Performance of AMSR_E Soil Moisture Data Assimilation in CLM4.5 Model for Monitoring Hydrologic Fluxes at Global Scale

Di Liu^{a,b}, Ashok K. Mishra^{b*}

Abstract

In this study, we evaluated the performance of community land surface model (CLM4.5) to simulate the hydrologic fluxes, such as, soil moisture (SM), evapotranspiration (ET) and runoff with (without) remote sensing data assimilation. The Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR_E) daily SM (both ascending and descending) are incorporated into the CLM4.5 model using data assimilation (DA) technique. The GLDAS data is used to validate the AMSR_E SM data and evaluate the performance of CLM4.5 simulations. The AMSR_E SM data are rescaled to meet the resolution of CLM4.5 model. By assimilating the AMSR_E SM data into the CLM4.5 model can improve the SM simulations, especially over the climate transition zones in Africa, East Australia, South South America, Southeast Asia, and East North America in summer season. The Local Ensemble Kalman Filter (LEnKF) technique improves the performance of CLM4.5 model compared to the directly substituted method. The improvement in ET and surface runoff simulations from CLM4.5 model assimilated with AMSR_E SM data shares similar spatial patterns with SM.

Key Words: Data assimilation; AMSR_E soil moisture; CLM4.5; Local Ensemble Kalman Filter

1. Introduction

The soil moisture is vital in controlling the partitioning of rainfall into runoff and percolation as well as the separation of incident solar energy into sensible heat and latent heat across the land-atmosphere interface (Koster et al., 2006; Liu et al., 2014a; Liu et al., 2014b). It is also notable for its influence on the evolution of weather, the rate of evaporation from the surface, eco-geomorphologic processes of arid hillside ecosystems, and vegetation photosynthetic functions (Houser et al., 1998; Katra et al., 2007; Mostovoy and Anantharaj, 2008; Entekhabi et al., 2010). The long memory of soil moisture impacts the exchanges of water, energy, and carbon between land surface and atmosphere, which are the necessary conditions for numerical weather prediction or hydrology cycles (e.g., Segele et al., 2013; Fu et al., 2014; Liu et al., 2014a).

Over the past decades, many studies have explored the mechanisms of soil moisture in the land-atmosphere water and energy cycles (e.g., Koster et al., 2006; Dirmeyer, 2011; Wei and Dirmeyer, 2012; Liu et al., 2014b). However, due to the complicated topography and limited in situ soil moisture observations, there is a challenge to quantify (measure) fine-scale observed soil moisture at global scale, especially in the root zone. To enhance the understanding of land atmosphere interactions, researchers have coupled land surface models (LSMs) and numerical weather prediction models over the past few years (e.g., Zhang et al., 2011; Mei et al., 2013; Liu et al., 2014b). Yet a lack of accurate information about spatial and temporal variation of soil properties, vegetation characteristics, moisture states, and many atmospheric forcing fields at high resolutions often hinders executing objective diagnoses on the accuracy of these coupling efforts and even results in employing erroneous initial states for simulations. Though LSMs can generate high resolution soil moisture at various depths, the simulated soil moisture relatively at higher spatial resolutions are often biased due to the uncertainties of their parameterization, numerical integration, and various physical processes represented (e.g., Han et al., 2014; Yin et al., 2014). To resolve these problems, new strategies are applied, such as the application of data assimilation techniques (e.g., Moradkhani

et al., 2005; Liu and Gupta, 2007; Kumar et al., 2008; Nie et al., 2011; Moradkhani et al., 2012; Yin et al., 2015), and the application of remote sensing data (e.g., Chen et al., 2013; Lievens et al., 2015; Lannoy and Reichle, 2016; Xiao et al., 2016).

Microwave remote sensing is an important tool for obtaining surface soil moisture data at local, regional and global scales (Kerr et al., 2001; Bartalis et al., 2007). Currently, multiple soil moisture products have been generated from data acquired by the Advanced Microwave Scanning Radiometer (AMSR-E) on board the Aqua satellite (Kawanishi et al., 2003). Application of near-surface soil moisture fields derived from AMSR-E have shown promising results over Europe (Rüdiger et al., 2009), the United States (McCabe et al., 2005; Crow and Zhan, 2007), Australia (Draper et al., 2009; Liu et al., 2009) and East Asia (Njoku et al., 2005).

To facilitate studies at a regional scale, some efforts have been made to simulate soil moisture using land surface models forced with precipitation and other atmospheric data. The Global Land Data Assimilation System (GLDAS) provides simulated soil moisture data at 1.0 degree and 0.25 degree resolution (Rodell et al., 2004b). The simulated soil moisture data provide spatially and temporally consistent time series, but their accuracy is hindered by model deficiencies, and uncertainties in both model parameters and atmospheric forcing variables (Draper et al., 2009; Chen et al., 2013). Recently, many studies evaluated the assimilation of satellite observations into land surface models to produce optimal soil moisture estimates (e.g., Reichle and Koster, 2005; Han et al., 2014; Lievens et al., 2015; Lannoy and Reichle, 2016). For example, (Han et al., 2014) evaluated the joint state and parameter estimation method for the coupled Community Land Model version 4.5 (CLM4.5) (Oleson, 2013) and Community Microwave Emission Model (CMEM) (de Rosnay et al., 2009) through assimilation of synthetic microwave brightness temperature data.

According to previous studies, CLM model is mainly adopted in the land-atmosphere interactions or the water balance studies at continental to global scales, however its performance on hydrological simulations (e.g., runoff) varies based on the regions. Currently there are several limitations in CLM4.0 (Li et al., 2011), for example, unrealistic behavior of subsurface runoff due to the exponential

form of the surface runoff parameterization in CLM4.0 as well as significant effects of spatial heterogeneity on runoff even at small watershed. CLM4.0 tends to produce unrealistically large temporal variations of runoff for applications at mountainous catchment. Additional limitations can be found in Li et al., (2011).

This study aims to investigate the efficiency of AMSR_E daily soil moisture (AMSR_E SM) assimilation in CLM4.5 Model for Monitoring Hydrologic Fluxes at Global Scale. The Qian et al. atmospheric (Qian ATM) data (Qian et al., 2006) are used as the forcing data. The GLDAS data from Noah model version 2 is used to evaluate the CLM4.5 simulations assimilated with AMSR_E SM data. The local Ensemble Kalman Filter (LEnKF) method is applied to assimilate AMSR_E SM data in the CLM4.5 modeling framework.

The main objectives of this study are: (1) to validate the efficiency of CLM4.5 simulations for soil moisture; (2) to evaluate the feasibility of AMSR_E SM to improve CLM4.5 models for soil moisture estimation with or without DA technique; and (3) to investigate the performance of AMSR_E SM assimilation in CLM4.5 model to simulate hydrometeorological variables (e.g., evapotranspiration (ET), runoff). Due to radio-frequency interference, dense vegetation, and frozen soil, the remote sensing products are not spatially or temporally complete. These deficiencies limit the application of soil moisture in land surface process simulation, climatic modeling, and global change research. In order to diminish such influence, the AMSR-E SM data are preprocessed to match the resolution of CLM4.5 in this study.

The remainder of this paper is organized as follows: the CLM4.5 model, GLDAS data and AMSR-E SM data are introduced in Section 2; the Methodologies, including the experimental design and evaluation criteria are described in Section 3; the results, including model validation and analysis of simulated soil moisture, ET and runoff are documented in Section 4; while the conclusions are summarized in Section 5.

