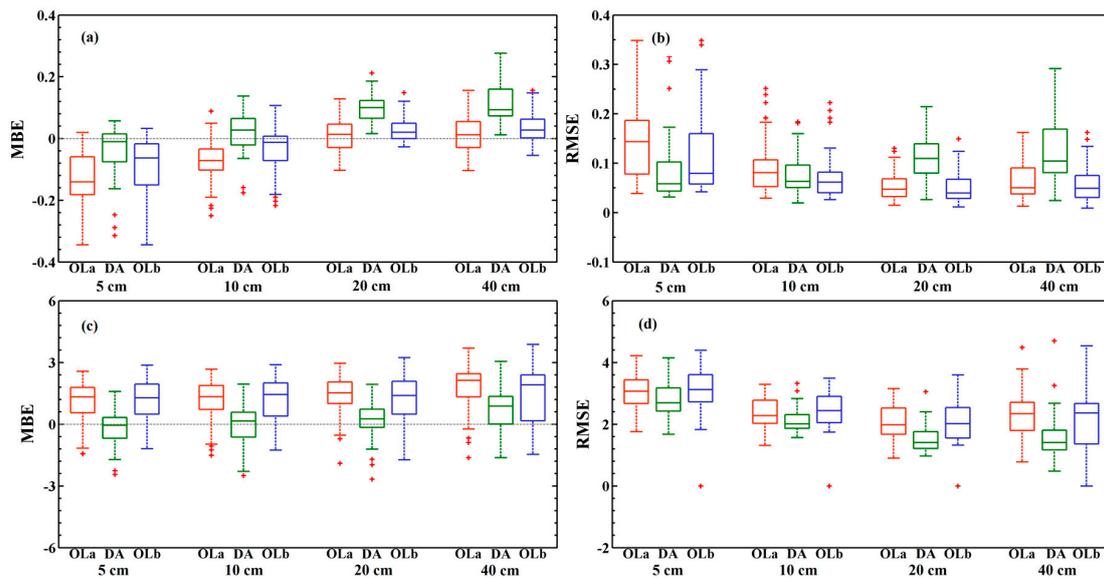


Figure 3. Box plots of the mean bias error and the root mean square error for soil moisture (a,b) and temperature (c,d) at four layers from simulation using default parameters (OLa), assimilation (DA) and simulation using updated parameters (OLb).

The soil moisture measurements of the calculating grids (0.05 degree) within a single AMSR-E grid appear to be diverged from each other in terms of their mean value and variability. Although an overwhelming improvement was demonstrated for the soil moisture at the first layer, such satisfied results did not benefit to every single grid among the experiment area. Therefore, we further investigated the relationship between the performance of the assimilation experiment and the characteristic of in situ measurements, taking the soil moisture at first layer as an example. The NER values of soil moisture at the first layer for 49 grids are plotted in Figure 4 with the mean and variability of the corresponding measurements. The colors of the dots denote the magnitude of the NER. By careful scrutiny of the relationship, we artificially split the plot by two dash lines ($0.25 \text{ cm}^3/\text{cm}^3$ for the mean of measurements and $0.05 \text{ cm}^3/\text{cm}^3$ for the variability of measurements). Clearly, all the weak results occurred at the grids with lower mean value of soil moisture measurements (less than $0.25 \text{ cm}^3/\text{cm}^3$). Conversely, the NERs distributing above the horizontal dash line showed the RMSE reduction of DA maximizing to 80% and surrounding 50% in most grids. A comparison of the two zones below the horizontal dash line suggested that the assimilation experiment completely failed to capture the soil moisture at the grids with both lower mean value and variability of in situ measurements. To investigate the reason for the discrepancy in the soil moisture measurements at different grid, we compare the precipitation of each grid which is the most sensitive forcing data for soil moisture. However, precipitation used to force CoLM showed small discrepancy among the calculating grids, thus either the grids with lower mean value and variability measurements failed to respond to the precipitation events or the precipitation data did not reflect the real condition at these grids. The former can be concluded as the error of measurements which can be the reason for the failure of assimilation experiment at these grids while the latter can be solved by data assimilation. Another reason for the poor results at these grids may explained by the representativeness from a 0.25 degree grid to a 0.05 degree grid (AMSR-E TB of one grid tends to capture the dominate information of soil moisture).

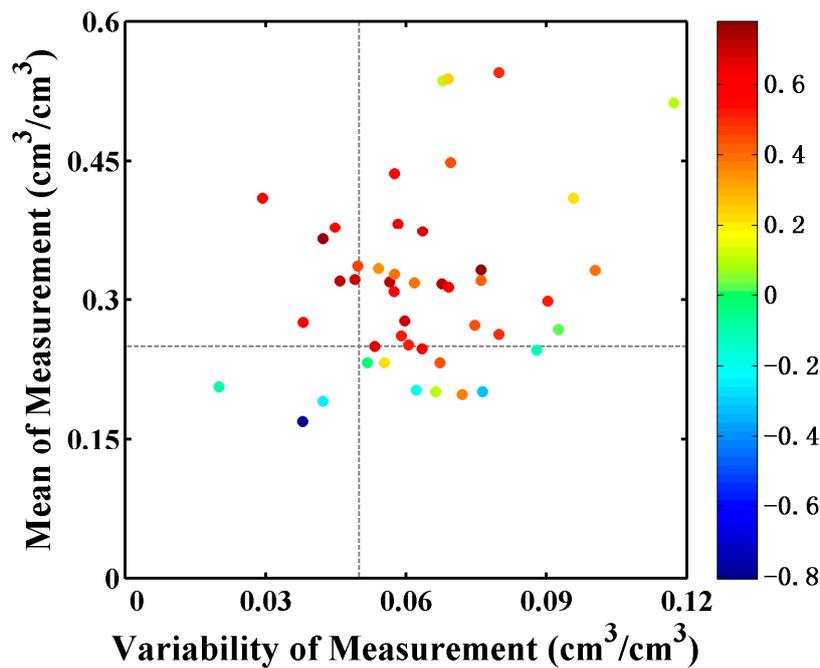


Figure 4. Normalized error reduction for all the 49 grids varying with the mean and variability of the measurement. The two dashed lines represent the specific values analyzed in the manuscript for the mean of measurement ($0.25 \text{ cm}^3/\text{cm}^3$) and the variability of measurement ($0.05 \text{ cm}^3/\text{cm}^3$).

