The Influence of Assimilating Leaf Area Index in a Land Surface Model on Global Water Fluxes and Storages

Xinxuan Zhang[1], Viviana Maggioni[1], Azbina Rahman[1], Paul Houser[1], Yuan Xue[1], Timothy Sauer[1], Sujay Kumar[2] and David Mocko[2]

1. George Mason University, Fairfax, VA, USA
2. Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Submit to:

Hydrology and Earth System Sciences

May, 2020

Abstract

Vegetation plays a fundamental role not only in the energy and carbon cycles, but also in the global water balance by controlling surface evapotranspiration (ET). Thus, accurately estimating vegetation-related variables has the potential to improve our understanding and estimation of the dynamic interactions between the water, energy, and carbon cycle. This study aims to assess to what extent a land surface model (LSM) can be optimized through the assimilation of leaf area index (LAI) observations at the global scale. Two Observing System Simulation Experiments (OSSEs) are performed to evaluate the efficiency of assimilating LAI into an LSM through an Ensemble Kalman Filter (EnKF) to estimate LAI, ET, canopy interception evaporation (CIE), canopy water storage (CWS), surface soil moisture (SSM), and terrestrial water storage (TWS). Results show that the LAI data assimilation framework not only effectively reduces errors in LAI model simulations, but also improves all the modeled water flux and storage variables considered in this study (ET, CIE, CWS, SSM, and TWS), even when the forcing precipitation is strongly positively biased (extremely wet condition). However, it tends to worsen some of the modeled water-related variables (SSM and TWS) when the forcing precipitation is affected by a dry bias. This is attributed to the fact that the amount of water in the LSM is conservative and the LAI assimilation introduces more vegetation, which requires more water than what available within the soil.

1. Introduction

Terrestrial vegetation plays a vital role in the global water cycle, as it controls the surface evapotranspiration (ET) and the state of the carbon cycle. As shown in past literature, there exists a strong relationship among vegetation, precipitation, and soil moisture (Di et al., 1994; Farrar et al., 1994; Richard and Poccard, 1998; Adegoke and Carleton, 2002). Nevertheless, the role that vegetation and its dynamics play in the water cycle (for instance on the variability of precipitation) is extremely complex (Wang and Eltahir 2000; Wang et al. 2011). In the past half-century, these land surface processes and feedbacks have been examined through numerical modeling experiments (Foley et al. 1996; Kim and Wang 2007; Druel et al. 2019). In early generation land surface models (LSMs), the development stage of vegetation was prescribed by regularly updating vegetation variables, based on fixed lookup tables to simplify the model computation (Foley et al. 1996). This approach uses constant vegetation indices, e.g., the leaf area index (LAI), while in reality the growth of vegetation continuously changes in response to weather and climate conditions. To overcome this deficiency, new generation LSMs are coupled with dynamic vegetation modules that comprehensively simulate several biogeochemical processes (Woodward and Lomas 2004; Gibelin et al. 2006; Fisher et al. 2018) and that are able to capture more detailed variations in plant productivity than traditional methods (Kucharik et al. 2000; Arora 2002; Krinner et al. 2005).

LAI can also be estimated through observations from satellite sensors, such as the Moderate Resolution Imaging Spectroradiometer (MODIS, Pagano and Durham 1993; Justice et al. 2002), the Système Probatoire d'Observation de la Terre VEGETATION (SPOT-VGT, Baret et al. 2007), and the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR, Cracknell 1997). LAI products retrieved from different satellite missions and sensors provide spatially and temporally varying LAI fields on a routine basis at regional and global scales, including the MODIS LAI (Myneni et al. 2002), the Global Land Surface Satellite (GLASS) LAI (Xiao et al. 2013), and the GLOBMAP LAI dataset (Liu et al. 2012), among others. Satellite-derived LAI products were found to be affected by uncertainties due to the limitation of retrieval algorithms and vegetation type sampling issues (Cohen and Justice 1999; Privette et al. 2002; Tian et al. 2002; Morisette et al. 2002).

A method to combine the inherently incorrect estimates from satellite observations and model simulations is data assimilation (DA). One of the most common DA systems — the Ensemble Kalman Filter (EnKF; Evensen 2003) — dynamically updates the model error covariance information by producing an ensemble of model predictions, which are individual model realizations perturbed by the assumed model error (Reichle et al. 2007). The ensemble approach is widely used in hydrology because of its flexibility with respect to the type of model error (Crow and Wood 2003) and well suited to the nonlinear nature of land surface processes (Reichle et al. 2002a, 2002b; Andreadis and Lettenmaier 2006; Durand and Margulis 2008; Kumar et al. 2008; Pan and Wood 2006; Pauwels and De Lannoy 2006; Zhou et al. 2006). However, the use of an EnKF for the assimilation of LAI in LSMs has not been thoroughly investigated in the past. Pauwels et al. (2007) proposed an Observing System Simulation Experiment (OSSE) to evaluate the performance of assimilating LAI in a hydrology-crop growth model with an EnKF algorithm. Other studies have also tested simplified 1D-VAR and extended Kalman filter methods for LAI assimilation (e.g., Sabater et al. 2008; Barbu et al. 2011; Fairbairn et al. 2017). Recently, Kumar et al. (2019) assimilated GLASS LAI in a land surface model with an EnKF across the continental U.S. Some water budget variables were improved through the assimilation procedure, particularly in agricultural areas where the assimilation added harvesting information to the model.

Ling et al. (2019) assimilated global LAI information with an Ensemble Adjust Kalman Filter (EAKF) algorithm and found that the assimilation is more effective during the growing season. LAI assimilation also had a positive impact on gross primary production (GPP) and ET in low latitude regions.

Nevertheless, most of the aforementioned studies mainly focused on the impact of LAI assimilation on the simulated LAI or vegetation biomass. Only a few studies discussed the influences of LAI assimilation on the estimation of water variables such as soil moisture or streamflow (Pauwels et al. 2007; Sabater et al. 2008) and most of them focused on limited regions. Most recently, Albergel et al. (2017) conducted a study on a much larger domain – Europe and the Mediterranean basin –and showed improvement in soil moisture at various depths thanks to LAI assimilation.

This work leverages upon these studies but aims to assess to what extent a land surface model, especially the simulation of water-related variables, can be optimized through the assimilation of LAI observations at the global scale. As this study serves as a feasibility test to quantify the impact of LAI assimilation on water cycle variables, an OSSE is chosen to investigate the model's behavior. This guarantees that reference variables (often referred to as the "truth"), which are synthetically produced, are available for quantifying the performance of the proposed framework. Specifically, two OSSEs that apply an EnKF algorithm to the Noah LSM with multi-parameterization options (Noah-MP, Niu et al. 2011; Yang et al. 2011) are performed to evaluate the efficiency of assimilating LAI observations for estimating ET, canopy interception evaporation (CIE), canopy water storage (CWS), surface soil moisture (SSM), and terrestrial water storage (TWS).

2. Methods and materials

2.1. Land surface model (Noah-MP)

The Noah-MP 3.6 (Niu et al. 2011; Yang et al. 2011) is adopted in this study. Noah-MP contains a separate vegetation canopy defined by a canopy top and bottom, crown radius, and leaves with defined dimensions, orientation, density, and radiometric properties (Niu et al. 2011). Multiple options are available for surface water infiltration, runoff, groundwater transfer and storage including water table depth to an unconfined aquifer (Niu et al. 2007), dynamic vegetation, canopy resistance, and frozen soil physics. Specifically, the prognostic vegetation growth combines a Ball-Berry photosynthesis-based stomatal resistance (Ball et al. 1987) with a dynamic vegetation model (Dickinson et al. 1998). The dynamic vegetation model calculates the carbon storages in various parts of the vegetation (leaf, stem, wood, and root) and the soil carbon pools.

The Noah-MP 3.6 LSM has been implemented into the National Aeronautics and Space Administration (NASA) Land Information System (LIS; Peters-Lidard et al. 2007; Kumar et al. 2006). LIS is a software that provides an interagency test bed for land surface modeling and data assimilation that allows customized systems to be built, assembled and reconfigured easily, using shared plugins and standard interfaces. All the experiments in this study are setup through LIS. The Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2; Gelaro et al. 2017) dataset serves as the meteorological forcings of Noah-MP. MERRA-2 is the latest atmospheric reanalysis produced by the NASA Global Modeling and Assimilation Office (GMAO) and includes updates from the Goddard Earth Observing System (GEOS). The meteorological variables selected from MERRA-2 include surface pressure, surface air temperature, surface specific humidity, incident radiations, wind speed, and precipitation rate.

Five model output variables that describe terrestrial water fluxes and storages are investigated in this work: ET (defined as the sum of evaporation and the plant transpiration [kg/m^2s]), CIE (defined as the evaporation of the canopy intercepted water [kg/m^2s]), CWS (defined as the amount of canopy intercepted water in both liquid and ice phases [kg/m^2]), SSM (defined as the water content in the top 10 cm of the soil column [m^3/m^3]), and TWS (defined as the sum of all water storage on the land surface and in the subsurface [mm]).

2.2. Experimental design

An OSSE is designed to understand the efficiency of assimilating LAI within Noah-MP version 3.6 using a one-dimensional EnKF algorithm (Reichle et al. 2010), when the precipitation forcing data are strongly biased. Being the major driving force of the hydrological cycle, the quality of input precipitation is critical for the accuracy of land surface model outputs. However, global precipitation datasets are far from being perfect and often affected by large regional biases. For example, the MERRA-2 precipitation dataset shows a widespread relative bias greater than 100% in South Asia (Ghatak et al. 2018). Although an EnKF is optimal only under the assumption of unbiasedness (which is not met in the proposed experimental setup), we want to investigate here to what extent the EnKF LAI assimilation (even if sub-optimal) can improve water storages and fluxes under two extreme conditions, i.e., a very dry and a very wet precipitation bias, knowing that such biases are very plausible in the real world and often unknown (and therefore difficult to remove). The proposed framework is evaluated through a global experiment (Antarctica excluded) at the 0.625° × 0.5° spatial resolution of the MERRA-2 forcing dataset (Figure 1).