2. Model and Data Description

2.1 CLM4.5 Description

The CLM4.5 coupled in the community earth system model (CESM) version

1.2 is adopted in this paper. The CESM is a coupled climate model for simulating Earth's climate system, composed of separate models simultaneously simulating the Earth's atmosphere, ocean, land, land-ice, and sea-ice, plus one central coupler component. CESM is suitable to conduct fundamental research into the Earth's past, present, and future climate states.

The CLM4.5 model is well-developed based on the water and energy balance simulation. The model parameterizes the interception, throughfall, canopy drip, snow accumulation and melt, water transfer between snow layers, infiltration, evaporation, surface runoff, sub-surface drainage, redistribution within the soil column, and groundwater discharge and recharge to simulate changes in canopy water, surface water, snow water, soil water, and soil ice, and water in the unconfined aquifer.

In CLM4.5, the soil profile is divided into 15 layers with different depth from 0.007100635 m, 0.027925 m, 0.06225858 m, 0.1188651 m, 0.2121934 m, 0.3660658 m, 0.6197585 m, 1.038027 m, 1.727635 m, 2.864607 m, 4.739157 m, 7.829766 m, 12.92532 m, 21.32647 m to 35.17762 m in each active grid cell and the soil moisture is only simulated for the first 10 layers. The modified Richard's equation is used to predict the one-dimensional multilayer vertical soil water flow and energy transport. The Monin-Obukhov similarity theory is used to derive the land surface fluxes. The spatial land surface heterogeneity is considered in CLM4.5 and soil moisture, soil temperature, infiltration, evapotranspiration, sensible heat flux, latent heat flux and soil heat flux are simulated (Oleson, 2013). CLM4.5 uses the pedotransfer functions of sand and clay (Cosby et al., 1984) and organic properties of the soil (Lawrence and Slater, 2007) to derive soil hydraulic and thermal parameters. Additional information on modeling components can be found in the CLM45_Tech_Note (http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM45_Tech_Note.pdf).

In this paper, the CLM4.5 model is coupled with Variable Infiltration Capacity (VIC) model (Liang et al., 1994) by incorporating following components: DATM; CLM; RTM; SICE; SOCN; SGLC; SWAV. The Satellite Phenology is chosen based on the industrial scenario driven by Qian ATM forcing data during 1972-2004. The model resolution is set with f19_g16 globally with 288 (longitude) *192 (latitude)

total grid cells. To get a stable model, the control run is simulated with the Qian ATM forcing and the default setting over the period of 1972-2004 for three cycles (total 99 years) with the first two cycles simulations (initial 66 years) considered as spin up. The model is validated using the last 33 years simulation based on the 3rd cycle.

2.2 GLDAS Data Description

GLDAS product combines satellite data and Land Surface Models (LSM) to estimate global distributions of land surface states such as soil water, runoff and ET (Rodell et al., 2004b; Hogue et al., 2005). GLDAS product has been widely used in many studies, including the initialization of climate/hydrology modeling experiments (e.g., Koster et al., 2004b; Seyyedi et al., 2015; Moiwo and Tao, 2015), land-atmosphere interactions (e.g., Zhang et al., 2008; Liu et al. 2014b; Liu et al. 2015), and the evaluation and interpretation of new satellite data records or terrestrial water storage (e.g., Syed et al., 2008; Anderson et al. 2005; Chen et al. 2013; Awange et al. 2014; Yang et al. 2013; Cheng et al. 2015). Though GLDAS is simulated with advanced simulation techniques, it is not bias free and contains some problems. For example, GLDAS amplified error uncertainties in hydrological modeling due to coarse resolution (Seyyedi et al. 2015); GLDAS product is neither calibrated against field measurements and/or observations globally, nor can account for anthropogenic contributions or simulated dynamic lakes or reservoirs (Ahmed et al. 2016). Additionally, GLDAS product does not simulate groundwater and river channel storages (Rodell et al. 2004b). Moreover, irrigation is not included in GLDAS land surface state simulations due to limited long-term irrigation measurements (Rodell et al. 2004b). The reasons for selecting GLDAS are its global coverage, relatively high temporal resolution, and long data record. Moreover, GLDAS version 2 constitutes significant improvements in terms of source code and forcing data, and is therefore reliable for hydrological studies (Rodell et al., 2004a, 2004b).

Due to the lack of more suitable globally spatial and temporal resolution of

observed data, the $0.25^{\circ} \times 0.25^{\circ}$ 3-hourly GLDAS data (version 2) from Noah model (<http://disc.sci.gsfc.nasa.gov/services/grads-gds/gldas>) are adopted in this study to evaluate the performance of AMSR_E SM assimilation in CLM4.5 modeling framework to simulate hydrometeorological variables (e.g., ET, runoff).

2.3 AMSR-E Soil Moisture Data Description

The Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR_E) sensor on NASA's Aqua Satellite is launched on 2 May 2002. Right now, the AMSR-E instrument may provide measurements of terrestrial, oceanic, and atmospheric parameters for the investigation of global water and energy cycles with the temporal coverage from 18 June 2002 through 04 October 2011.

The gridded Level-3 land surface product can provide daily measurements of surface soil moisture and vegetation/roughness water content interpretive information, as well as brightness temperatures and quality control variables. The soil microwave emission originates from the top surface layer with the emitting depth governed by the dielectric characteristics of the near-surface moisture profile and smaller for higher frequencies and for wetter soils Njoku et al., (2003). The emitting depth is about 1 cm or less at C-band while about 2-5 cm at L-band (Ulaby et al., 1986). The soil moisture and temperature profiles are more uniform at night than in the early afternoon. Therefore, the descending AMSR_E soil moisture (1:30 A.M.) retrievals are expected to have less modeling error and be more representative of the deeper-layer soil moisture than the ascending observations .

In this study, the CLM4.5 model is driven with Qian ATM forcing data during 1972-2004, therefore, to match the same time period, the 25 km x 25 km AMSR_E soil moisture data at daily resolution during 2003-2004 years (both ascending and descending) (http://nsidc.org/data/docs/daac/ae_land3_l3_soil_moisture.gd.html) are adopted to update the CLM4.5 model simulations. The ascending and descending soil moisture will be analyzed to assess their relative accuracies with respect to the expected surface layer diurnal moisture variations derived from GLDAS and CLM4.5 control run.

3. Methodology

To improve the simulation of soil moisture, several studies focused on updating model parameters like the saturated hydraulic conductivity or the Mualem-van Genuchten parameters of the governing Richards equation (Montzka et al., 2011; Nie et al., 2011; Han et al., 2013). However land surface or hydrologic models also use soil properties (i.e., soil texture and soil organic matter density) combined with model internal pedotransfer functions to derive soil hydraulic and thermal parameter values (Han et al., 2014). In this study, the CLM4.5 model is used to simulate the time series of soil moisture and other hydrometeorological variables (i.e., runoff, evapotranspiration (ET)). The AMSR_E soil moisture data is assimilated into the CLM4.5 model with DA technique to improve the simulation results.

3.1 Experimental Design

The CLM4.5 model with satellite phenology scenario is selected. The control run is simulated with the Qian ATM forcing and the default setting over the period of 1972-2004 based on three cycles (total = $3 \times 33 = 99$ years) with the first two cycles simulations (initial 66 years) used as spin up. The simulation in the last two years generated with the Qian ATM forcing data during 2003-2004 from the 3rd cycle in the open loop is used as the control run (EXP1) to evaluate the performance of CLM4.5 modeling without AMSR_E data assimilation. The simulations of CLM4.5 driven with CRUNCEP ATM forcing data during 2003-2004 (EXP2) is compared with EXP1.