4.2. Evaluation of Parameters Estimation

Considering the resolution of the TBs used to estimate parameters, the comparison of the updated parameters was conducted at the scale of 0.25 degree. The mean value of the estimated and default parameters (sand content, clay content, soil porosity, SOC fraction and RMS) for all the 16 AMSR-E grids within the experiment area are displayed in Figure 5. The parameters measurements are the average values derived from a number of soil samples as described by Yang et al. [66]. First, five parameters showed a large discrepancy with the default parameters which indicated the necessity of performing parameters estimation and demonstrated the effect of TBs for update these parameters. The updated parameters approached the measurements when compared with the corresponding default value, except the clay content. Perfect match to the parameters measurements cannot be expected due to the errors from model structure, etc. Moreover, parameters estimated by data assimilation technique tend to produce results that can assist the state assimilation. To verify the usefulness of parameters estimation, the updated parameters produced by DA were re-input into CoLM while the others conditions were kept exact same as OLa. Then, the simulations of soil moisture and temperature were compared with the in situ measurements. The error statistics are also displayed in Figure 3 and marked by OLa, OLa successively reduced the MBE and RMSE for soil moisture particularly at the first two layers. Little improvement was found in the deeper layers because the parameters retrieved by assimilating surface observations were not supposed to always improve the state accuracy in the deeper layers due to the vertical heterogeneity of the solum. However, negative influence of the updated parameters for soil temperature was also presented in Figure 3c,d. TBs used to estimate parameters highly relate to the surface soil moisture rather than the soil temperature. Thus, the retrieved parameters, contributing in match the surface soil moisture to the in situ measurements but rarely confirmed the simultaneous improvements for multiple states. The above results also indicated that the parameters calibration by a single type of observation conspires against the global optimum for state variables simulation [67]. In conclusion, there is still an ample space to explore for the robust interaction between states and states as well as states and parameters to achieve the improved estimation of multiple-states.

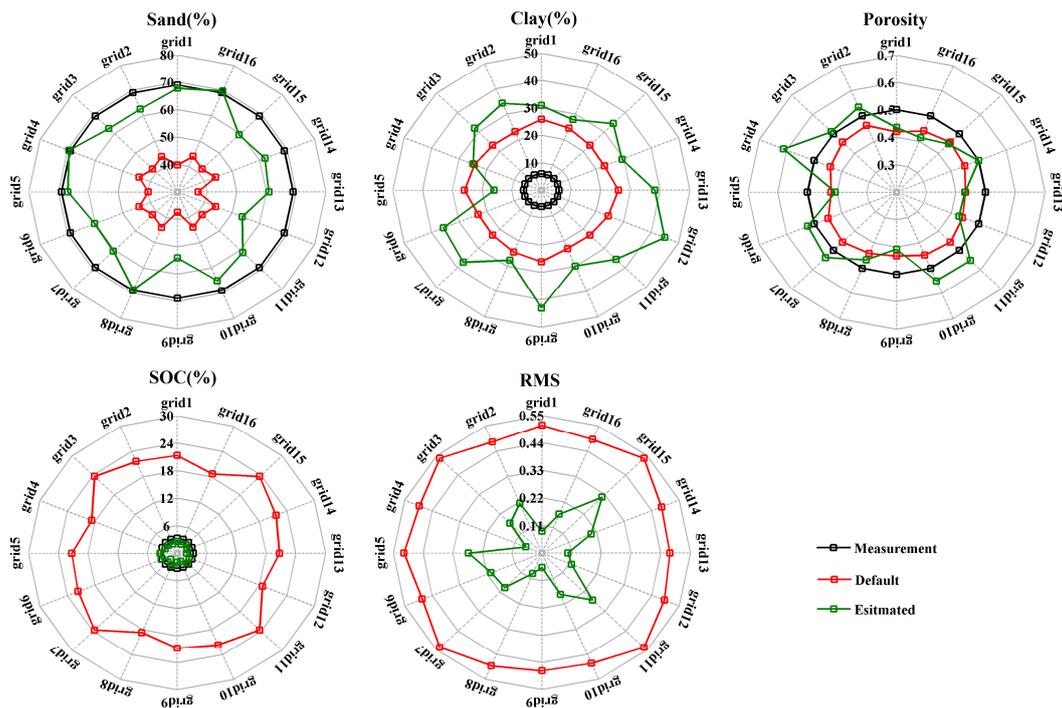


Figure 5. Comparison of the estimated parameters from DA and the default parameters with the measurements which are from Yang et al. [66].