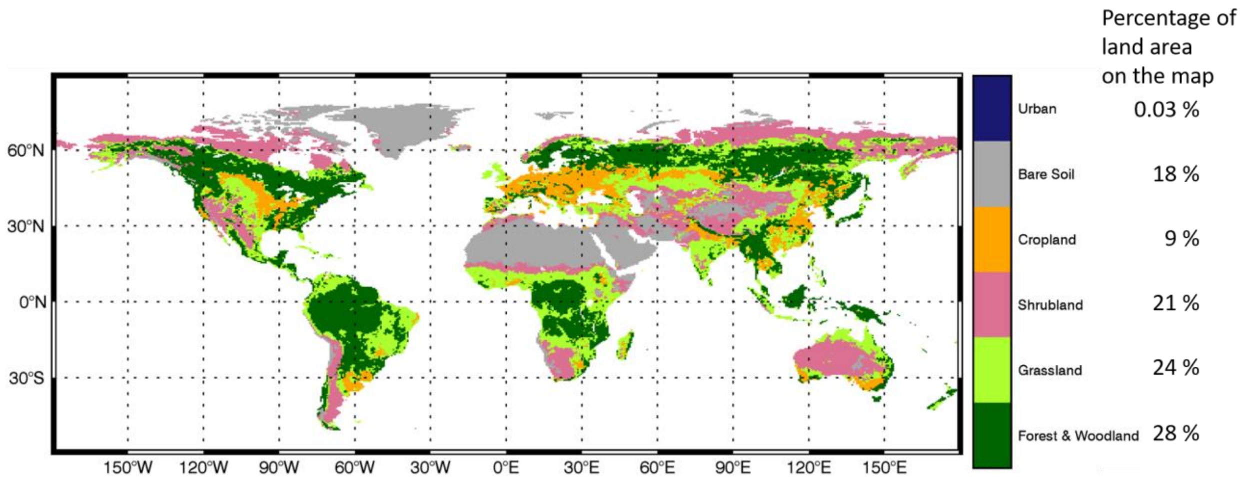


Figure 1. Study domain and land cover types (Hansen et al. 2000).

Figure 2 shows a schematic diagram of the experiments. First, the Noah-MP model is spun-up for a 10-year period (2001-2010) to ensure a physically realistic state of equilibrium. Second, the model is run for a 29-month period (January 2011 – May 2013) to conduct the Nature Run (NR) with the same configuration as the spin-up one. By definition, an OSSE is a controlled experiment that does not assimilate any real observation. Instead, it treats all the model outputs from the NR as the "true" condition (denoted as the "synthetic truth"). The "true" LAI (i.e., the LAI output from NR) is then perturbed via a simple additive error model to produce the synthetic observations to be assimilated into the DA runs. The spin-up run and NR are forced by the original MERRA-2 precipitation data. Third, two Open Loop (OL) runs (no DA) are conducted for the same 29-month period under two conditions: i) "extremely dry" condition (the model is forced by halving the MERRA-2 precipitation data; OL-dry), and ii) "extremely wet" condition (the model is forced by doubling the MERRA-2 precipitation; OL-wet). The biased forcing precipitation data in OL mimic typical precipitation biases in current precipitation reanalysis and satellite products (e.g., Ghatak et al. 2018; Yoon et al. 2019).

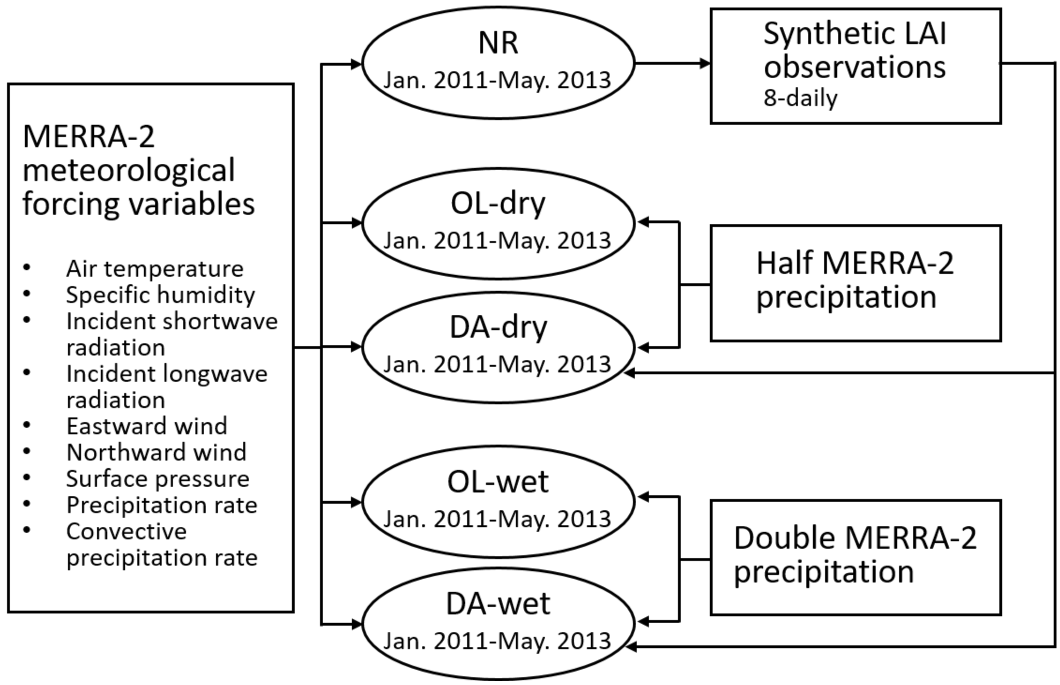


Figure 2. Schematic diagram of the OSSE design.

The two DA runs are then conducted under the two same conditions (DA-dry and DA-wet) using a one-dimensional EnKF assimilation algorithm, which is a built-in DA method in LIS. The EnKF DA algorithm is suitable for non-linear and intermittent land surface processes (Reichle et al. 2002a, 2002b). Details of the algorithm can be found in numerous previous studies (Reichle et al. 2010; De Lannoy et al. 2012; Liu et al. 2015; Kumar et al. 2019a).

The model ensemble is generated by perturbing a set of meteorological forcing. To select the optimal ensemble size, a sensitivity test is performed for ensemble sizes spanning from 2 to 24 members (not shown here). The number of ensemble members has a strong impact on the model results at small sizes, while the model performance tends to become steady when more than 20 ensemble members are considered. Thus, all the DA simulations are run for 20 members.

The synthetic LAI observations are obtained from the NR and assimilated to the DA system at 8-daily frequency. The synthetic LAI observation has the same temporal resolution as the

MODIS LAI product but with full coverage over the study domain. In real case studies, satellite LAI products contain a substantial amount of missing data, mainly due to the cloud obscuration gaps. Based on the vegetation type in the model, the leaf mass fields are also updated. Random perturbations of MERRA-2 meteorological forcings and synthetic LAI observations are applied to create an ensemble of land surface conditions that represent the uncertainties of in the LSM.

Similar to previous work (Kumar et al. 2014, 2019a, 2019b), the MERRA-2 forcing inputs such as shortwave/longwave radiations and precipitation are perturbed hourly. Multiplicative perturbations are applied to the shortwave radiation and precipitation with a mean of 1 and standard deviations of 0.3 and 0.5, respectively. The longwave radiation is perturbed via an additive perturbation with a standard deviation of 50 W/m^2. The perturbations of the three meteorological forcing variables also include cross correlations: cross correlation between shortwave radiation and precipitation is -0.8, cross correlation between longwave radiation and precipitation is 0.5; and cross correlation between shortwave and longwave radiations is -0.5. The synthetic LAI observations are perturbed via an additive model with a standard deviation of 0.1.

2.3. Evaluation and error metrics

Output variables from the OL and DA runs are evaluated against the "truth" from the NR at daily, monthly, and seasonal temporal scales. Besides LAI, five more water fluxes and storages are evaluated in the results section: ET, CIE, CWS, SSM, and TWS.

The initial condition for the OL and DA runs is generated by a spin-up run that uses the original MERRA-2 precipitation as input. However, the OL and DA runs are forced by either doubled or halved precipitation, which is not consistent with the spin-up run and the model needs some time to stabilize. The first 5-month model outputs are therefore eliminated from the evaluation to avoid the model systematic instability at the beginning of the OL and DA simulations and the evaluation, thus, focused only on model outputs from 2011-06-01 to 2013-05-31. Results are discussed using maps and time series of global averaged values and anomalies. Each of the anomaly time series is computed relative to the mean of its respective model run. Moreover, two error metrics are employed to quantify the difference between OL (and DA) with respect to the reference variables (from the NR). The first one is the Normalized and Centered Root Mean Square Error (NCRMSE), defined as follows:

$$E = \frac{\left\{\frac{1}{N}\sum_{i=1}^{N}\left[\left(X_i - mean(X)\right) - \left(O_i - mean(O)\right)\right]^2\right\}^{\frac{1}{2}}}{mean(O)} \qquad \text{Eq. 1}$$

where E is the NCRMSE, O is the NR output variable, and X is the output variable from the OL runs or DA runs. N is the total number of X values, and i represents the index of each X value. Second, to investigate the improvement (or degradation) due to the DA of LAI observations, we adopt the Normalized Information Contribution (NIC, similar to the NIC in Kumar et al. 2016) index based on NCRMSE and defined as:

$$C = \frac{E_{DA} - E_{OL}}{0 - E_{OL}} \qquad \text{Eq. 2}$$

where C represents the NIC index and E is the NCRMSE for OL or DA runs. NIC equals to 1 means that DA realizes the maximum possible improvement over the OL; NIC equals to zero means that DA and OL show the same performance skills; and negative NIC indicates a model degradation through DA.