As described, the soil moisture is estimated in 10 soil layers in CLM4.5. However, the AMSR_E SM data refers to the volumetric soil moisture data within 1 cm in the top soil layer in winter season while it is 2.5 cm in spring and summer season when the weather is warm enough. Therefore, to make it simple, only the 1st layer soil moisture is substituted with AMSR_E SM directly or assimilated with DA technique at selectd time steps, while the deeper layers evolve freely with the soil moisture movement defined in CLM4.5.

The daily AMSR-E SM data during 2003-2004 are assimilated to update the 1st layer soil moisture in the CLM4.5 at the first time step of each day during 2003-2004

with DA technique. Considering the number of grid cells and computational efficiency, the number of ensemble member is set to 15. As stated in previous literatures (e.g., Kumar et al., 2008; Pan and Wood, 2010; De Lannoy et al., 2012; Yin et al., 2015), the performance of EnKF relies on the ensemble size. For example, (Kumar et al., 2008; Yin et al., 2015) indicated that when the ensemble size is close to 12, it may lead to efficient DA updating process, while (Pan and Wood, 2010; De Lannoy et al., 2012) suggested 20 ensemble members. The observation error and model output error is assumed to be random Gaussian distribution. The liquid and ice soil moisture are separately updated with the volumetric soil moisture in the same layer in the same grid cell with the constraint of soil temperature within that layer at that time step. If the soil temperature is above zero degrees Celsius, the liquid soil moisture is updated; otherwise the ice soil moisture is updated.

Several experiments are conducted to evaluate and compare the state and parameter estimation strategies with the AMSR-E data and DA technique. All the experiments are driven with the same ATM forcing (Qian forcing during 2003-2004). The details of each experiment are given as follows:

1. **Open Loop (OL):** CLM4.5 model simulated with Qian ATM forcing data during 1972-2004 time periods (3*33 years), with the initial 2 cycles (2*33 = 66 years) used as spin up. This can be seen as the control run.
2. **EXP1:** Abstracted from the last two years (2003-2004) in the 3rd cycle in the open loop with Qian ATM forcing data without updating soil moisture with AMSR_E SM data.
3. **EXP2:** Similar to EXP1, but CRUNCEP ATM forcing data are used without updating soil moisture with AMSR_E SM during 2003-2004.
4. **ASC:** substitute the volumetric soil moisture at the 1st layer (within 0.71 cm) with daily AMSR_E ascending soil moisture data at the first time step of each day during 2003-2004.
5. **DSC:** similar to ASC, but substitute the volumetric soil moisture at the 1st layer (within 0.71 cm) with daily AMSR_E descending soil moisture data at the first time step of each day during 2003-2004.

6. **DDA**: assimilate the volumetric soil moisture at the 1st layer with daily AMSR_E descending soil moisture data and LEnKF technique at the first time step of each day during 2003-2004.

The flowchart for the CLM4.5 modeling and DA updating process is given in [Figure 1](#). The CLM4.5 model is conducted using different ATM forcing data in the open loop to select the best initial model parameters. The AMSR-E SM data used to update the CLM4.5 model by substituting the simulated soil moisture in the 1st layer directly or assimilating with DA technique.

3.2 Evaluation Criteria

The GLDAS data was used to validate the CLM4.5 modeling framework in the open loop as well as using assimilated AMSR_E SM data. The statistical criteria (e.g., mean and standard deviation) are used to assess the CLM4.5 model performance at different conditions. In addition, the root mean square error (RMSE), mean absolute error (MAE), correlation coefficient (R), and the assimilation efficiency (AE) index, are also adopted to evaluate the efficiency of assimilated AMSR_E SM data in the improvement of CLM4.5 model. The calculation of RMSE, MAE, and R can be found in (Yu et al., 2012). The AE is calculated in Equation (1) as in (Yin et al., 2014).

$$AE = \frac{RMSE_{UP}}{RMSE_{OL}} - 1 \quad (1)$$

Here, the $RMSE_{UP}$ is the RMSE between GLDAS data and the simulations from CLM4.5 model assimilated with AMSR_E SM data (i.e., ASC, DSC, DDA); while the $RMSE_{OL}$ is the RMSE between GLDAS data and the simulations from CLM4.5 control run (EXP1). As stated in (Yin et al., 2014), the negative AE index indicates the improvements of CLM4.5 model assimilated with the AMSR_E SM data while the positive AE index signifies the degradations in the assimilation. The results are illustrated in [Section 4.2](#).

4. Results Analysis

4.1 Model Validation

The initial ATM forcing data (e.g., precipitation, 2 m air temperature (T2m)) from Qian and CRUNCEP data are compared with CRU and GLDAS data sets during 1991-2004 years (Figure S1-4). By comparison, the mean and standard deviation of precipitation and T2m from GLDAS data agrees well with the CRU data not only in the spatial patterns but also in terms of magnitude at both annual and seasonal time scales. The grid-based mean precipitation and T2m of Qian forcing data resembles similar spatial patterns and shares almost the same amplitude with the CRUNCEP data in both the annual and seasonal time scale with the only obvious difference exists in winter season (DJF) when Qian forcing data has slightly higher precipitation magnitude (about 20 mm/mon) over the north Africa and southeast Asia.

Compared to CRU and GLDAS data, both Qian and CRUNCEP overestimate the precipitation over the Southeast Asia, North Europe, Australia, North Africa and East North America in each season. The standard deviation of precipitation from Qian data has higher magnitude over these regions while the other three data sets share similar amplitude, implying the significant inter-annual and intra-seasonal trends of Qian data. The standard deviation of T2m differs little among different data sets either at the annual or the seasonal time scale.

Considering the interannual or intraseasonal influence on the quality of data, a long period of AMSR_E SM data are validated with GLDAS data. Here, the monthly, seasonal and annual mean and standard deviation of AMSR_E daily SM (mm^3/mm^3) data (both ascending and descending) during 2003-2007 are compared with GLDAS data and the simulations from CLM45 EXP1 and EXP2 (Figure S5-8). As shown in Figure S5, the mean soil moisture simulated by CLM4.5 at the 1st soil layer driven with either Qian (EXP1) or CRUNCEP (EXP2) data have higher magnitude compared to GLDAS or AMSR_E SM data (both ascending or descending) in each month, especially over Australia, North Africa, South South America and East North America. This may be caused by the ATM forcing data which shows obvious deviation of precipitation with respect to GLDAS data over these regions. However, the difference

between EXP1 and EXP2 are small and can be neglected over most of the grid cells in each month. Unlike the CLM4.5 simulations, the AMSR_E ascending or descending SM shares similar spatial patterns with GLDAS but have lower magnitude especially over the climate transition zones in Southeast Asia, Northeast Asia, Mid Africa and East North America. As indicated previously (Koster et al. 2004a; Koster et al. 2006; Seneviratne et al. 2006; Zhang et al. 2008; Liu et al. 2014b; Liu et al. 2015), the land-atmosphere interaction is stronger over the climate transition zones, such as the central Great Plains of North America, northern India, southeast China, western China, the Sahel, equatorial Africa, southern Europe (Koster et al. 2004a; Koster et al. 2006; Zhang et al. 2008). Therefore, the improvement of SM over these regions may greatly improve the CLM4.5 modeling of hydrometeorological variables. The difference between ascending and descending SM is small and can be ignored. Consistently, the standard deviation of CLM4.5 simulations share similar spatial patterns with GLDAS except Australia, Africa and South America at each month; while the AMSR_E SM presents lower standard deviation especially from June to October (Figure S6). The annual or seasonal mean and standard deviation of SM among each data set shares similar spatial patterns with the monthly results (Figure S7-S8).