4.3. Evaluation of States within Three Coarser Grids

Two densely observed grids covering 0.1 degree and 0.25 degree are marked by the blue and green rectangle in Figure 2. The results from these two grids and the entire experiment area (referred to as the small, medium and large grid) are presented in detail to demonstrate the performance of the assimilation experiment. The spatially-averaged and hourly value of soil moisture at four layers (5 cm, 10 cm, 20 cm and 40 cm) for the small, medium and large grid are plotted in Figure 6a–l, respectively. OLa showed large system bias between the soil moisture simulations and the measurements at the first layer in Figure 6a,e,i while DA accomplished great improvement in the soil moisture estimation. DA cut down the MBE and RMSE value for soil moisture at the first layer from -0.1547 and 0.1594 to -0.0155 and 0.0397 in the small grid, from -0.1324 and 0.1363 to -0.0299 and 0.0371 in the medium grid and from -0.1308 and 0.1327 to -0.0567 and 0.0598 in the large grid, respectively. The dramatic decrease of the RMSE value produced by assimilation both exceeded 70% at the small and medium grid while inferior results of the RMSE reduction at the large grid still reached up to 55%. The shortage of ground stations in a major portion of the large grid potentially led to an irrational comparison using simple arithmetic mean of measurements which will be discussed later in this section. The smoother windows for soil moisture and parameters estimation reduced the feedback frequency of the updated soil moisture and parameters into CoLM which led to an unapparent improvement of soil moisture during the earlier stage of the experiment period. However, the curves of DA rapidly approached the curves of measurements after a short adjustment period, which proved the effectiveness of parameters estimation in the assimilation experiment. The promising results of DA demonstrated the feasibility to obtain highly accurate soil moisture information at surface by assimilating the microwave observations.

The depth of the second layer is relatively shallow (10 cm), and the measurements at the first two layer are highly correlated which allows the correction of the soil moisture at the first layer to coincidentally improve the accuracy at the second layer. It was obvious that soil moisture from DA at the second layer also achieved satisfied results with the MBE for the three grids declining to below 0.03 and the reduction of RMSE varying from 36% to 72%. The raise of soil moisture at the first layer inevitably caused the increase of soil moisture at the deeper layers. However, the original

simulations at deeper layers (OLa) were closer to the measurements which eventually caused the problem of overcorrection for soil moisture. In addition, judging from the soil moisture at different layers in the same grid, the variability of soil moisture at deeper layers resembled to the variation at the first layer after the assimilation experiment. However, as indicated by the measurements, soil moisture at the last two layers seemed to be insusceptible to the water quantity changes due to natural processes in the upper layers. Obviously, the process of assimilation violated the inherent water flow delivery mechanism in reality which also indicate the spurious correlations between observation and soil moisture at the deeper layers. Thus, the utility of the observations that only imply the surface information to retrieve the information of soil profile can be unreasonable or require a more sophisticated assimilation scheme.

In Figure 6, OLB stands for the model simulations derived by using the updated parameters to examine the impact of parameters estimation. Parameter calibration ameliorated the underestimation of soil moisture at the first layer, reducing the RMSE by 46% in the small grid, by 33% in the medium grid and by 25% in the large grid. The curves of the soil moisture profile were generally upraised after the parameters substitution, which induced the overestimation of the soil moisture simulations at deeper layers at the same time. Unsurprisingly, the parameters were reversed to optimize the surface soil moisture using TB observations that were not suitable for the deeper layers. The heterogeneity of the solum must be treated more carefully in further research. The curves of OLa and OLB responded to precipitation similarly as the measurements but the latter can correctly grape the peaks. Differ from the measurements at the surface with steady high soil moisture that can maintain the water in soil, both OLa and OLB showed swift change of water with a rapid decrease after the water intake process (e.g., precipitation) which reflected the unreasonable infiltration scheme inherent in the CoLM for the experiment area.

Upscaled soil moisture derived by MODIS ATI-based upscaling algorithm was also applied to validate the performance of the assimilation experiment. Considering the restricted coverage and the sufficient quantity of ground stations in the small grid, the comparison with the upscaled measurements was only conducted for the medium and large grid. Figure 7 displays the comparison of soil moisture at the surface of the medium (Figure 7a) and large grid (Figure 7b) during the experiment period. OBS and upOBS denote the daily soil moisture measurements calculated by arithmetic-average and the upscaling algorithm for multiple stations, respectively. OL and DA are the daily soil moisture from the model simulation and assimilation which present the station-average value in the corresponding grid. DA enhanced the features capture with the dashed black line and maintained good agreement with the upscaled measurements. The MBE and RMSE for soil moisture at the medium grid descended to -0.007 and 0.025 (DA) from -0.109 and 0.113 (OL). Unlike the inferior results compared with the small and medium grids when took the arithmetic-average ground station measurements as truth hereinbefore, the improvement of soil moisture estimation in the large grid based on the upscaled measurements were remarkable, with the MBE dropping from -0.089 to -0.015 and the RMSE dropping from 0.092 to 0.027 . Such results confirmed the inference mentioned above and the representativeness issue (from point to area) must be properly resolved, especially in the area with disperse in situ measurements (such as the large grid in this study). The correlation coefficients for the two grids were promoted to 0.94 and 0.95 by an increment amount of 0.05 from OL.