3. Results and discussion

3.1. LAI

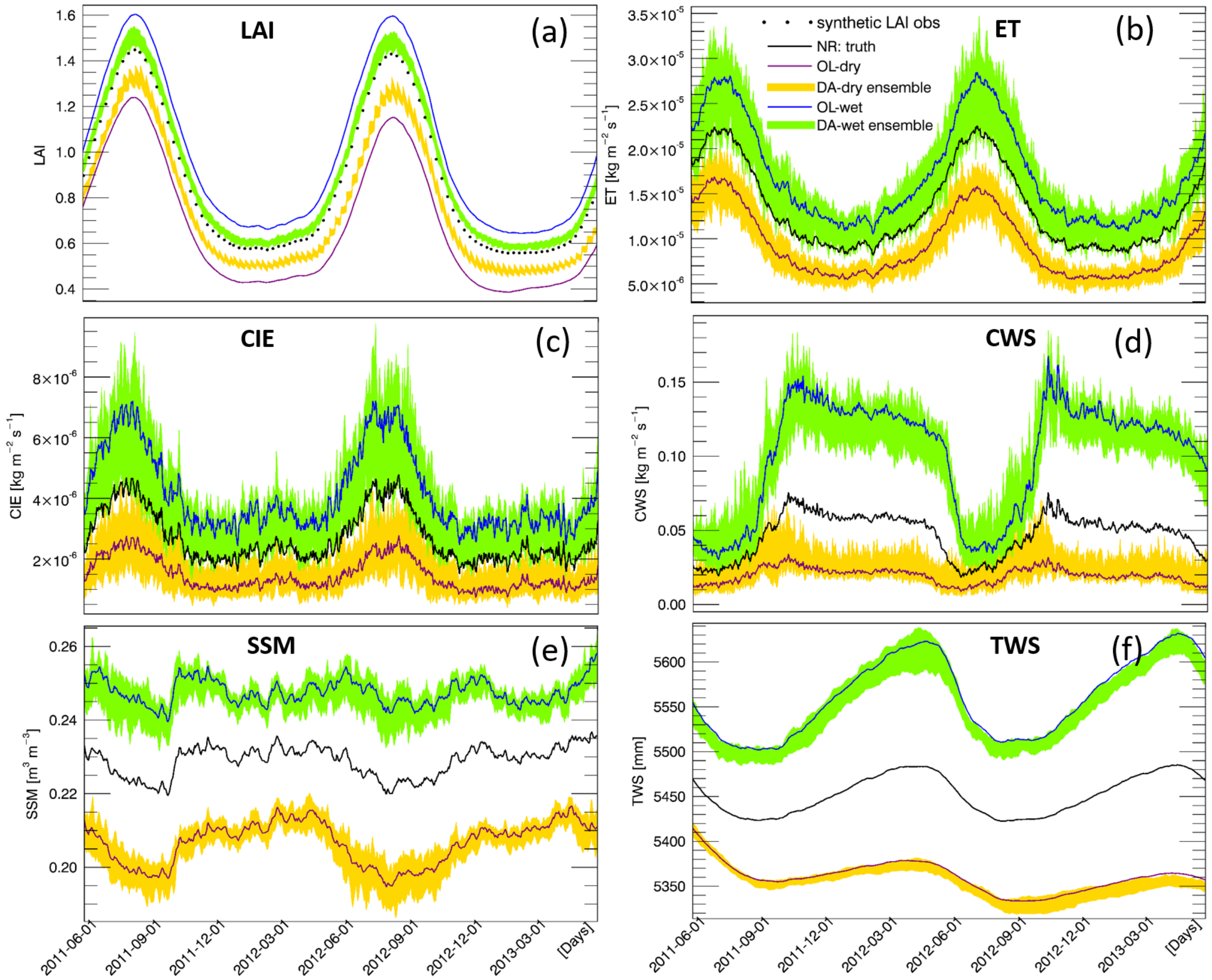


Figure 3. Global averaged daily values of LAI and five water variables (2011-06-01 to 2013-05-30).

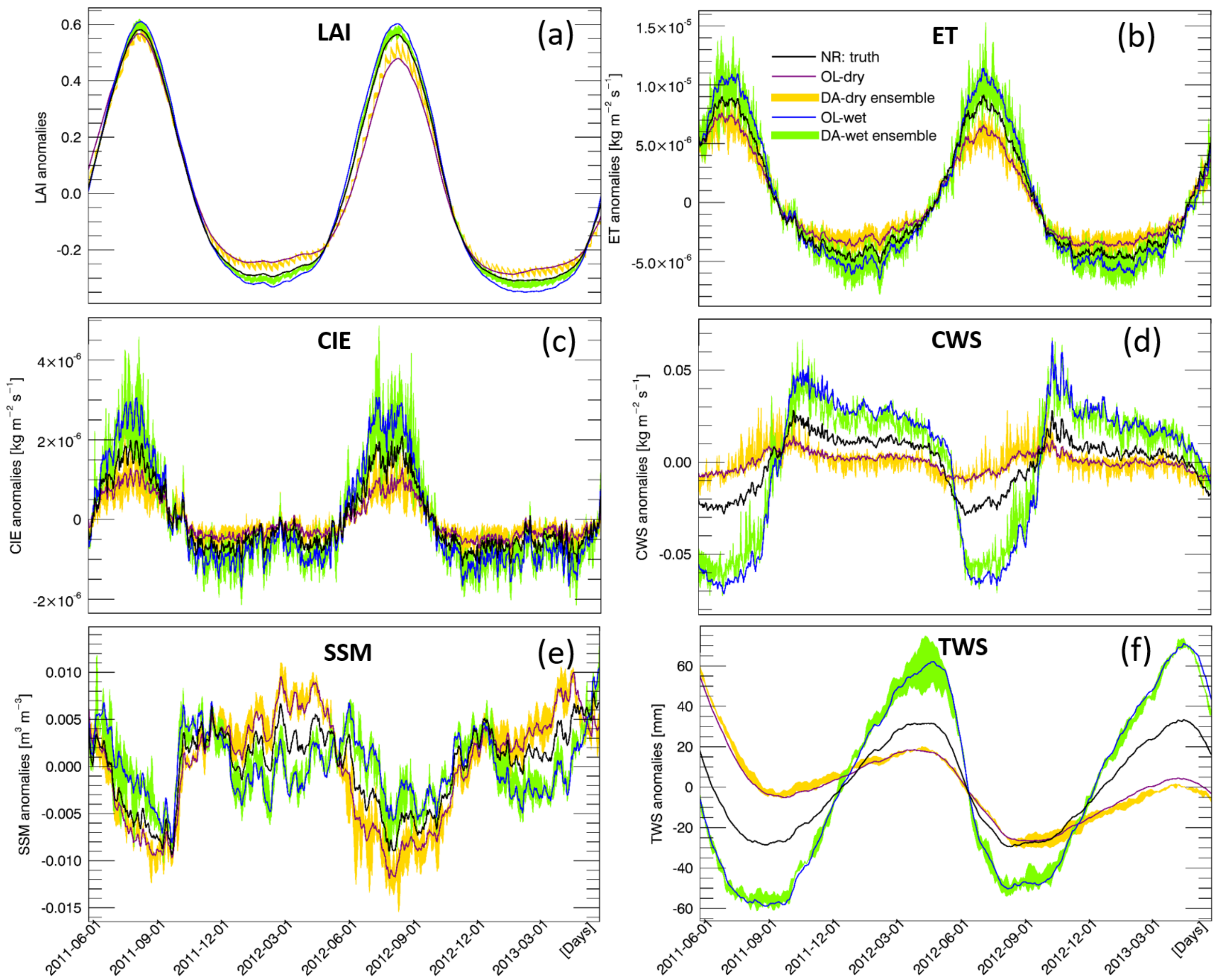


Figure 4. Global averaged daily anomalies of LAI and five water variables (2011-06-01 to 2013-05-30).

Figure 3a and Figure 4a show time series of global averaged LAI values and corresponding anomalies, respectively. As expected, LAI values are largely impacted by the extreme precipitation conditions. The wet condition introduces more vegetation, while the dry condition limits the vegetation growth throughout the two-year period. The DA procedure effectively corrects the LAI errors caused by the biased precipitation input. The seasonality of LAI anomalies is evident, showing larger variations in DJF and JJA than during the transition periods (MAM and SON). The OL-wet condition simulation shows larger LAI anomalies than the NR reference, while the OL-dry condition has smaller LAI anomalies than NR. The LAI anomalies obtained from DA runs under both wet and dry conditions are closer to the reference anomalies than the corresponding OL runs. In general, DA performs better in the wet condition experiment than in the dry case. Moreover, the DA runs show lower NCRMSEs than the corresponding OL runs across the globe (Figure 5a), especially over shrublands and grasslands (refer to Figure 1 for land covers).