GLDAS data is derived using DA technique by merging multiple remote sensing data or observed data. This simulated SM data provide spatially and temporally consistent time series, but their accuracy is hindered by model deficiencies, and uncertainties in both model parameters and atmospheric forcing variables (Draper et al., 2009; Chen et al., 2013). The consistent spatial pattern of AMSR_E SM with GLDAS but with lower magnitude provides opportunity to reduce the uncertainties of CLM4.5 modeling with DA methods.

By comparison, the difference between AMSR_E ascending and descending SM is small for both the spatial patterns and magnitude. The simulations from EXP1 and EXP2 also differ little. Therefore, the Qian data are selected as the CLM4.5 ATM forcing data in the CLM4.5 sensitivity experiments with AMSR_E SM data assimilation.

4.2 Evaluation of Soil Moisture

4.2.1 Annual and Seasonal Mean Comparison

The annual and seasonal mean of soil water in the upper 10 cm (SWL1) (Unit: mm^3/mm^3) from the CLM4.5 experiments (OL, ASC, DSC, DDA) are compared in [Figure 2](#). As shown in [Figure 2](#), the simulations from CLM4.5 open loop has higher soil water content over the South America, Africa, Southeast America and Northwest Europe compared to the remaining grid cells. The CLM4.5 model simulations with AMSR_E SM data assimilation at each condition (ASC, DSC, DDA) underestimated the SWL1 over most of these areas while overestimated over the remaining areas in both the annual and seasonal time scale. The difference between CLM4.5 assimilated with the Ascending (ASC) or Descending (DSC) AMSR_E SM data is small and can be neglected over most of the grid cells with an exception over the Northeast Asia. The difference between simulations with DSC and DDA is stronger compared to the difference between ASC and DSC without LEnKF technique. The major difference between DDA and DSC exists over the North America, South America, Southeast-Northeast Asia, North Europe with the magnitude greater than $0.045 (\text{mm}^3/\text{mm}^3)$ in the annual or each seasonal time scale. The comparison in each month reflects that most of these significant differences between DDA and DSC come from March, April, May, October, November and December ([Figure S9 in the supplementary](#)).

Compared to the GLDAS data ([Figure 3](#)), the CLM4.5 simulations from open loop overestimate the SWL1 over the South America, East North America, Africa, Northwest Europe, East Australia and part of North China (e.g., located in the Yellow River Basin) in both the annual and seasonal time scales. The assimilation of AMSR_E SM (both ascending and descending) minimizes the bias with GLDAS data in each condition. Compared to the open loop, the simulations from either ASC or DSC or DDA minimize the deviation with GLDAS from above $0.15 (\text{mm}^3/\text{mm}^3)$ to the magnitude smaller than $0.09 (\text{mm}^3/\text{mm}^3)$ or even lower than $0.03 (\text{mm}^3/\text{mm}^3)$ over most of the grid cells in both the annual and seasonal time scale. Such findings indicate the efficiency of AMSR_E SM assimilation in the improvement of CLM4.5 model. In addition, the deviation between DDA and GLDAS is smaller (within 0.03

mm³/mm³) compared to the deviation between DSC and GLDAS over the Southeast Asia, North Europe, South Africa and South America, demonstrating the improvement of LEnKF technique in the assimilation process. The improvement over Southeast Asia (Indochina) and Mid-Africa is comparatively low with AMSR_E SM assimilation though it does have little improvement compared to the open loop (deviation within 0.15 mm³/mm³).

At the seasonal time scale, the SM simulations from CLM4.5 assimilated with AMSR_E SM data (both ascending and descending) in summer and autumn seasons are slightly better compared to the spring and winter seasons due to smaller bias (within 0.03 mm³/mm³) with respect to GLDAS data. In spring and winter, the bias is greater than 0.03 mm³/mm³ over most grids located in Africa (North and South), North America (most part), South America (South part), Australia and Mid-high Asia. The performance of data assimilation differs in space and time, which may be due to the gaps in remote sensing data. As stated, the satellite soil moisture products are not spatially or temporally complete, primarily due to track changes, radio-frequency interference, dense vegetation, and frozen soil. Such deficiencies would limit the application of soil moisture in land surface process simulation. Overall, the assimilated AMSR_E SM data (both ascending and descending) improve the simulations of SWL1 in CLM4.5 open loop.

In the CLM4.5 model, soil moisture propagates between the surface and deep layers according to the modeled soil moisture dynamics. Therefore, the soil moisture information may be transferred from surface to root zone. Consistently, the deviation of soil water contents in the deep soil layers (SWL2 within 10-40 cm depth, SWL3 within 40-100 cm depth and SWL4 within 100-200 cm depth) from CLM4.5 simulations shares similar spatial patterns with the surface soil layer within 10 cm depth (SWL1) with respect to GLDAS data ([Figures S10-S12](#)). The assimilation of AMSR_E SM (both ascending and descending) with or without LEnKF technique greatly improved the root zone soil water simulations over the typical areas (including South America, East North America, Africa, Northwest Europe, East Australia and part of Northeast Asia along the Yellow River Basin). The DDA

improves slightly over the climate transition zones in Southeast Asia, part of South America and South Africa. The improvement from DDA may be due to the reduction of both AMSR_E SM uncertainties and CLM4.5 modeling uncertainties with DA technique over these area.

4.2.2 Evaluation based on goodness of fits

The goodness of fits, including MAE, RMSE, and R, between simulated SM using CLM4.5 model at different condition and GLDAS for the surface layer (SWL1) is provided in [Figure 4](#). By comparison, the simulation based on CLM4.5 open loop (OL) has higher MAE and RMSE values than ASC, DSC and DDA with AMSR_E SM assimilation, especially over the East North America, Southeast South America, Africa (especially North part), Northeast Europe, North India and part of Northeast Asia above the Yellow river basin and East Australia with the magnitude above 0.24 mm³/mm³. Additionally, the correlation coefficient (R) over these areas shows lower positive or even negative signals based on the open loop scenario. Such information indicates higher bias associated with CLM4.5 simulations in the open loop scenario with respect to GLDAS data. This indicates poor performance of CLM4.5 over these areas. However, the updated CLM4.5 simulations with AMSR_E SM data with or without LEnKF technique compares well with the GLDAS data based on the decreased MAE and/or RMSE values combined with increase in positive R values over these areas.

The spatial patterns of the MAE, RMSE and R for the soil water in the deep layers (within 10-40 cm (SWL2), 40-100 cm (SWL3) and 100-200 cm (SWL4)) agree well with the surface layer (SWL1). Additionally, they also resemble similar magnitude over most grid cells from the surface to root zones ([Figures S13-S15](#)). By comparison, the simulations from CLM4.5 open loop seem to deviate less from surface to root zone. However, the root zone (SWL4) with the updated CLM4.5 simulations with AMSR_E SM data (ASC, DSC, DDA) leads to higher R values over East North America, South America, North and South part of Africa and Southeast Australia. In the CLM4.5, the soil water is predicted from a multi-model, in which the vertical soil moisture transport is governed by infiltration, surface and

sub-surface runoff, gradient diffusion, gravity, canopy transpiration through root extraction, and interactions with groundwater followed the methods provided in Z.-L. Yang (1998, unpublished manuscript) and modified by Zeng and Decker (2009) (described in Section 7.4 in CLM45_Tech_Note http://www.cesm.ucar.edu/models/cesm1.2/clm/CLM45_Tech_Note.pdf). Therefore, the simulations of soil moisture in the deeper layer may not only influenced by the surface layer, but also by the land use pattern or other terrestrial factors.