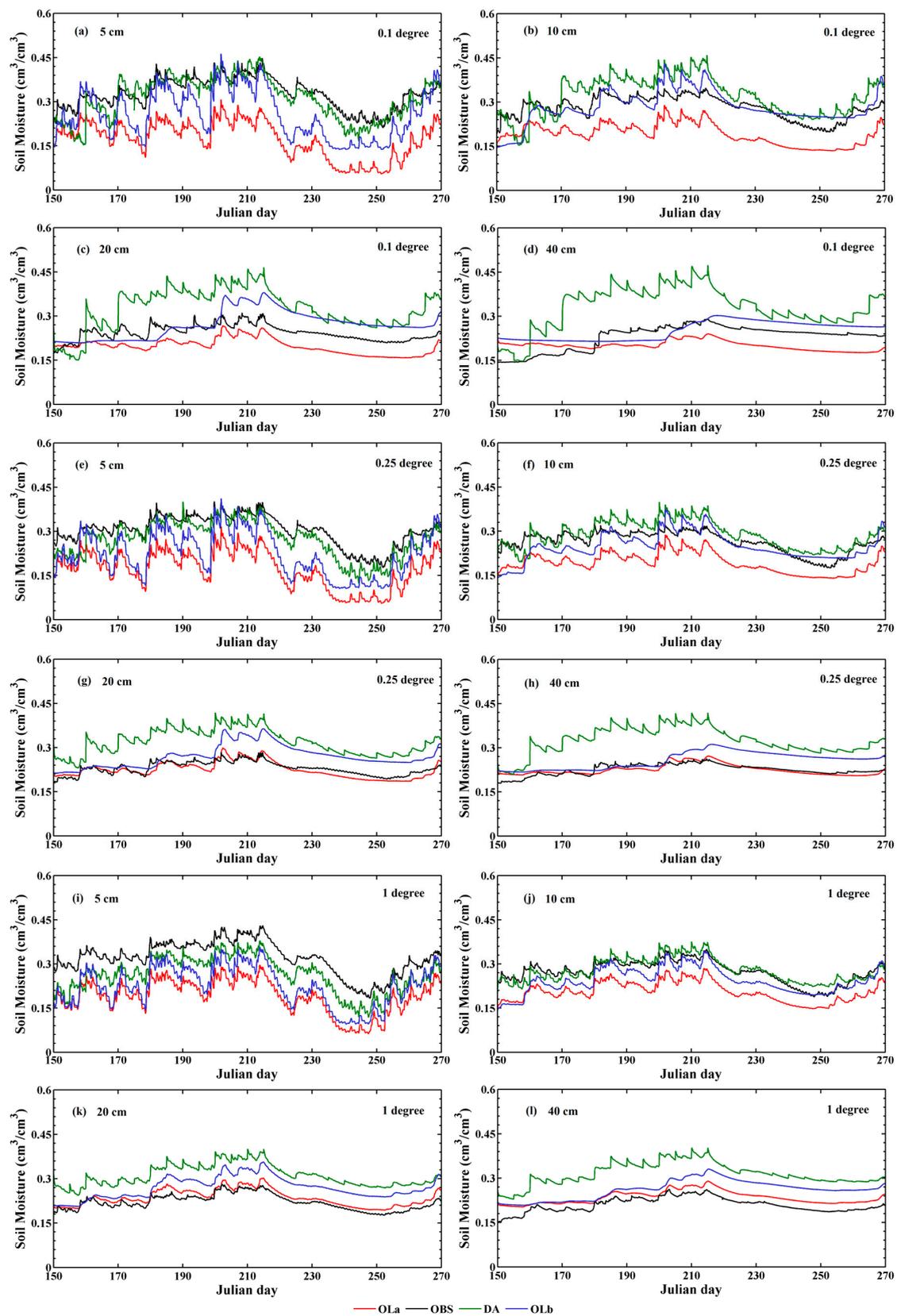


Figure 6. Evaluation of the soil moisture at four layers derived from simulation and assimilation with the measurements in the: small grid (a–d), medium grid (e–h), and large grid (i–l).

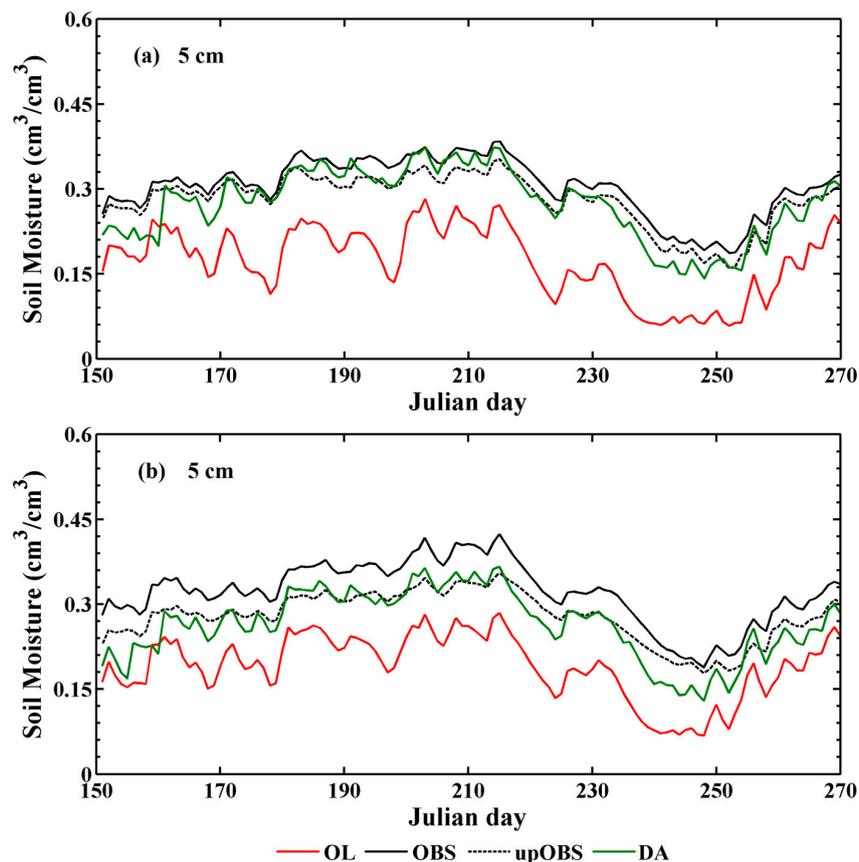


Figure 7. Evaluation of the daily soil moisture at surface derived from simulation and assimilation with the average measurements and upscaled measurements in the: medium grid (a) and large grid (b).