In order to illustrate how LAI assimilation performs for different seasons, Figure 6a and Figure 7a show monthly averages of NCRMSE for LAI across the northern and southern hemispheres, respectively. In the northern hemisphere (Figure 6a), the NCRMSE time series follow clear seasonal patterns. First, the NCRMSE is higher in DJF/MAM and is lower in JJA/SON for both extreme precipitation conditions. The highest NCRMSE values are in March and April, and the lowest values are in July, August, and September. The differences of NCRMSE between OL and the corresponding DA runs tend to be much larger in MAM than in any other seasons, which means that LAI assimilation is more effective in the vegetation growth period. Moreover, the NCRMSE is constantly higher in the dry condition runs than the wet ones, which is due to the fact that the growth of vegetation is sensitive to the lack of water. Differences between wet and dry conditions are much smaller in JJA than in other seasons. In JJA, the vegetation leaves in the north hemisphere are fully developed and the plants can use stomatal closure to preserve water under water limited condition (dry condition). Thus, the NCRMSE of dry condition becomes smaller and does not show much difference from the wet condition. The southern hemisphere (Figure 7a), which does not have a strong climate seasonality, shows more modest seasonal NCRMSE patterns than the northern regions. In general, the NCRMSEs in the southern hemisphere are smaller than the ones in the northern hemisphere all year around. Specifically, NCRMSEs in the southern hemisphere are slightly higher in October, November, and December, when the differences between OL and DA runs are also larger.

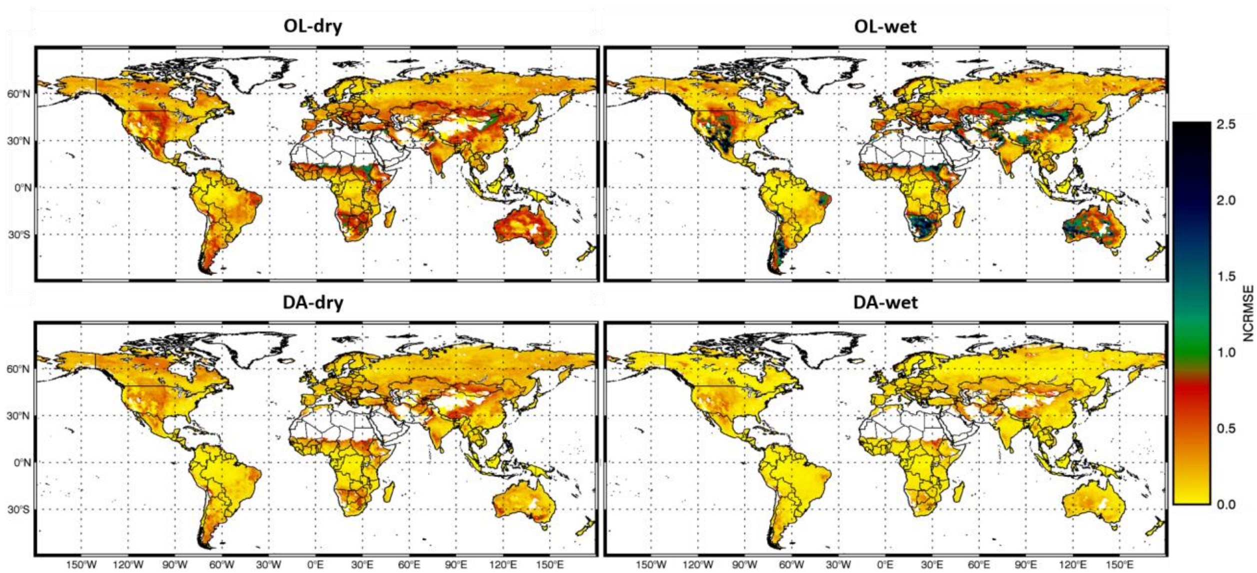


Figure 5. Maps of LAI NCRMSE for the OL and DA runs.

3.2. Water fluxes and storages

As mentioned in section 2.3, we focus on five water-related variables from the Noah-MP output to evaluate the impact of LAI assimilation on simulating the water cycle (ET, CIE, CWS, SSM, and TWS). Daily time series of global averaged values and corresponding anomalies of the five water variables are shown in Figure 3(b-f) and Figure 4(b-f), respectively. The model well simulates the seasonality of all water fluxes/storages considered here. The OL runs reveal that global average values of all five variables are impacted by the highly biased precipitation conditions. The variations of anomalies for ET, CIE, CWS, and TWS tend to be amplified by the wet condition and tend to be dampened by the dry condition. On the contrary, the anomalies of SSM become larger in dry conditions and become smaller in wet conditions, which is probably due to the limited soil water capacity. The surface soil is more likely to get saturated in wet conditions when the precipitation doubles the original amount, but SSM cannot get larger once the soil is saturated, even if there is more precipitation added to the system. Thus, the range of SSM anomalies in the wet experiment is limited and narrower than in the dry condition. The green and yellow shaded areas represent the ensemble of the DA runs. The anomaly ensembles of the five water variables show slight improvements through DA when precipitation is severely positively biased (wet condition). However, none of these variables shows improvement when the precipitation is severely negatively biased (dry condition) – the anomalies either have no change through the LAI DA (ET, CIE, and CWS) or worsen the OL-dry run (SSM and TWS).

To further investigate the efficiency of assimilating LAI in Noah-MP, time series of monthly NCRMSE averages are shown in Figure 6(b-f) and Figure 7(b-f) for all five water variables. The five variables can be divided into two main groups based on their performances: ET/CIE/CWS and SSM/TWS. For the wet bias experiment, DA improves the NCRMSE for all variables. However, LAI assimilation is not able to correct the model when the input precipitation is negatively biased (dry condition). A dry precipitation bias means that the system has (erroneously) less water than in reality (NR in the synthetic experiment). Since no water is otherwise added to the system, LAI DA cannot fully correct water-related model states (such as soil moisture). The NCRMSEs of DA runs are either the same as in the OL runs (ET/CIE/CWS) or worse (SSM/TWS). Specifically, ET/CIE/CWS have larger NCRMSE in the northern hemisphere and much smaller NCRMSEs in the southern hemisphere, but SSM/TWS do not show large differences between north and south. Moreover, ET/CIE/CWS in the northern hemisphere follow a seasonal pattern: NCRMSEs are lower in warm season (JJA) and higher in the colder season (DJF and March). In the southern hemisphere the three variables also have relative higher

NCRMSE in the colder season (JJA). On the contrary, SSM/TWS show a different seasonal pattern that NCRMSEs are larger in the warmer season (April, May, and June) over northern hemisphere. In southern hemisphere, TWS also has larger NCRMSEs in warmer season (October to April), but SSM shows higher NCRMSEs in colder season (similar to the ET/CIE/CWS group).

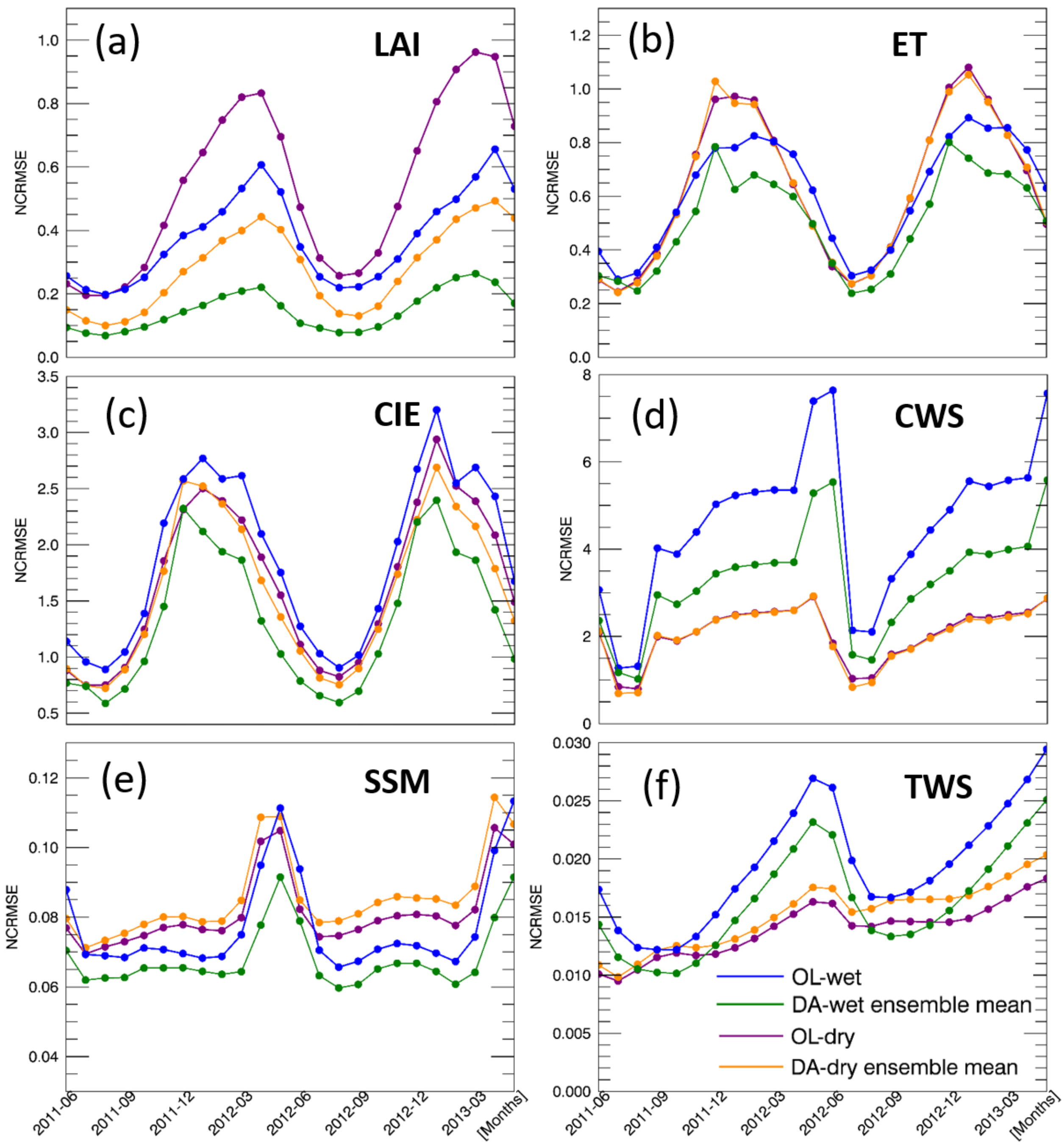


Figure 6. Monthly averaged NCRMSE for LAI and five water variables over the Northern hemisphere.