To better illustrate the efficiency of CLM4.5 model with AMSR_E SM data assimilation, the AE index is applied to document the improvement with AMSR_E SM data. The results for soil water in different layers from surface to root zone are plotted in Figure 5. As indicated in previous section, the negative AE signals indicate improvements while the positive AE signals presents degradations in the assimilation performance. As shown, the AE is dominated with negative signal over most of the grid cells from each CLM4.5 simulations assimilated with AMSR_E SM data (ASC, DSC and DDA) in each soil layers from surface to root zone (L1, L2, L3 and L4). The spatial patterns from different CLM4.5 experiments or different soil layers in the same experiment are almost consistent with each other. Consistently, the lower negative magnitude of AE (higher absolute negative values) mainly located over the East North America, Southeast South America, North and South Africa, Northeast Europe, North India and part of Northeast Asia above the Yellow river basin, and East Australia.

The positive AE signals were observed in some parts of the Southeast and Northeast Asia (including India, Indochina, Southeast China, and Mongolia) and part of west North America, reflecting the degradation of AMSR_E SM assimilation in CLM4.5 model. This may be due to the uncertainties of assimilated AMSR_E SM over these areas. The LEnKF technique seems to improve less over these areas. For example, the positive magnitude of AE increased over the northeast Asia in the first soil layer from DDA compared to ASC and DSC (Figure 5).

4.3 Evaluation of ET and Runoff

The increase in difference of ET between GLDAS and CLM4.5 open loop and CLM4.5 assimilated with AMSR_E SM data (ASC, DSC, DDA) are located over Africa, Australia, part of Mid-Asia along the Yellow River basin and North India, and part of South America and South North America with the absolute deviation greater than 1.5 mm/day or even 2.5 mm/day (Figure 6). The CLM4.5 assimilated with AMSR_E SM data greatly reduces the bias with respect to GLDAS over these areas. By comparison, the simulation between ASC and DSC differs little and can be ignored while the DDA is a little better than DSC and ASC for some of the regions (WHERE?) based on the decrease in deviation compared to GLDAS. Consistently, the assimilated AMSR_E SM data also improves the CLM4.5 runoff simulations over most part of the areas with respect to GLDAS data in both the annual and seasonal time scales (Figure 7). The absolute deviation of runoff is reduced from more than 0.15 mm/day in the CLM4.5 open loop to lower than 0.03 mm/day based on the ASC, DSA or DDA.

By comparison, the distribution of ET or runoff deviation between the CLM4.5 sensitivity experiments with open loop or with GLDAS share similar spatial patterns with that of soil water content. The significant improvements are mainly located over the areas with higher soil water content, indicating the dominant role of soil moisture in the local hydrological processes.

4.4 Regional SM, ET and Runoff Analysis

To further analyze the improvement of CLM4.5 model assimilated with AMSR_E SM data, 5 typical regions are selected: East North America [40-60 N, -90~-60E] (Reg1), South South America [20~40S, -50~-40E] (Reg2), North Africa [10~40N, -10~-30E] (Reg3), East Australia [20~40S, 135~150E] (Reg4), and South East Asia [10~35N, 90~120E] (Reg5).

The time series of soil moisture at different depths from surface to root zone (SWL1, SWL2, SWL3 and SWL4) derived from CLM4.5 simulations and GLDAS data are given in Figure 8. As shown, the soil moisture in Reg1 follows similar pattern based on different CLM4.5 simulations and GLDAS data in each soil layer from surface to deep zone. Clearly, the assimilated AMSR_E SM data in CLM4.5

reduce the bias with respect to GLDAS data. The fluctuation of trends in the surface soil layer is obvious due to seasonal weather pattern while it tends to be stable with soil depth increased. The Reg2 reflects similar results with Reg1. However, the soil moisture pattern in regions Reg3 and Reg4 from CLM4.5 assimilated with AMSR_E SM data (ASC, DSC, DDA) agrees well with the GLDAS data with an exception of the initial 1-3 months, which may be due to the initial model spin up with AMSR_E SM data assimilation. The simulations in Reg1 to Reg4 based on open loop scenario overestimate the soil moisture in each layer while the results from ASC, DSC and DDA assimilated with AMSR_E SM data reduces such higher bias with respect to the GLDAS data, which is considered as the benchmark for comparison studies (e.g., Liu et al., 2014a; Liu et al., 2014b). In Reg5, the performance of simulations from CLM4.5 open loop and DDA matches well with the GLDAS data though the ASC and DSC tend to increase the bias at some times. Overall, the AMSR_E SM data assimilated in CLM4.5 model improves the performance of CLM4.5 model by minimizing the bias with respect to GLDAS data.

Similar to the changes in the grid cells, the time series of ET ([Figure S16](#)) derived from the CLM4.5 model assimilated with AMSR_E SM data did not show significant improvement in Reg1, Reg2 and Reg5 regions while decrease the deviation with respect to GLDAS in Reg3 and Reg4 regions. The runoff simulations from the CLM4.5 open loop have higher magnitude compared to the GLDAS while the CLM4.5 model assimilated with AMSR_E SM data reduced such bias over each region ([Figure S17](#)), indicating the improvement of CLM4.5 model performance when assimilated with AMSR_E SM data.

In all, the simulations of soil moisture, ET or runoff over each region compare well with the grid cells. Both results prove that the assimilation of AMSR_E SM data in CLM4.5 can improve the performance of CLM4.5 modeling in the simulations of local hydrometeorological fluxes over most grid cells (except Reg5 over the South East Asia). The poor performance over Reg5 may be due to the complicated topography over that region, which may increase the uncertainties of AMSR_E SM data. Another reason may come from the uncertainties of GLDAS

data. The GLDAS data used in this paper to act as a kind of benchmark to evaluate the efficiency of CLM4.5 modeling performance is useful, however, GLDAS itself is a kind of assimilated data which may still have uncertainties. In the future, more observation data need to be explored in the assessment or assimilation of modeling.

5. Summary

This paper aims to investigate the performance of CLM4.5 model to simulate hydrologic fluxes with or without remote sensing data assimilation. The AMSR_E SM data (both ascending and descending) is assimilated into the CLM4.5 model to evaluate the performance of CLM4.5 model.

The CLM4.5 simulations in the open loop overestimate the soil moisture over most of South America, East North America, Africa, Northwest Europe, East Australia and part of Northeast Asia along the Yellow River Basin and North India in the surface soil layer as well as in the root zone depths. The CLM4.5 model assimilated with AMSR_E SM data (both ascending and descending) improves the soil moisture simulations over most part of these areas. The CLM4.5 model with AMSR_E SM data assimilation using LEnKF technique outweighs the model directly substituted with AMSR_E SM data over Southeast Asia, North Europe, South Africa and South America.