Figure 8 shows the scatter plots for soil temperature at 5 cm, 10 cm, 20 cm and 40 cm from the three grids which are present with spatially-averaged and hourly values. The layout of Figure 8 is same as that of Figure 6. OL and DA are the daily soil moisture from the model simulation and assimilation. The MBE and RMSE from OL and DA are distinguished by the same color as the dots and marked within the corresponding plot. Soil temperature from OL indicated good agreement between the simulations and the measurements, with the MBE for the small and medium grid reaching approximate 2 K (extremely small for the large grid) and the RMSE slightly larger than the MBE. The overestimated soil temperature expressed by the discrete dots from OL was evidently modified after assimilation and the soil temperature from DA concentrated along 1:1 line. The MBE declined by more than 1 K in the small and medium grid which suggested the evident correction of the bias. The RMSEs was reduced by 20% to 50% with the increase of the soil depth. However, less improvement was observed in the large grid because space for the advancement of soil temperature estimation was limited. Since the RMSE for OL was similar to the predetermined standard deviation error of observations, MODIS LST has almost no effect on the assimilation. Thus, the equivalent improvement of the soil temperature at the deeper layers is predominantly ascribed to the increase of soil moisture which in turn decreases the soil temperature because of the high heat capacity of water. Overall, the updated soil temperature from DA appeared to be underestimated as the LST observations acquired by satellite were generally lower than the in situ observations due to the field of view, atmospheric attenuation and imperfect inversion algorithm.

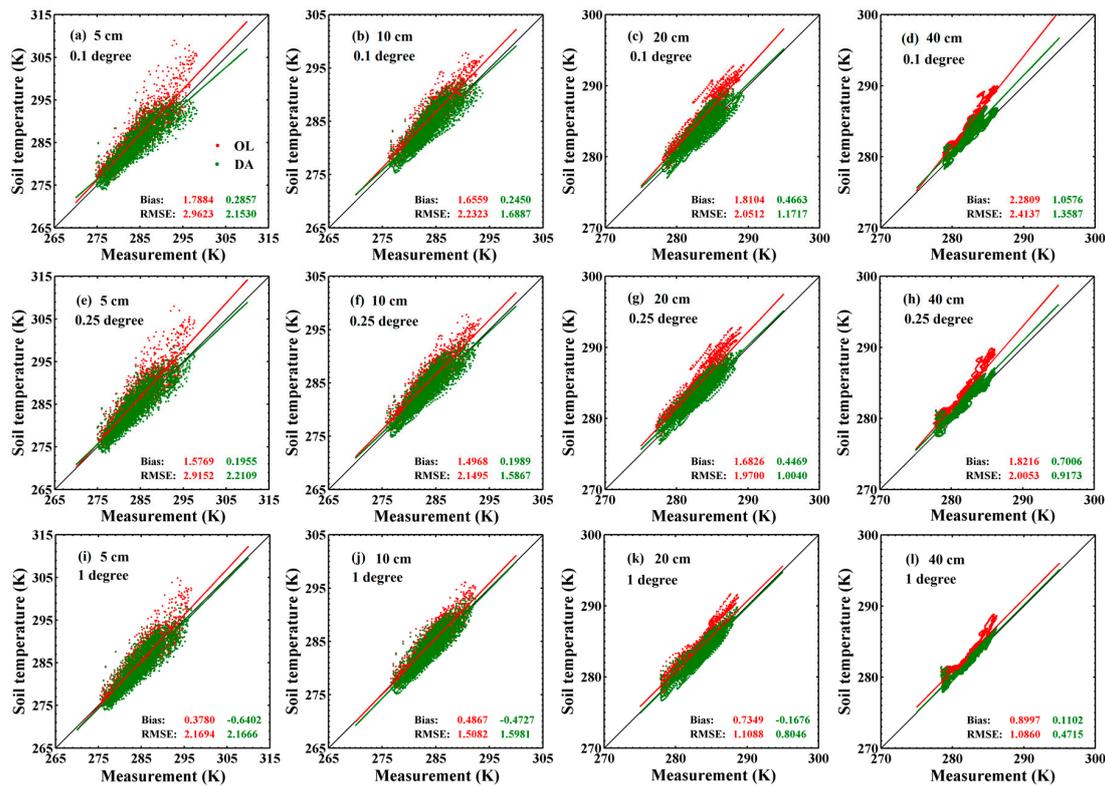


Figure 8. Evaluation of the soil temperature at four layers derived from simulation and assimilation with the measurements in the: small grid (a–d), medium grid (e–h), and large grid (i–l).

5. Conclusions

Microwave observations have significant advantage in the acquisition of soil moisture information at land surface, especially in terms of the low frequency. Considering that the coarser spatial resolution of AMSR-E observations/products cannot fulfill the needs of applications in most cases, this study investigates the possibility to downscale the soil moisture by simultaneously assimilating AMSR-E brightness temperatures and MODIS LST product. Here, MODIS LST product is employed to provide the information of surface temperature at the fine spatial scale of CoLM. Thus, the combined assimilation experiment can retrieve soil moisture at a high resolution despite the coarse-scale TB observations and resulting in a better accuracy than modeling. Moreover, dealing with the spatial scale issue during the data assimilation process in this study avoids a number of problems arising in the downscaling algorithms in which the coarse observations are usually disaggregated before assimilation [68]. The assimilation experiment produced quite encouraging improvement in the estimation of soil moisture at the surface as well as soil temperature profile, but the estimation of soil moisture at deeper layers deteriorated. A superior model parameterization scheme to achieve the effective delivery of the surface information to the deeper layers is expected in further studies.

To consider the uncertainty in parameters, five parameters were chosen to be calibrated using Ensemble Kalman Smoother. The soil moisture simulation with updated parameters reduced the original bias to the measurements and performed better than the results with default soil parameter values. The estimated parameter values were potentially different from the measured values due to discrepancy between the parameterizations of CoLM and natural processes. Finally, parameters estimated by assimilation were also likely to compensate for other uncertainties implied in initial conditions, forcing data etc.