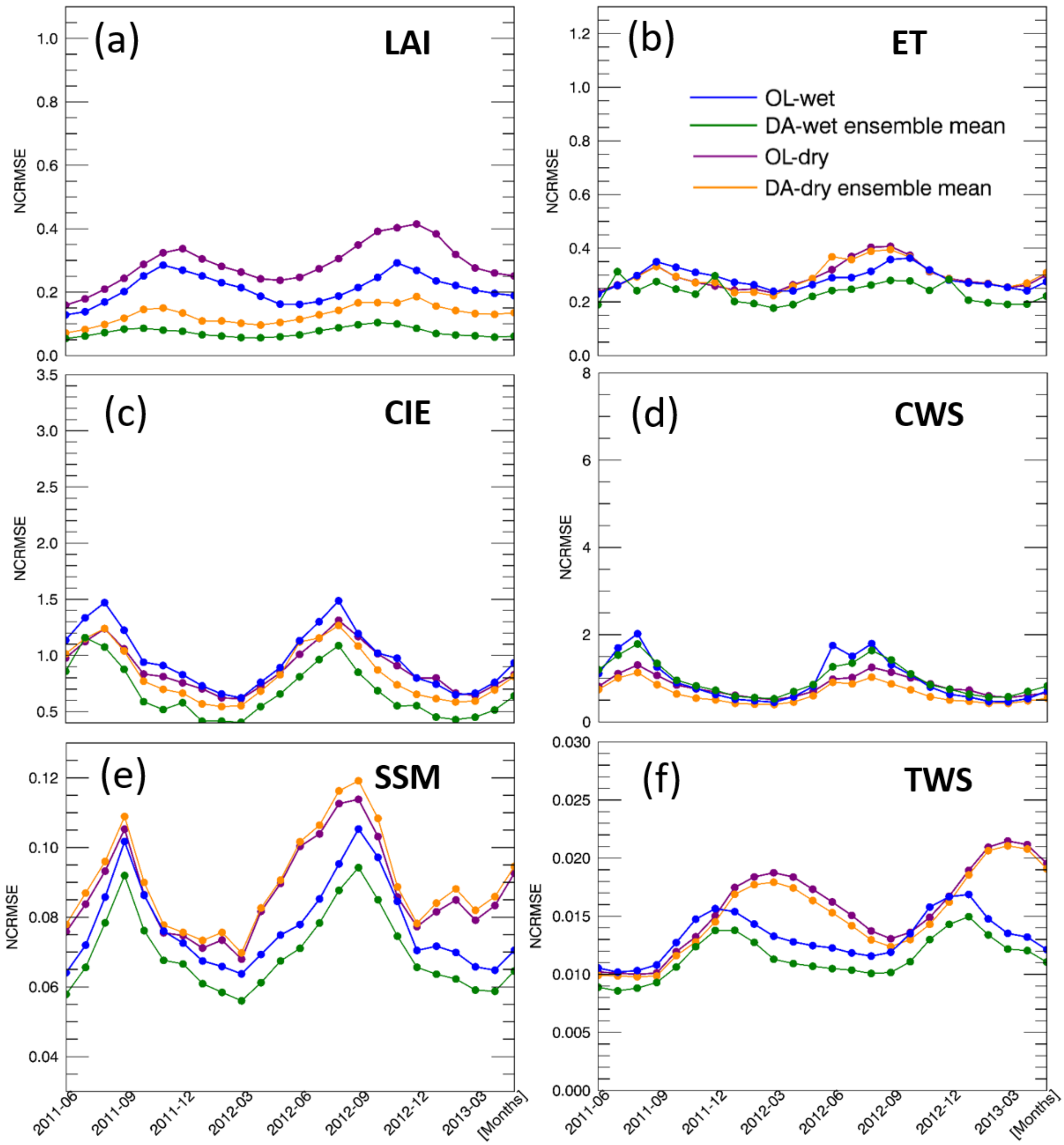


Figure 7. Same as in Figure 6, but for the Southern hemisphere.

The improvements in the model water fluxes and storages through LAI DA are also quantified by the NIC index (defined in Eq. 2). Figure 8 presents comparisons among NIC indices for each water variable analyzed in this study across areas with four different land cover types:

forest & woodland, grassland, shrubland, and cropland. In general, LAI DA improves the NIC indices with positively biased input precipitation (DA-wet) but worsens the NIC when negatively biased input precipitation (DA-dry) is considered. Specifically, in wet condition, ET, CIE, and CWS have higher variability over areas with different land cover types, while SSM and TWS have similar NIC values across different land covers. Shrubland and cropland tend to perform better in wet condition except for TWS. In dry condition, the NICs of ET, CIE, and TWS have higher variability than the ones of CWS and SSM. SSM and TWS show very low NIC values in dry condition for almost all land covers. Overall the NIC values of ET, CIE, and CWS are better than the ones of SSM and TWS for all land cover types, though the NICs of ET and CIE over forest & woodland perform very poorly.

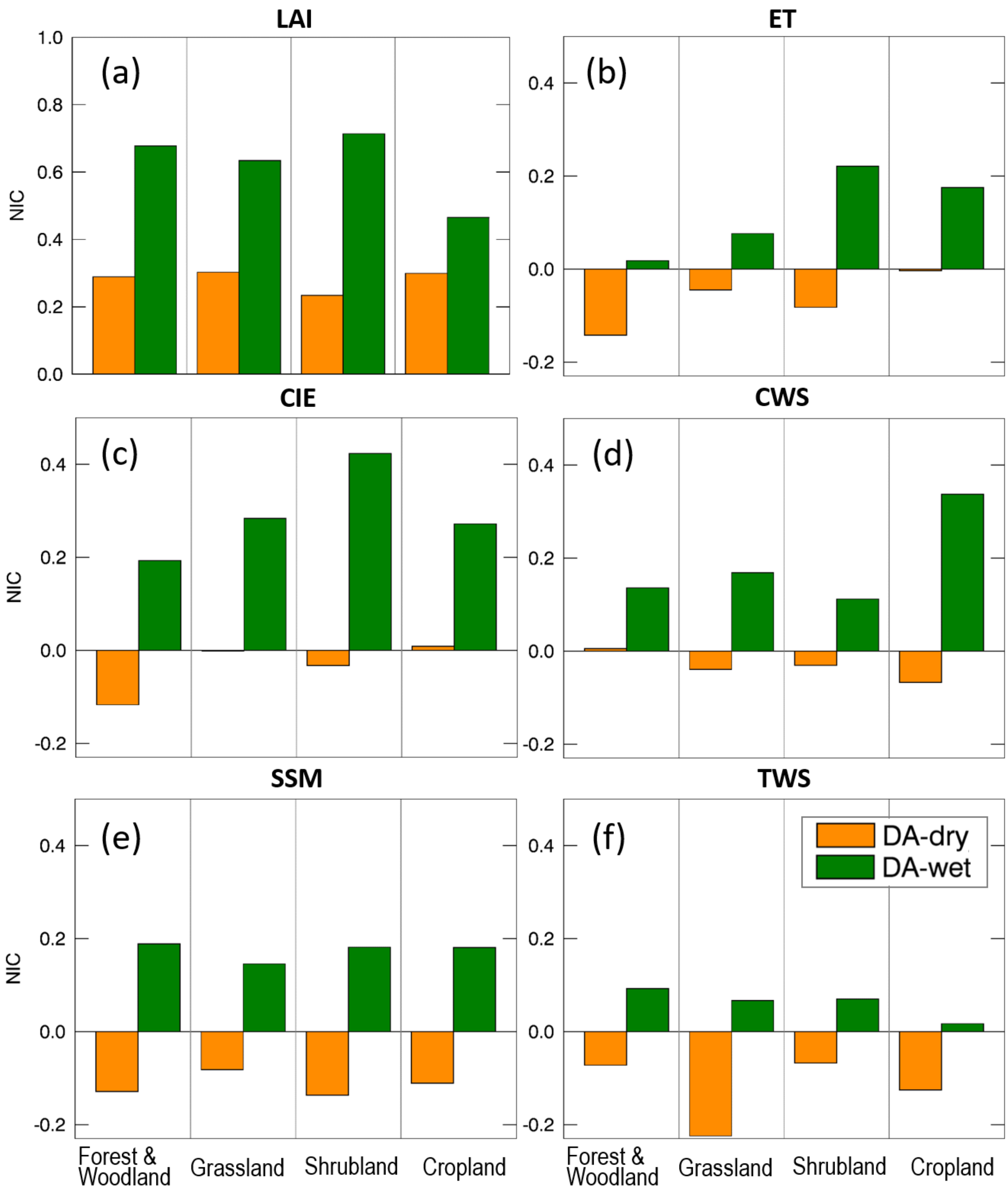


Figure 8. NIC for different variables and different land cover types for the two DA runs.