The assimilation of AMSR_E SM into CLM4.5 model slightly improved the overall performance of the CLM4.5 model. Apparently, the improvement of ET and runoff simulations from CLM4.5 model assimilated with AMSR_E SM data are obvious over the regions with significant soil moisture contents in both the annual and seasonal time scales. The results from the sub regions agree well with the grid cells within the same region.

Finally, it should be remarked that the optimality of the assimilation scenarios depends to a large extent on the characterization of the model and observation errors, and the selected method for bias-correction. Both aspects clearly require further research. The model and observation errors used in this study were not specifically optimized with respect to performance of each of the assimilation scenarios. The

568 AMSR_E SM is rescaled to the same resolution with CLM4.5 modeling scale
569 without bias-corrected. However, a different evaluation could be obtained after
570 rigorous optimization. The bias-corrected of AMSR_E SM may require more
571 elaborated methods which account for the increased variability over time due to
572 contributions of temperature and vegetation, verified over longer data records. In this
573 context, a major science challenge lies in the development of techniques that
574 mitigate biases in SM while producing unbiased estimates of SM and associated
575 water fluxes.

576

577 **Acknowledgement**

578 This study was supported by the United States Department of Agriculture
579 (USDA) award 2015-68007-23210, the Fundamental Research Funds for the Central
580 Universities of China, the Special Fund of State Key Laboratory of Hydrology-Water
581 Resources (No.20145027312), and the National Key R&D Program of China (Grant
582 No.2016YFC0402706, 2016YFC0402710). The GLDAS data used in this study were
583 acquired as part of the mission of NASA's Earth Science Division and archived and
584 distributed by the Goddard Earth Sciences (GES) Data and Information Services
585 Center (DISC).

586

References

- Ahmed, M., Sultan M., Yan E., and Wahr J., 2016. Assessing and improving land surface model outputs over Africa using GRACE, field, and remote sensing data. *Surveys in Geophysics*, 37(3), 529-556.
- Andersen, O.B., Seneviratne S.I., Hinderer J., and Viterbo P., 2005. GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave. *Geophys. Res. Lett.* 32, L18405.
- Awange, J.L., Gebremichael M., Forootan E., Wakbulcho G., Anyah R., V. Ferreira G., and Alemayehu T., 2014. Characterization of Ethiopian mega hydrogeological regimes using GRACE, TRMM and GLDAS datasets. *Advances in Water Resources* 74, 64-78.
- Bartalis, Z., Wagner, W., Naeimi, V., et al., 2007. Initial soil moisture retrievals from the METOPA Advanced Scatterometer (ASCAT). *Geophys. Res. Lett.* 34, L20401.
- Chen, Y., Yang, K., Qin, J., et al., 2013. Evaluation of AMSR-E retrievals and GLDAS simulations against observations of a soil moisture network on the central Tibetan plateau. *J. Geophys. Res. Atmos.* 118, 4466-4475.
- Cheng, S., Guan X., Huang J., Ji F., and Guo R., 2015. Long-term trend and variability of soil moisture over East Asia. *Journal of Geophysical Research: Atmospheres*, 120(17), 8658-8670.
- Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. *Water Resour. Res.* 14, 601-604.
- Cosby, B.J., Hornberger, G.M., Clapp, R.B., et al., 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* 20, 682-690.
- Crow, W., Zhan, X., 2007. Continental-scale evaluation of remotely sensed soil moisture products. *IEEE Geoscience and Remote Sensing Letters* 4, 451 - 455
- De Lannoy, G.J.M., Reichle, R.H., Arsenault, K.R., et al., 2012. Multiscale assimilation of Advanced Microwave Scanning Radiometer-EOS snow water

equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado. *Water Resour. Res.* 48, W01522.

de Rosnay, P., Drusch, M., Boone, A., et al., 2009. AMMA land surface model intercomparison experiment coupled to the community microwave emission model: ALMIP-MEM. *J. Geophys. Res.* 114, D05108.

Dirmeyer, P.A., 2011. The terrestrial segment of soil moisture-climate coupling. *Geophys. Res. Lett.* 38, L16702.

Draper, C., Walker, J., Steinle, P., et al., 2009. An evaluation of AMSR-E derived soil moisture over Australia. *Remote Sens. Environ.* 113, 703-710.

Entekhabi, D., Reichle, R., Koster, R., et al., 2010. Performance metrics for soil moisture retrieval and application requirements. *J. Hydrometeorol.* 11, 832-840.

Fu, X., Yu, Z., Luo, L., et al., 2014. Investigating soil moisture sensitivity to precipitation and evapotranspiration errors using SiB2 model and ensemble Kalman filter. *Stochastic Environmental Research and Risk Assessment* 28, 681-693.

Han, X., Franssen, H.-J.H., Montzka, C., et al., 2014. Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations. *Water Resour. Res.* 50, 6081-6105.

Hogue T.S., Bastidas L., Gupta H., Sorooshian S., Mitchell K., Emmerich W., 2005. Evaluation and transferability of the Noah land surface model in semiarid Environments. *J. Hydrometeorol.* 6, 68–84.

Houser, P., Shuttleworth, W., Famiglietti, J., et al., 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation *Water Resour. Res.* 34, 3405-3420.

Katra, I., Blumberg, D., Lavee, h., et al., 2007. Spatial distribution dynamics of topsoil moisture in shrub microenvironment after rain events in arid and semiarid areas by means of high-resolution maps. *Geomorphology* 86, 455-464.

Kawanishi, T., Sezai, T., Ito, Y., et al., 2003. The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASDA's contribution to the EOS for global energy and water cycle studies. *IEEE Trans. Geosci.*

647 Remote Sens. 41, 184-194.

648 Kerr, Y.H., Waldteufel, P., Wigneron, J.P., et al., 2001. Soil moisture retrieval from
649 space: The Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Trans.
650 Geosci. Remote Sens. 39, 1729-1735.

651 Koster, R.D., Dirmeyer, P.A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C.T.,
652 Kanae, S., Kowalczyk, E., Lawrence, D. and Liu, P., 2004a. Regions of strong
653 coupling between soil moisture and precipitation. Science 305(5687), 1138-1140.

654 Koster, R.D., Guo, Z.C., Dirmeyer, P.A., et al., 2006. GLACE: The global
655 land-atmosphere coupling experiment. Part I: Overview. J. Hydrometeor. 7,
656 590-610.

657 Koster, R.D., Suarez M.J., Liu P., Jambor U., Kistler M., Berg A., Reichle R., Rodell
658 M., and Famiglietti J., 2004b. Realistic initialization of land surface states:
659 Impacts on subseasonal forecast skill, J. Hydrometeorol. 5(6), 1049–1063,
660 doi:10.1175/JHM-387.1.

661 Kumar, S.V., Reichle, R.H., Peters-Lidard, C.D., et al., 2008. A land surface data
662 assimilation framework using the land information system: Description and
663 applications. Adv. Water Resour. 31, 1419-1432.

664 Lannoy, G.J.M.D., Reichle, R.H., 2016. Global Assimilation of Multiangle and
665 Multipolarization SMOS Brightness Temperature Observations into the GEOS-5
666 Catchment Land Surface Model for Soil Moisture Estimation. Journal of
667 Hydrometeorology 17, 669-691.

668 Lawrence, D.M., Slater, A.G., 2007. Incorporating organic soil into a global climate
669 model. Clim. Dyn. 30, 145-160.