An upscaling method was implemented to obtain the footprint-averaged soil moisture as the ground truth to resolve the representative problem brought about from point to area. Compared to the

ATI-based upscaled soil moisture, the improvement of soil moisture at surface become more significant which further proved the positive performance of TB assimilation in soil moisture estimation and the necessity of correctly handling the representative problem.

In summary, this study demonstrates the feasibility of downscaling the soil moisture through assimilating coarse-scale microwave TB observations and fine-scale LST observations and describes the potential to calibrate parameters and optimize state together in the same framework. The combination assimilation of multi-scale observations produces effective information and positively affects the states estimation in fine scale. Complicated mechanism in such data assimilation can be expected in the further work. In addition, it is worthy to take account of multiple sources of uncertainties such as parameters in state assimilation.

Acknowledgments: This work is supported by the National Science Foundation of China under Grants 91425303 and 41671375, and the Science and Technology Service Network Initiative Project of the Chinese Academy of Sciences.

Author Contributions: Weijing Chen designed the research, performed the research and wrote the manuscript. Huanfeng Shen, Chunlin Huang and Xin Li participated in the design and revision of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Seneviratne, S.I.; Corti, T.; Davin, E.L.; Hirschi, M.; Jaeger, E.B.; Lehner, I.; Orlowsky, B.; Teuling, A.J. Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Sci. Rev.* **2010**, *99*, 125–161. [[CrossRef](#)]
2. Hupet, F.; Vanclooster, M. Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field. *J. Hydrol.* **2002**, *261*, 86–101. [[CrossRef](#)]
3. Jung, M.; Reichstein, M.; Ciais, P.; Seneviratne, S.I.; Sheffield, J.; Goulden, M.L.; Bonan, G.; Cescatti, A.; Chen, J.; de Jeu, R.; et al. Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature* **2010**, *467*, 951–954. [[CrossRef](#)] [[PubMed](#)]
4. Robock, A.; Vinnikov, K.Y.; Srinivasan, G.; Entin, J.K.; Hollinger, S.E.; Speranskaya, N.A.; Liu, S.; Namkhai, A. The global soil moisture data bank. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 1281–1299. [[CrossRef](#)]
5. Entekhabi, D.; Njoku, E.G.; O’Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The soil moisture active passive (SMAP) mission. *Proc. IEEE*. **2010**, *98*, 704–716. [[CrossRef](#)]
6. Njoku, E.G.; Jackson, T.J.; Lakshmi, V.; Chan, T.K.; Nghiem, S.V. Soil moisture retrieval from AMSR-E. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 215–229. [[CrossRef](#)]
7. Rodell, M.; Houser, P.R.; Jambor, U.E.A.; Gottschalck, J.; Mitchell, K.; Meng, C.J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The global land data assimilation system. *Bull. Am. Meteorol. Soc.* **2004**, *85*, 381–394. [[CrossRef](#)]
8. Houser, P.R.; Shuttleworth, W.J.; Famiglietti, J.S.; Gupta, H.V.; Syed, K.H.; Goodrich, D.C. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resour. Res.* **1998**, *34*, 3405–3420. [[CrossRef](#)]
9. Liu, Y.; Gupta, H.V. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.* **2007**, *43*. [[CrossRef](#)]
10. Reichle, R.H. Data assimilation methods in the Earth sciences. *Adv. Water Resour.* **2008**, *31*, 1411–1418. [[CrossRef](#)]
11. Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **1994**, *99*, 10143–10162. [[CrossRef](#)]
12. Reichle, R.H.; McLaughlin, D.B.; Entekhabi, D. Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Weather Rev.* **2002**, *130*, 103–114. [[CrossRef](#)]
13. Houtekamer, P.L.; Mitchell, H.L. Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.* **1998**, *126*, 796–811. [[CrossRef](#)]