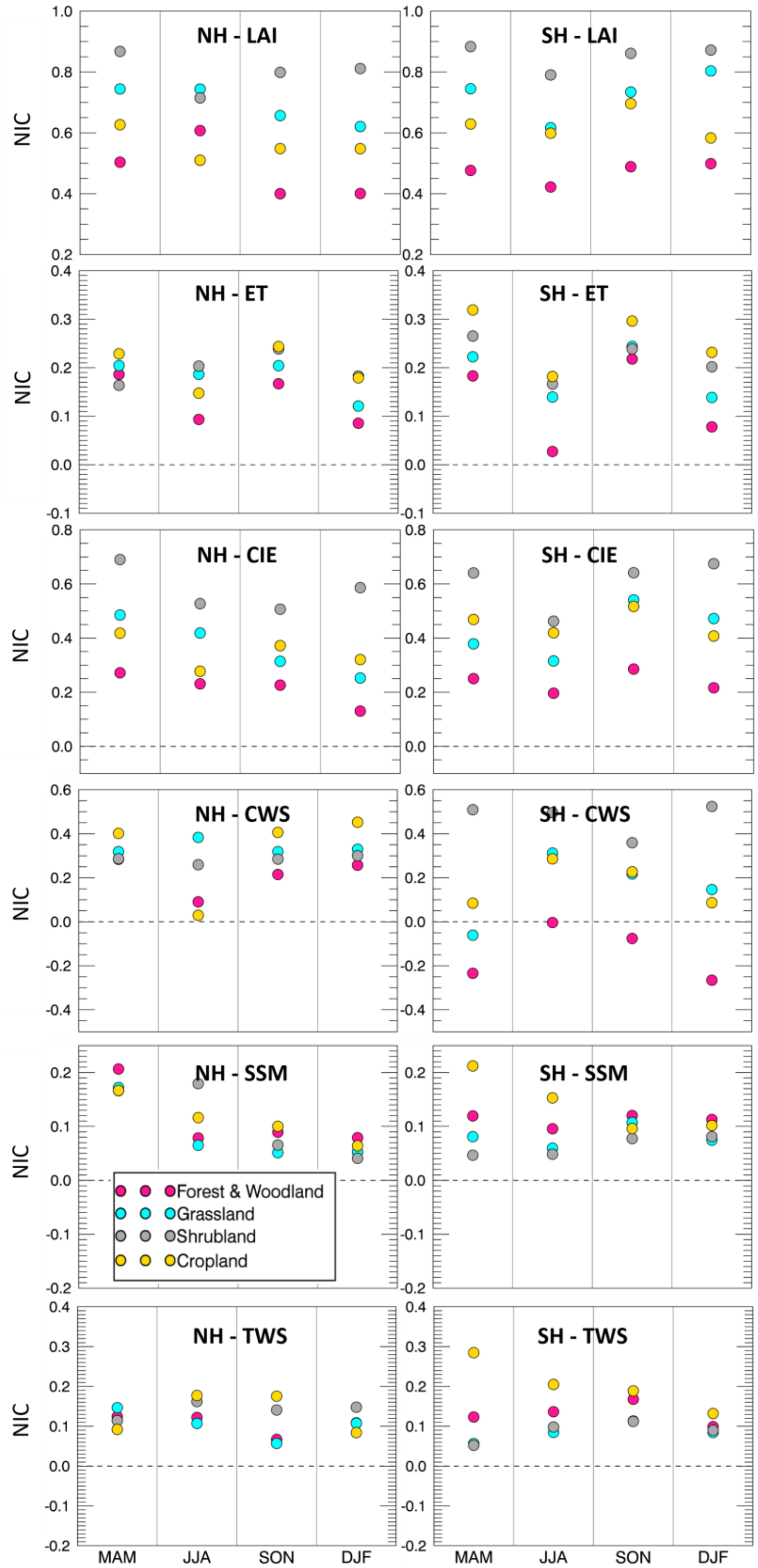
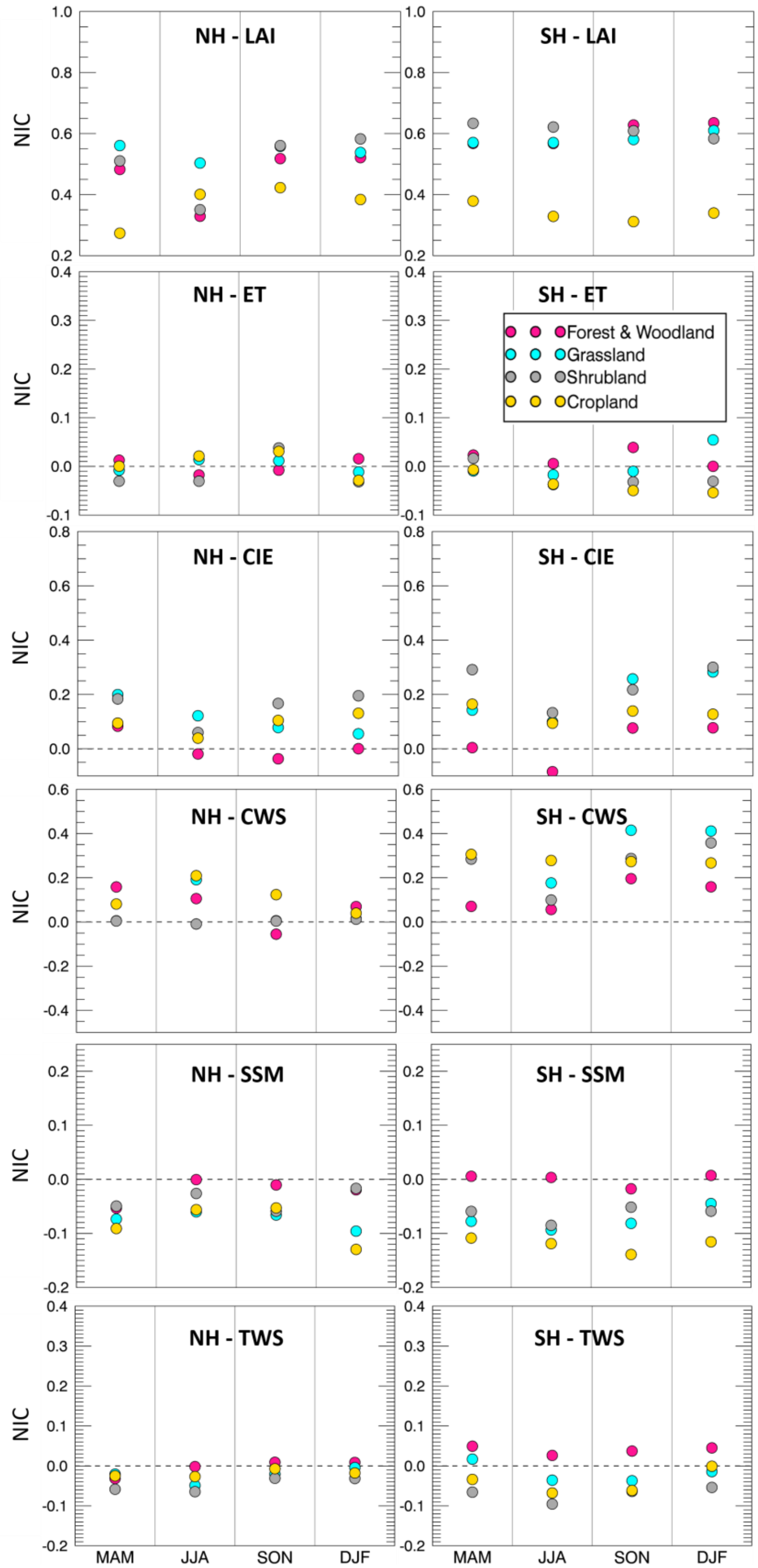


Figure 9. NIC of five water variables under wet precipitation conditions over northern and southern hemispheres (NH and SH) during different seasons (MAM, JJA, SON, and DJF)



334

335 Figure 10. Same as in Figure 9, but for the dry precipitation experiment.

The effectiveness of LAI DA therefore varies across the northern and southern hemispheres, different land cover types, as well as different input precipitation biases. To further investigate the influence of LAI assimilation, Figures 8 and 9 present NIC values for each hemisphere, each season, and each of the input precipitation conditions – wet and dry, respectively. For the wet case (Figure 9), NIC is positive in most cases, which means that the five water variables benefit from the LAI assimilation in all seasons and in both hemispheres. The only exception is CWS which has negative NIC values in the southern hemisphere over grassland (in MAM season) and over forest & woodland (in all seasons). In fact, the forest & woodland region tends to have the least improvement through the LAI assimilation among all land cover types. This is probably because forests and woodlands have large water-holding capacity; thus, the change of water amount caused by LAI DA is not enough to improve the water-related variables. In other words, forest and woodland tend to have lower sensitivity in response to the change of precipitation conditions because of their large rooting depth. On the contrary, cropland is very sensitive to precipitation and it benefits the most from the assimilation of LAI for most of the variables. Moreover, NICs of ET/CIE/CWS tend to be smaller than the NICs of SSM and TWS. There is no clear seasonality in the NIC values, though it has a weak tendency to be lower in warm seasons.

For the dry condition case (Figure 10), NIC values are much lower than in the wet bias case. Nearly half of the NIC values for the five water-related variables are negative, meaning that DA degrades the OL estimates. Nevertheless, the forest & woodland regions tend to perform better than other land covers in dry condition for SSM and TWS. This is due to large soil reservoir of forests and woodlands, which keeps the model water storage more stable when the input precipitation is affected by large negative biases.

3.3. Discussion

As a key factor in land surface processes, precipitation greatly affects surface water fluxes and states and, consequently, affects the vegetation development. Furthermore, changes in vegetation also have considerable impact on the surface water condition. Sections 3.1 and 3.2 quantified changes in five water variables (ET, CIE, CWS, SSM, and TWS) due to the LAI assimilation in Noah-MP. Among the five variables, CIE and CWS are directly related to LAI, while the relationships between LAI and ET, SSM, and TWS are more complex (and indirect) and involve several other factors. For example, ET counts the water losses via both vegetation and soil; SSM is impacted by precipitation, temperature, soil characteristics, etc.; TWS considers all the water storage in the land surface and subsurface, including CWS and SSM.

The performance of the proposed LAI assimilation largely varies depending on the modeled variable, land cover type, errors in the model input (e.g., wet or dry bias in the forcing precipitation), and season. This is due to the complex relationships between vegetation and land water condition. Specifically, results in this study indicate that assimilating LAI in Noah-MP improves the model estimates of water fluxes and storages under positively biased precipitation input, but does not benefit most of the selected water variables when the precipitation input is characterized by a negative bias.