670 Li, H., Huang, M., Wigmosta, M.S., et al., 2011. Evaluating runoff simulations from
671 the Community Land Model 4.0 using observations from flux towers and a
672 mountainous watershed. J. Geophys. Res. 116, D24120.

673 Liang, X., Lettenmaier, D.P., Wood, E.F., et al., 1994. A simple hydrologically based
674 model of land surface water and energy fluxes for GSMs. J. Geophys. Res. 99,
675 14,415-414,428.

676 Lievens, H., Lannoy, G.J.M.D., Bitar, A.A., et al., 2015. Assimilation of SMOS soil

moisture and brightness temperature products into a land surface model. Remote Sensing of Environment.

Liu, D., Wang, G.L., Mei, R., et al., 2014a. Impact of initial soil moisture anomalies on climate mean and extremes over Asia. *J. Geophys. Res.* 119, 1-14.

Liu, D., Wang, G.L., Mei, R., et al., 2014b. Diagnosing the strength of land-atmosphere coupling at sub-seasonal to seasonal time scales in Asia. *J. Hydrometeor.* 15, 320-339.

Liu, D., Yu Z., and Zhang J. Y., 2015. Diagnosing the Strength of Soil Temperature in the Land-Atmosphere Interactions over Asia Based on RegCM4 Model. *Global and Planetary Change* 130, 7-21.

Liu, Y., Gupta, H.V., 2007. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.* 43, W07401.

Liu, Y.Y., Dijk, A.I.J.M.v., Jeu, R.A.M.d., et al., 2009. An analysis of spatiotemporal variations of soil and vegetation moisture from a 29-year satellite-derived data set over mainland Australia. *Water Resour Res* 45, W07405.

McCabe, M., Wood, E.H., Gao, H., 2005. Initial soil moisture retrievals from AMSR-E: Multiscale comparison using in situ data and rainfall patterns over Iowa. *Geophysical Research Letters* 32, 6.

Mei, R., Wang, G.L., Gu, H.H., 2013. Summer Land-Atmosphere Coupling Strength over the United States: Results from the Regional Climate Model RegCM4-CLM3.5. *Journal of Hydrometeorology* 14, 946-962.

Moiwo J.P. and Tao F.L., 2015. Contributions of precipitation, irrigation and soil water to evapotranspiration in (semi)-arid regions. *Int. J. Climatol.* 35, 1079–1089. Doi: 10.1002/joc.4040.

Montzka, C., Moradkhani, H., Weihermuller, L., et al., 2011. Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter. *J. Hydrol.* 399, 410-421.

Moradkhani, H., DeChant, C.M., Sorooshian, S., 2012. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. *Water Resour. Res.* 48, W12520.

707 Moradkhani, H., Sorooshian, S., Gupta, H.V., et al., 2005. Dual state–parameter
 708 estimation of hydrological models using ensemble Kalman filter. *Adv. Water*
 709 *Resour.* 28, 135-147.

710 Mostovoy, G., Anantharaj, V., 2008. Observed and simulated soil moisture
 711 variability over the lower Mississippi Delta Region. *Amer. Meteor. Soc.* 9,
 712 1125-1150.

713 Nie, S., Zhu, J., Luo, Y., 2011. Simultaneous estimation of land surface scheme
 714 states and parameters using the ensemble Kalman filter: Identical twin
 715 experiments. *Hydrol. Earth Syst. Sci.* 15, 2437-2457.

716 Njoku, E., Ashcroft, P., Chan, T., et al., 2005. Global survey of statistics of
 717 radiofrequency interference in AMSR-E land observations. *IEEE Transactions*
 718 *on Geoscience and Remote Sensing* 43, 938-947.

719 Njoku, E.G., Jackson, T.J., Lakshmi, V., et al., 2003. Soil moisture retrieval from
 720 AMSR-E. *IEEE Trans. Geosci. Remote Sens.* 41, 215-229.

721 Oleson, K., et al., 2013. Technical description of version 4.5 of the Community Land
 722 Model (CLM), NCAR Tech.l Note NCAR/TN-5031STR. 422. Natl. Cent. for
 723 Atmos. Res., Boulder, Colo.

724 Pan, M., Wood, E.F., 2010. Impact of accuracy, spatial availability, and revisit time
 725 of satellite-derived surface soil moisture in a multiscale ensemble data
 726 assimilation system. *IEEE J. STARS* 3, 49-56.

727 Qian, T., Dai, A., Trenberth, K.E., et al., 2006. Simulation of global land surface
 728 conditions from 1948 to 2004: Part I: Forcing data and evaluations. *J.*
 729 *Hydrometeor.* 7, 953-975.

730 Reichle, R.H., Koster, R.D., 2005. Global assimilation of satellite surface
 731 soilmoisture retrievals into the NASA catchment land surface model. *Geophys.*
 732 *Res. Lett.* 32, L02404.

733 Rodell M., Famiglietti J.S., Chen J., Seneviratne S.I., Viterbo P., Holl S., Wilson C.R..
 734 2004a. Basin scale estimates of evapotranspiration using GRACE and other
 735 observations. *Geophys. Res. Lett.* 31, L20504.

736 Rodell M., Houser P.R., Jambor U., Gottschalck J., Mitchell K., Meng C.J., Arsenault

- K., Cosgrove B., Radakovich J., Bosilovich M., Entin J.K., Walker J.P., Lohmann D., Toll D., 2004b. The global land data assimilation system. *Bull. Am. Meteorol. Soc.* 85, 381–394.
- Rüdiger, C., Calvet, J.C., Gruhier, C., et al., 2009. An Intercomparison of ERS-Scat and AMSR-E Soil Moisture Observations with Model Simulations over France. *Journal of Hydrometeorology* 10, 431-447.
- Segele, Z.T., Leslie, L.M., Lamb, P.J., 2013. Weather Research and Forecasting Model simulations of extended warm-season heavy precipitation episode over the US Southern Great Plains: data assimilation and microphysics sensitivity experiments. *Tellus A* 65, 1-28.
- Seneviratne S.I., Lüthi D., Litschi M. and Schär C., 2006. Land–atmosphere coupling and climate change in Europe. *Nature* 443(7108): 205-209.
- Seyyedi H., Anagnostou E.N., Beighley E., McCollum J., 2015. Hydrologic evaluation of satellite and reanalysis precipitation datasets over a mid-latitude basin. *Atmospheric Research*, 164-165, 37-48.
- Swenson, S.C., Lawrence, D.M., Lee, H., 2012. Improved simulation of the terrestrial hydrological cycle in permafrost regions by the Community Land Model. *J. Adv. Model. Earth Syst.* 4, M08002.
- Syed, T.H., Famiglietti J.S., Rodell M., Chen J.L., and Wilson C.R., 2008. Analysis of terrestrial water storage changes from GRACE and GLDAS, *Water Resour. Res.* 44, W02433, doi:10.1029/2006WR005779.
- Ulaby, F.T., Moore, R.K., Fung, A.K., 1986. *Microwave Remote Sensing, Active and Passive*. Dedham, MA: Artech, 1986, vol. 3, From Theory to Applications.
- Wei, J.F., Dirmeyer, P.A., 2012. Dissecting soil moisture-precipitation coupling. *Geophys. Res. Lett.* 39, L19711.
- Xiao, Z.Q., Jiang, L.M., Zhu, Z.L., et al., 2016. Spatially and Temporally Complete Satellite Soil Moisture Data Based on a Data Assimilation Method. *Remote Sens.* 8, 1-16.
- Yang, T., Wang C., Yu Z., and Xu F., 2013. Characterization of spatio-temporal patterns for various GRACE-and GLDAS-born estimates for changes of global

terrestrial water storage. *Global and planetary change*, 109, 30-37.