14. Clark, M.P.; Rupp, D.E.; Woods, R.A.; Zheng, X.; Ibbitt, R.P.; Slater, A.G.; Schmidt, J.; Uddstrom, M.J. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Adv. Water Resour.* **2008**, *31*, 1309–1324. [[CrossRef](#)]
15. Houtekamer, P.L.; Mitchell, H.L.; Pellerin, G.; Buehner, M.; Charron, M.; Spacek, L.; Hansen, B. Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Weather Rev.* **2005**, *133*, 604–620. [[CrossRef](#)]
16. Huang, C.; Li, X.; Lu, L.; Gu, J. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. *Remote Sens. Environ.* **2008**, *112*, 888–900. [[CrossRef](#)]
17. Huang, C.; Li, X.; Lu, L. Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. *Remote Sens. Environ.* **2008**, *112*, 1320–1336. [[CrossRef](#)]
18. Montzka, C.; Moradkhani, H.; Weihermüller, L.; Franssen, H.J.H.; Canty, M.; Vereecken, H. Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter. *J. Hydrol.* **2011**, *399*, 410–421. [[CrossRef](#)]
19. Noh, S.; Tachikawa, Y.; Shiiba, M.; Kim, S. Applying sequential Monte Carlo methods into a distributed hydrologic model: Lagged particle filtering approach with regularization. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3237–3251. [[CrossRef](#)]
20. Moradkhani, H.; DeChant, C.M.; Sorooshian, S. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. *Water Resour. Res.* **2012**, *48*. [[CrossRef](#)]
21. Yan, H.; DeChant, C.; Moradkhani, H. Improving soil moisture profile prediction with the particle filter-Markov chain Monte Carlo method. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6134–6147. [[CrossRef](#)]
22. Dumedah, G.; Coulibaly, P. Evaluating forecasting performance for data assimilation methods: The ensemble Kalman filter, the particle filter, and the evolutionary-based assimilation. *Adv. Water Resour.* **2013**, *60*, 47–63. [[CrossRef](#)]
23. Liu, P.W.; Bongiovanni, T.; Monsivais-Huertero, A.; Judge, J.; Steele-Dunne, S.; Bindlish, R.; Jackson, T.J. Assimilation of Active and Passive Microwave Observations for Improved Estimates of Soil Moisture and Crop Growth. *IEEE J. Sel. Top. Earth Obs. Remote Sens.* **2016**, *9*, 1357–1369. [[CrossRef](#)]
24. Crow, W.T.; Wood, E.F. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Adv. Water Resour.* **2003**, *26*, 137–149. [[CrossRef](#)]
25. Reichle, R.H.; Koster, R.D. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophys. Res. Lett.* **2005**, *32*. [[CrossRef](#)]
26. Pan, M.; Wood, E.F.; Wójcik, R.; McCabe, M.F. Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation. *Remote Sens. Environ.* **2008**, *112*, 1282–1294. [[CrossRef](#)]
27. Lievens, H.; Tomer, S.K.; Al Bitar, A.; de Lannoy, G.J.M.; Drusch, M.; Dumedah, G.; Hendricks Franssen, H.J.; Kree, Y.H.; Martens, B.; Pan, M.; et al. SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. *Remote Sens. Environ.* **2015**, *168*, 146–162. [[CrossRef](#)]
28. De Lannoy, G.J.M.; Reichle, R.H. Global Assimilation of Multiangle and Multipolarization SMOS Brightness Temperature Observations into the GEOS-5 Catchment Land Surface Model for Soil Moisture Estimation. *J. Hydrometeorol.* **2016**, *17*, 669–691. [[CrossRef](#)]
29. Draper, C.S.; Reichle, R.H.; de Lannoy, G.J.M.; Liu, Q. Assimilation of passive and active microwave soil moisture retrievals. *Geophys. Res. Lett.* **2012**, *39*, L04401. [[CrossRef](#)]
30. Lu, H.; Koike, T.; Yang, K.; Hu, Z.; Xu, X.; Rasmy, M.; Kuria, D.; Tamagawa, K. Improving land surface soil moisture and energy flux simulations over the Tibetan plateau by the assimilation of the microwave remote sensing data and the GCM output into a land surface model. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *17*, 43–54. [[CrossRef](#)]
31. Cho, E.; Choi, M.; Wagner, W. An assessment of remotely sensed surface and root zone soil moisture through active and passive sensors in northeast Asia. *Remote Sens. Environ.* **2015**, *160*, 166–179. [[CrossRef](#)]
32. De Lannoy, G.J.; Reichle, R.H.; Pauwels, V.R. Global calibration of the GEOS-5 L-band microwave radiative transfer model over nonfrozen land using SMOS observations. *J. Hydrometeorol.* **2013**, *14*, 765–785. [[CrossRef](#)]
33. Mo, T.; Choudhury, B.J.; Schmugge, T.J.; Wang, J.R.; Jackson, T.J. A model for microwave emission from vegetation-covered fields. *J. Geophys. Res.* **1982**, *87*, 11229–11237. [[CrossRef](#)]

34. Wigneron, J.P.; Kerr, Y.; Waldteufel, P.; Saleh, K.; Escorihuela, M.J.; Richaume, P.; Ferrazzoli, P.; de Rosnay, P.; Gurney, R.; Calvet, J.-C.; et al. L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields. *Remote Sens. Environ.* **2007**, *107*, 639–655. [[CrossRef](#)]
35. Han, X.; Hendricks Franssen, H.J.; Li, X.; Zhang, Y.; Montzka, C.; Vereecken, H. Joint assimilation of surface temperature and L-band microwave brightness temperature in land data assimilation. *Vadose Zone J.* **2013**, *12*. [[CrossRef](#)]
36. Madsen, H. Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. *Adv. Water Resour.* **2003**, *26*, 205–216. [[CrossRef](#)]
37. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91. [[CrossRef](#)]
38. Cao, W.; Bowden, W.B.; Davie, T.; Fenemor, A. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol. Proc.* **2006**, *20*, 1057–1073. [[CrossRef](#)]
39. Moradkhani, H.; Sorooshian, S.; Gupta, H.V.; Houser, P.R. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Adv. Water Resour.* **2005**, *28*, 135–147. [[CrossRef](#)]
40. Chen, F.; Crow, W.T.; Ryu, D. Dual forcing and state correction via soil moisture assimilation for improved rainfall–runoff modeling. *J. Hydrometeorol.* **2014**, *15*, 1832–1848. [[CrossRef](#)]
41. Vrugt, J.A.; Diks, C.G.; Gupta, H.V.; Bouten, W.; Verstraten, J.M. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resour. Res.* **2005**, *41*, W01017. [[CrossRef](#)]
42. Lu, H.; Yang, K.; Koike, T.; Zhao, L.; Qin, J. An improvement of the radiative transfer model component of a land data assimilation system and its validation on different land characteristics. *Remote Sens.* **2015**, *7*, 6358–6379. [[CrossRef](#)]
43. Zhao, L.; Yang, K.; Qin, J.; Chen, Y.; Tang, W.; Lu, H.; Yang, Z.L. The scale-dependence of SMOS soil moisture accuracy and its improvement through land data assimilation in the central Tibetan Plateau. *Remote Sens. Environ.* **2014**, *152*, 345–355. [[CrossRef](#)]
44. Chen, W.; Huang, C.; Shen, H.; Li, X. Comparison of ensemble-based state and parameter estimation methods for soil moisture data assimilation. *Adv. Water Resour.* **2015**, *86*, 425–438. [[CrossRef](#)]
45. Chu, N.; Huang, C.; Li, X.; Du, P. Simultaneous estimation of surface soil moisture and soil properties with a dual ensemble Kalman smoother. *Sci. China Earth.* **2015**, *58*, 2327–2339. [[CrossRef](#)]
46. Dai, Y.; Zeng, X.; Dickinson, R.E.; Baker, I. The common land model. *Bull. Am. Meteorol. Soc.* **2003**, *84*, 1013–1023. [[CrossRef](#)]
47. Lawrence, D.M.; Slater, A.G. Incorporating organic soil into a global climate model. *Clim. Dyn.* **2008**, *30*, 145–160. [[CrossRef](#)]
48. Jackson, T.J.; Schmugge, T.J. Vegetation effects on the microwave emission of soils. *Remote Sens. Environ.* **1991**, *36*, 203–212. [[CrossRef](#)]
49. Paloscia, S.; Pampaloni, P. Microwave polarization index for monitoring vegetation growth. *IEEE Trans. Geosci. Remote Sens.* **1988**, *26*, 617–621. [[CrossRef](#)]
50. Kun, Y.; Watanabe, T.; Koike, T.; Li, X.; Fujii, H.; Tamagawa, K.; Ma, Y.; Ishikama, H. Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget. *J. Meteorol. Soc. Jpn. Ser. II* **2007**, *85A*, 229–242.
51. Wang, J.R.; Choudhury, B.J. Remote sensing of soil moisture content, over bare field at 1.4 GHz frequency. *J. Geophys. Res.* **1981**, *86*, 5277–5282. [[CrossRef](#)]
52. Dobson, M.C.; Ulaby, F.T.; Hallikainen, M.T.; El-Rayes, M.A. Microwave dielectric behavior of wet soil-Part II: Dielectric mixing models. *IEEE Trans. Geosci. Remote Sens.* **1985**, *23*, 35–46. [[CrossRef](#)]
53. Ulaby, F.T.; Moore, R.K.; Fung, A.K. *Microwave Remote Sensing Active and Passive-Volume III: From Theory to Applications*; Artech House Inc.: Norwood, MA, USA, 1986.
54. Evensen, G.; van Leeuwen, P.J. An ensemble Kalman smoother for nonlinear dynamics. *Mon. Weather Rev.* **2000**, *128*, 1852–1867. [[CrossRef](#)]
55. Lei, F.; Huang, C.; Shen, H.; Li, X. Improving the estimation of hydrological states in the SWAT model via the ensemble Kalman smoother: Synthetic experiments for the Heihe River Basin in northwest China. *Adv. Water Resour.* **2014**, *67*, 32–45. [[CrossRef](#)]