In the dry condition runs, Noah-MP is fed by only half of the original MERRA-2 precipitation used in the NR. Considering that the amount of water in Noah-MP is conservative (since based on a water balance equation), the model has no additional water source in the system, even though the LAI assimilation pushes the model towards more vegetation (that should result in more water). As a matter of fact, introducing more vegetation in the system results in more ET and more root water uptake from the soil, which is most likely the cause for the poor performance of most water fluxes and storages in the DA-dry experiment.

On the other hand, the LAI assimilation is found to improve the original OL runs when the input precipitation is positively biased. This is because LAI assimilation is able to help constrain the partitioning of model water storage when there is abundant water in the system, thus, improving the performance of water-related variables. In summary, although the EnKF is run here in a sub-optimal mode (not satisfying the unbiasedness assumption), the assimilation of LAI is shown to have a positive impact on multiple variables and in several regions of the world.

Overall the improvement of water variables through LAI assimilation is not remarkable enough to compensate the model degradation caused by the biased precipitation forcing data. Previous studies (Pauwels et al. 2007; Sabater et al. 2007; Barbu et al. 2011; Fairbairn et al. 2017; Albergel et al. 2017) have tested the performance of the joint assimilation of LAI and soil moisture over regional domains and showed promising results. However, no experiment was performed at the global scale. Future work could investigate a multi-variate data assimilation system that concurrently merges both LAI and soil moisture (or TWS) observations globally.

4. Conclusions

This study evaluates the efficiency of assimilating vegetation information (i.e., LAI synthetic observations) within a land surface model (Noah-MP 3.6) when the precipitation forcing data are strongly biased (either positively or negatively). Two OSSEs that use an EnKF algorithm for LAI assimilation are performed at global scale during June 2011 – May 2013. The experiments use MERRA-2 as meteorological forcing data. The OL and DA runs are evaluated against a synthetic "truth" from a nature run, in which the MERRA-2 precipitation is neither perturbed nor biased.

The performance of the proposed framework is evaluated for several model output, including LAI estimates and five water-related variables (ET, CIE, CWS, SSM, and TWS).

Overall the EnKF LAI assimilation procedure effectively reduces the LAI error under positively (wet case) and the negatively (dry case) biased precipitation conditions. For the five selected water flux or storage variables, LAI DA improves the model estimates when the model input precipitation is positively biased, but tends to worsen the OL estimates for some of those variables when the input precipitation is negatively biased. Specifically, SSM and TWS estimates are degraded in the DA-dry run with respect to the OL-dry run, while ET, CIE, and CWS do not present large changes when LAI is assimilated in the dry bias run. The poor performance of LAI DA under dry condition is mainly attributed to the fact that the amount of water in Noah-MP is conservative. The LAI assimilation in dry condition introduces more vegetation, which requires more water in the system to replenish the soil water supply. However, the model has no additional source of water, since the input precipitation is negatively biased.

Although a blind bias case (e.g., unknown biases in the precipitation forcing dataset) is presented here in which the EnKF is run in a sub-optimal mode, the assimilation of LAI observations is proven useful to improve several model output variables. Future research should focus on alternative DA methods, such as updating other related model states while assimilating LAI observations, perturbing the model initial condition and model parameters, and/or assimilating actual satellite-based LAI observations (e.g., MODIS, GLASS) at the global scale to verify the efficiency of the proposed vegetation DA framework. This may be particularly useful in agricultural areas, where the vegetation conditions are largely impacted by cropping schedules (Kumar et al. 2019b). Moreover, future work could investigate multi-variate DA techniques that combine the assimilation of several variables (such as LAI, soil moisture, and TWS) at the global scale.

Acknowledgements: This research is sponsored by the NASA Modeling, Analysis, and Prediction (MAP) Program (80NSSC17K0109). We would also like to acknowledge the computational resources and support from the ARGO HPC Cluster team at George Mason University.

References

Adegoke, J. O. and Carleton, A. M.: Relations between soil moisture and satellite vegetation indices in the US Corn Belt, J. Hydrometeorol., 3, 395-405, https://doi.org/10.1175/1525-7541(2002)003<0395:RBSMAS>2.0.CO;2, 2002.

Albergel, C., Munier, S., Leroux, D.J., Dewaele, H., Fairbairn, D., Barbu, A.L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C. and Le Moigne, P.: Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8. 0: LDAS-Monde assessment over the Euro-Mediterranean area, Geosci. Model Dev., 10, 3889-3912, https://doi.org/10.5194/gmd-10-3889-2017, 2017.Andreadis, K. M. and Lettenmaier, D. P.: Assimilating remotely sensed snow observations into a macroscale hydrology model, Adv. Water Resour., 29, 872-886, https://doi.org/10.1016/j.advwatres.2005.08.004, 2006.

Arora, V.: Modeling vegetation as a dynamic component in soil ‑ vegetation ‑ atmosphere transfer schemes and hydrological models, Rev. Geophy., 40, 3-1, https://doi.org/10.1029/2001RG000103, 2002.

Ball, J. T., Woodrow, I. E. and Berry, J. A.: A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in: Progress in photosynthesis research, Springer, Dordrecht, 221-224, https://doi.org/10.1007/978-94-017-0519-6_48, 1987.

Barbu, A. L., Calvet, J. C., Mahfouf, J. F., Albergel, C., and Lafont, S.: Assimilation of Soil Wetness Index and Leaf Area Index into the ISBA-A-gs land surface model: grassland case study, Biogeosciences, 8, 1971-1986, https://doi.org/10.5194/bg-8-1971-2011, 2011.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O. and Roujean, J. L.: LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm, Remote Sens. Environ., 110, 275-286, https://doi.org/10.1016/j.rse.2007.02.018, 2007.

Cohen, W. B. and Justice, C. O.: Validating MODIS terrestrial ecology products: linking in situ and satellite measurements, Remote Sens. Environ., 70, 1-3, 1999.

Cracknell, A. P.: Advanced very high resolution radiometer AVHRR, CRC Press, 543, 1997.

Crow, W. T. and 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., 26, 137-149, https://doi.org/10.1016/S0309-1708(02)00088-X, 2003.

De Lannoy, G.J., Reichle, R.H., Arsenault, K.R., Houser, P.R., Kumar, S., Verhoest, N.E. and Pauwels, V.R.: 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, https://doi.org/10.1029/2011WR010588, 2012

Di, L., Rundquist, D. C. and Han, L.: Modelling relationships between NDVI and precipitation during vegetative growth cycles, Int. J. Remote Sens., 15, 2121-2136, https://doi.org/10.1080/01431169408954231, 1994.

Dickinson, R. E., Shaikh, M., Bryant, R. and Graumlich, L.: Interactive canopies for a climate model, J. Climate, 11(, 2823-2836, https://doi.org/10.1175/1520-0442(1998)011<2823:ICFACM>2.0.CO;2, 1998.

Druel, A., Ciais, P., Krinner, G., and Peylin, P.: Modeling the vegetation dynamics of northern shrubs and mosses in the ORCHIDEE land surface model, J. Adv. Model Earth Sy., 11, 2020-2035, https://doi.org/10.1029/2018MS001531, 2019.

Durand, M. and Margulis, S. A.: Effects of uncertainty magnitude and accuracy on assimilation of multiscale measurements for snowpack characterization, J. Geophys. Res.: Atmos., 113, D02105, https://doi.org/10.1029/2007JD008662, 2008.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean dynam., 53, 343-367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Fairbairn, D., Barbu, A., Napoly, A., Albergel, C., Mahfouf, J. F., and Calvet, J. C.: The effect of satellite-derived surface soil moisture and leaf area index land data assimilation on streamflow simulations over France. HESS, 21, 2015-2033, https://doi.org/10.5194/hess-21-2015-2017, 2017

Farrar, T. J., Nicholson, S. E. and Lare, A. R.: The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. II. NDVI response to soil oisture, Rremote Sens. Environ., 50, 121-133, https://doi.org/10.1016/0034-4257(94)90039-6, 1994.

Fisher, R. A., Koven, C. D., Anderegg, W. R., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J. and Lichstein, J. W.: Vegetation demographics in Earth System Models: A review of progress and priorities, Global Change Biol., 24, 35-54, https://doi.org/10.1111/gcb.13910, 2018.

Foley, J. A., Prentice, I. C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S. and Haxeltine, A.: An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics, Global Biogeochem. Cy., 10, 603-628, https://doi.org/10.1029/96GB02692, 1996.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R. and Wargan, K.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), J. Climate, 30, 5419-5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Ghatak, D., Zaitchik, B., Kumar, S., Matin, M., Bajracharya, B., Hain, C. and Anderson, M.: Influence of Precipitation Forcing Uncertainty on Hydrological Simulations with the NASA South Asia Land Data Assimilation System, Hydrology, 5, 57, https://doi.org/10.3390/hydrology5040057, 2018.

Gibelin, A. L., Calvet, J. C., Roujean, J. L., Jarlan, L. and Los, S. O.: Ability of the land surface model ISBA - A - gs to simulate leaf area index at the global scale: Comparison with satellites products, J. Geophys. Res.: Atmos., 111, D18102, https://doi.org/10.1029/2005JD006691, 2006.

Hansen, M. C., DeFries, R. S., Townshend, J. R. and Sohlberg, R.: Global land cover classification at 1 km spatial resolution using a classification tree approach, Int. J. Remote Sens., 21, 1331-1364, https://doi.org/10.1080/014311600210209, 2000.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P. and Morisette, J. T.: An overview of MODIS Land data processing and product status, Remote Sens. Environ., 83, 3-15, https://doi.org/10.1016/S0034-4257(02)00084-6, 2002.