Yang, Z.-L. 1998. Technical note of a 10-layer soil moisture and temperature model. Unpublished manuscript.

Yin, J., Zhan, X., Zheng, Y., et al., 2015. Optimal ensemble size of ensemble Kalman filter in sequential soil moisture data assimilation. *Geophys. Res. Lett.* 42.

Yin, J., Zhan, X., Zheng, Y., et al., 2014. Impact of quality control of satellite soil moisture data on their assimilation into land surface model. *Geophys. Res. Lett.* 41, 7159-7166.

Yu, Z., Liu, D., Lü, H., et al., 2012. A multi-layer soil moisture data assimilation using support vector machines and ensemble particle filter. *J. Hydro.* 475 53-64.

Zeng, X., and Decker M., 2009. Improving the numerical solution of soil moisture-based Richards equation for land models with a deep or shallow water table. *J. Hydrometeor.* 10, 308-319.

Zhang, J., Wang W.C., and Wei J., 2008. Assessing land-atmosphere coupling using soil moisture from the Global Land Data Assimilation System and observational precipitation. *J. Geophys. Res.* 113, D17119, doi:10.1029/2008JD009807.

Zhang, J.Y., Wu, L.Y., Dong, W.J., 2011. Land-atmosphere coupling and summer climate variability over East Asia. *J. Geophys. Res.* 116, 1-14.

Figure Captions:

Fig. 1: Framework of CLM4.5-VIC model data assimilation with AMSR_E soil moisture.

Fig. 2: (a) Mean annual and seasonal soil water in the upper 10 cm (SWL1) (mm^3/mm^3) from CLM4.5 open loop simulations during 2003–2004 (Open Loop). As in (a), (b)~(d) for the difference between sensitive CLM4.5 experiments with open loop: (b) ASC (ASC-OL); (c) DSC (DSC-GL); (d) DDA (DDA-GL); (e) difference between ASC and DSC (ASC-DSC); (f) difference between DDA and DSC (DDA-DSC). ANN, MAM, JJA, SON, and DJF for annual mean, spring, summer, autumn and winter seasons, respectively.

Fig. 3: (a) Mean annual and seasonal soil water in the upper 10 cm (SWL1) (mm^3/mm^3) from GLDAS data during 2003–2004 (GLDAS). As in (a), but for the difference between CLM4.5 experiments with GLDAS: (b) open loop (OL-GL); (c) ASC (ASC-GL); (d) DSC (DSC-GL); and (e) DDA (DDA-GL). ANN, MAM, JJA, SON, and DJF for annual mean, spring, summer, autumn and winter seasons, respectively.

Fig. 4: MAE, RMSE and R between CLM4.5 simulations with GLDAS data during 2003-2004: (A1) between CLM4.5 open loop and GLDAS; (A2) between ASC with GLDAS; (A3) between DSC with GLDAS; (A4) between DDA with GLDAS. The GLDAS data are upscaled to the same resolution as CLM4.5.

Fig. 5: AE for soil water at different depths (L1, L2, L3, L4) derived from the sensitivity CLM4.5 simulations updated with AMSR_E SM data: ASC (A1); DSC (A2); DDA (A3). All the data used here are accumulated into monthly data. The GLDAS data are upscaled to the same resolution as CLM4.5.

Fig. 6: (a) Mean annual and seasonal ET (mm/day) from GLDAS data during 2003–2004 (GLDAS). As in (a), but for the difference between CLM4.5 experiments with GLDAS: (b) open loop (OL-GL); (c) ASC (ASC-GL); (d) DSC (DSC-GL); and (e) DDA (DDA-GL). ANN, MAM, JJA, SON, and DJF for annual mean, spring, summer, autumn and winter seasons, respectively.

Fig. 7: Same as Fig. 6, but for the surface runoff (mm/day).

Fig. 8: Monthly time series of soil water content within different depth (L1, L2, L3 and L4) derived from GLDAS, CLM4.5 open loop (OL), ASC (ASC), DSC (DSC), DDA (DDA) during 2003~2004 years in the five sub-regions: (Reg1) East North America [40~60 N, -90~-60E]; (Reg2) South South America [20~40S, -50~-40E]; (Reg3) North Africa [10~40N, -10~-30E]; (Reg4) East Australia [20~40S, 135~150E]; and (Reg5) South East Asia [10~35N, 90~120E].

Figures

Fig. 1: Framework of CLM4.5-VIC model data assimilation with AMSR_E soil moisture.

Fig. 2: (a) Mean annual and seasonal soil water in the upper 10 cm (SWL1) (mm^3/mm^3) from CLM4.5 open loop simulations during 2003–2004 (Open Loop). As in (a), (b)~(d) for the difference between sensitive CLM4.5 experiments with open loop: (b) ASC (ASC-OL); (c) DSC (DSC-GL); (d) DDA (DDA-GL); (e) difference between ASC and DSC (ASC-DSC) ; (f) difference between DDA and DSC (DDA-DSC). ANN, MAM, JJA, SON, and DJF for annual mean, spring, summer, autumn and winter seasons, respectively.

Fig. 3: (a) Mean annual and seasonal soil water in the upper 10 cm (SWL1) (mm^3/mm^3) from GLDAS data during 2003–2004 (GLDAS). As in (a), but for the difference between CLM4.5 experiments with GLDAS: (b) open loop (OL-GL); (c) ASC (ASC-GL); (d) DSC (DSC-GL); and (e) DDA (DDA-GL). ANN, MAM, JJA, SON, and DJF for annual mean, spring, summer, autumn and winter seasons, respectively.

Fig. 4: MAE, RMSE and R between CLM4.5 simulations with GLDAS data during 2003-2004: (A1) between CLM4.5 open loop and GLDAS; (A2) between ASC with GLDAS; (A3) between DSC with GLDAS; (A4) between DDA with GLDAS. The GLDAS data are upscaled to the same resolution as CLM4.5.

Fig. 5: AE for soil water at different depths (L1, L2, L3, L4) derived from the sensitivity CLM4.5 simulations updated with AMSR_E SM data: ASC (A1); DSC (A2); DDA (A3). All the data used here are accumulated into monthly data. The GLDAS data are upscaled to the same resolution as CLM4.5.

Fig. 6: (a) Mean annual and seasonal ET (mm/day) from GLDAS data during 2003–2004 (GLDAS). As in (a), but for the difference between CLM4.5 experiments with GLDAS: (b) open loop (OL-GL); (c) ASC (ASC-GL); (d) DSC (DSC-GL); and (e) DDA (DDA-GL). ANN, MAM, JJA, SON, and DJF for annual mean, spring, summer, autumn and winter seasons, respectively.

Fig. 7: Same as Fig. 6, but for the surface runoff (mm/day).

Fig. 8: Monthly time series of soil water content within different depth (L1, L2, L3 and L4) derived from GLDAS, CLM4.5 open loop (OL), ASC (ASC), DSC (DSC), DDA (DDA) during 2003~2004 years in the five sub-regions: (Reg1) East North America [40-60 N, -90~-60E]; (Reg2) South South America [20~40S, -50~-40E]; (Reg3) North Africa [10~40N, -10~-30E]; (Reg4) East Australia [20~40S, 135~150E]; and (Reg5) South East Asia [10~35N, 90~120E].