56. Huang, C.; Chen, W.; Li, Y.; Shen, H.; Li, X. Assimilating multi-source data into land surface model to simultaneously improve estimations of soil moisture, soil temperature, and surface turbulent fluxes in irrigated fields. *Agric. For. Meteorol.* **2016**, *230*, 142–156. [[CrossRef](#)]
57. Whitaker, J.S.; Hamill, T.M. Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Weather Rev.* **2012**, *140*, 3078–3089. [[CrossRef](#)]
58. Wanders, N.; Bierkens, M.F.; de Jong, S.M.; de Roo, A.; Karssenberg, D. The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models. *Water Resour. Res.* **2014**, *50*, 6874–6891. [[CrossRef](#)]
59. Yang, K.; Qin, J.; Zhao, L.; Chen, Y.; Tang, W.; Han, M.; Zhu, L.; Chen, Z.; Lv, N.; Ding, B.; Wu, H.; Lin, C. A multiscale soil moisture and freeze-thaw monitoring network on the third pole. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1907–1916. [[CrossRef](#)]
60. Zhao, L.; Yang, K.; Qin, J.; Chen, Y.; Tang, W.; Montzka, C.; Wu, H.; Lin, C.; Han, M.; Vereecken, H. Spatiotemporal analysis of soil moisture observations within a Tibetan mesoscale area and its implication to regional soil moisture measurements. *J. Hydrol.* **2013**, *482*, 92–104. [[CrossRef](#)]
61. Chen, Y.; Yang, K.; He, J.; Qin, J.; Shi, J.; Du, J.; He, Q. Improving land surface temperature modeling for dry land of China. *J. Geophys. Res.* **2011**, *116*, D20104. [[CrossRef](#)]
62. He, J.; Yang, K. *China Meteorological Forcing Dataset*; Cold and Arid Regions Science Data Center at Lanzhou: Lanzhou, China, 2011.
63. Liston, G.E.; Elder, K. A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). *J. Hydrometeorol.* **2006**, *7*, 217–234. [[CrossRef](#)]
64. Qin, J.; Yang, K.; Lu, N.; Chen, Y.; Zhao, L.; Han, M. Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia. *Remote Sens. Environ.* **2013**, *138*, 1–9. [[CrossRef](#)]
65. Reichle, R.H.; Bosilovich, M.G.; Crow, W.T.; Koster, R.D.; Kumar, S.V.; Mahanama, S.P.; Zaitchik, B.F. Recent advances in land data assimilation at the NASA Global Modeling and Assimilation Office. In *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 407–428.
66. Yang, K.; Zhu, L.; Chen, Y.; Zhao, L.; Qin, J.; Lu, H.; Tang, W.; Han, M.; Ding, B.; Fang, N. Land surface model calibration through microwave data assimilation for improving soil moisture simulations. *J. Hydrol.* **2016**, *533*, 266–276. [[CrossRef](#)]
67. Vrugt, J.A.; ter Braak, C.J.F.; Diks, C.G.H.; Schoups, G. Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. *Adv. Water Resour.* **2013**, *51*, 457–478. [[CrossRef](#)]
68. De Lannoy, G.J.; Reichle, R.H.; Houser, P.R.; Arsenault, K.R.; Verhoest, N.E.; Pauwels, V.R. Satellite-scale snow water equivalent assimilation into a high-resolution land surface model. *J. Hydrometeorol.* **2010**, *11*, 352–369. [[CrossRef](#)]