Kim, Y. and Wang, G.： Impact of vegetation feedback on the response of precipitation to antecedent soil moisture anomalies over North America J. Hydrometeorol., 8, 534-550, https://doi.org/10.1175/JHM612.1, 2007.

Krinner, G., Viovy, N., de Noblet‐Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere‐biosphere system, Global Biogeochem. Cy., 19, GB1015, https://doi.org/10.1029/2003GB002199, 2005.

Kucharik, C. J., Foley, J. A., Delire, C., Fisher, V. A., Coe, M. T., Lenters, J. D., Young‐Molling, C., Ramankutty, N., Norman, J. M. and Gower, S. T.: Testing the performance of a dynamic global ecosystem model: water balance, carbon balance, and vegetation structure, Global Biogeochem. Cy., 14, 795-825, https://doi.org/10.1029/1999GB001138, 2000.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., Lighty, L., Eastman, J. L., Doty, B., Dirmeyer, P. and Adams, J.: Land information system: An interoperable framework for high resolution land surface modeling, Environ. Modell. Softw., 21, 1402-1415, https://doi.org/10.1016/j.envsoft.2005.07.004, 2006.

Kumar, S. V., Peters-Lidard, C., Tian, Y., Reichle, R., Geiger, J., Alonge, C., Eylander, J. and Houser, P.: An integrated hydrologic modeling and data assimilation framework, Computer, 41, 52-59, https://doi.org/10.1109/MC.2008.475, 2008.

Kumar, S. V., Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., Xia, Y., Ek, M., Riggs, G., Livneh, B. and Cosh, M.: Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation, J. Hydrometeorol., 15, 2446-2469, https://doi.org/10.1175/JHM-D-13-0132.1, 2014.

Kumar, S.V., Zaitchik, B.F., Peters-Lidard, C.D., Rodell, M., Reichle, R., Li, B., Jasinski, M., Mocko, D., Getirana, A., De Lannoy, G. and Cosh, M.H.: Assimilation of gridded GRACE terrestrial water storage estimates in the North American Land Data Assimilation System, J. Hydrometeorol., 17, 1951-1972, https://doi.org/10.1175/JHM-D-15-0157.1, 2016

Kumar, S. V., Jasinski, M., Mocko, D. M., Rodell, M., Borak, J., Li, B., Beaudoing, H. K. and Peters-Lidard, C. D.: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment, J. Hydrometeorol., 20, 1571-1593, https://doi.org/10.1175/JHM-D-17-0125.1, 2019.

Kumar, S. V., Mocko, D. M., Wang, S., Peters-Lidard, C. D. and Borak, J.: Assimilation of remotely sensed Leaf Area Index into the Noah-MP land surface model: Impacts on water and carbon fluxes and states over the Continental US, J. Hydrometeorol., 20, 1359-1377, https://doi.org/10.1175/JHM-D-18-0237.1, 2019.

Ling, X. L., Fu, C. B., Guo, W. D. and Yang, Z. L.: Assimilation of remotely sensed LAI into CLM4CN using DART, J. Adv. Model. Earth Sy., https://doi.org/10.1029/2019MS001634, 2019.

Liu, Y., Liu, R. and Chen, J. M.: Retrospective retrieval of long‐term consistent global leaf area index (1981–2011) from combined AVHRR and MODIS data, J. Geophys. Res.: Biogeo., 117, G04003, https://doi.org/10.1029/2012JG002084, 2012.

Liu, Y., Peters‐Lidard, C.D., Kumar, S.V., Arsenault, K.R. and Mocko, D.M.: Blending satellite‐based snow depth products with in situ observations for streamflow predictions in the Upper Colorado River Basin, Water Resour. Res., 51, 1182-1202, https://doi.org/10.1002/2014WR016606, 2015

Morisette, J. T., Privette, J. L. and Justice, C. O.: A framework for the validation of MODIS land products, Remote Sens. Environ., 83, 77-96, https://doi.org/10.1016/S0034-4257(02)00088-3, 2002.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R. and Lotsch, A.: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, Remote Sens. Environ., 83, 214-231, https://doi.org/10.1016/S0034-4257(02)00074-3, 2002.

Niu, G. Y. and Yang, Z. L.: An observation‐based formulation of snow cover fraction and its evaluation over large North American river basins, J. Geophys. Res.: Atmos., 112, D21101, https://doi.org/10.1029/2007JD008674, 2007.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E. and Tewari, M.: The community Noah land surface model with multiparameterization options (Noah‐MP): 1. Model description and evaluation with local‐scale measurements, J. Geophys. Res.: Atmos., 116, D12109, https://doi.org/10.1029/2010JD015139, 2011.

Pagano, T. S. and Durham, R. M.: Moderate resolution imaging spectroradiometer (MODIS), in: Proceedings of SPIE 1939 Sensor Systems for the Early Earth Observing System Platforms, International Society for Optics and Photonics, Orlando, FL, United States, 25 August 1993, 2-17, https://doi.org/10.1117/12.152835, 1993

Pan, M. and Wood, E. F.: Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter, J. Hydrometeorol., 7, 534-547, https://doi.org/10.1175/JHM495.1, 2006.

Pauwels, V. R. and De Lannoy, G. J.: Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, J. Hydrometeorol., 7, 458-477, https://doi.org/10.1175/JHM490.1, 2006.

Pauwels, V. R., Verhoest, N. E., De Lannoy, G. J., Guissard, V., Lucau, C. and Defourny, P.: Optimization of a coupled hydrology–crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter, Water Resour. Res., 43, W04421, https://doi.org/10.1029/2006WR004942, 2007.

Peters-Lidard, C. D., Houser, P. R., Tian, Y., Kumar, S. V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J. and Mitchell, K. High-performance Earth system modeling with NASA/GSFC's Land Information System, Innovations Syst. Softw. Eng., 3, 157-165, https://doi.org/10.1007/s11334-007-0028-x, 2007.

Privette, J. L., Myneni, R. B., Knyazikhin, Y., Mukelabai, M., Roberts, G., Tian, Y., Wang, Y. and Leblanc, S. G.: Early spatial and temporal validation of MODIS LAI product in the Southern Africa Kalahari, Remote Sens. Environ., 83, 232-243, https://doi.org/10.1016/S0034-4257(02)00075-5, 2002.

Reichle, R. H., McLaughlin, D. B. and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, Mon. Weather Rev., 130, 103-114, https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2, 2002.

Reichle, R. H., Walker, J. P., Koster, R. D. and Houser, P. R.: Extended versus ensemble Kalman filtering for land data assimilation, J. Hydrometeorol., 3, 728-740, https://doi.org/10.1175/1525-7541(2002)003<0728:EVEKFF>2.0.CO;2, 2002.

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P., Njoku, E. G. and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR‐E) and the Scanning Multichannel Microwave Radiometer (SMMR), J. Geophys. Res.: Atmos., 112, D09108, https://doi.org/10.1029/2006JD008033, 2007.

Reichle, R. H., Kumar, S. V., Mahanama, S. P., Koster, R. D. and Liu, Q.: Assimilation of satellite-derived skin temperature observations into land surface models, J. Hydrometeorol., 11, 1103-1122, https://doi.org/10.1175/2010JHM1262.1, 2010.

Richard, Y. and Poccard, I.: A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in Southern Africa, Int. J. Remote Sens., 19, 2907-2920, https://doi.org/10.1080/014311698214343, 1998.

Sabater, J. M., Rüdiger, C., Calvet, J. C., Fritz, N., Jarlan, L. and Kerr, Y.: Joint assimilation of surface soil moisture and LAI observations into a land surface model, Agr. Forest Meteorol., 148, 1362-1373, https://doi.org/10.1016/j.agrformet.2008.04.003, 2008.

Tian, Y., Woodcock, C. E., Wang, Y., Privette, J. L., Shabanov, N. V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., Veikkanen, B. and Häme, T.: Multiscale analysis and validation of the MODIS LAI product: I. Uncertainty assessment, Remote Sens. Environ., 83, 414-430, https://doi.org/10.1016/S0034-4257(02)00047-0, 2002.

Wang, G. and Eltahir, E. A.: Role of vegetation dynamics in enhancing the low‐frequency variability of the Sahel rainfall, Water Resour Res., 36, 1013-1021, https://doi.org/10.1029/1999WR900361, 2000.

Wang, G., Sun, S. and Mei, R.: Vegetation dynamics contributes to the multi‐decadal variability of precipitation in the Amazon region, Geophys. Res. Lett., 38, L19703, https://doi.org/10.1029/2011GL049017, 2011.

Woodward, F. I. and Lomas, M. R.: Vegetation dynamics–simulating responses to climatic change, Biol. Rev., 79, 643-670, https://doi.org/10.1017/S1464793103006419, 2004.

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L. and Song, J. Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance, IEEE T. Geosci. Remote, 52, 209-223, https://doi.org/10.1109/TGRS.2013.2237780, 2013.

Yang, Z. L., Niu, G. Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah - MP): 2. Evaluation over global river basins, J. Geophys. Res.: Atmos., 116, D12110, https://doi.org/10.1029/2010JD015140, 2011.

Yoon, Y., Kumar, S.V., Forman, B.A., Zaitchik, B.F., Kwon, Y., Qian, Y., Rupper, S., Maggioni, V., Houser, P., Kirschbaum, D. and Richey, A.: Evaluating the uncertainty of terrestrial water budget components over High Mountain Asia, Front. Earth Sci., 7, https://doi.org/10.3389/feart.2019.00120, 2019

Zhou, Y., McLaughlin, D. and Entekhabi, D.: Assessing the performance of the ensemble Kalman filter for land surface data assimilation, Mon. Weather Rev., 134, 2128-2142, https://doi.org/10.1175/MWR3153.1, 2006.