Assimilation of Remotely Sensed LAI Into CLM4CN Using DART

X. L. Ling1,2,3, C. B. Fu1,2, W. D. Guo1,2, and Z.-L. Yang3,4

1Institute for Climate and Global Change Research and School of Atmospheric Sciences, Nanjing University, Nanjing, China, 2Joint International Research Laboratory of Atmospheric and Earth System Sciences of Ministry of Education, Nanjing, China, 3Department of Geological Sciences, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, Austin, TX, USA, 4CAS Key Laboratory of Regional Climate-Environment Research for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Abstract

Plant leaves play an important role in water, carbon, and energy exchanges between terrestrial ecosystems and atmosphere. Assimilating remotely sensed leaf area index (LAI) into land surface models is a promising approach to improve our understanding of those processes. Toward this goal, this study uses the Community Land Model with carbon and nitrogen components (CLM4CN) coupled with the Data Assimilation Research Testbed (DART). Global Land Surface Satellite (GLASS) LAI data are assimilated via the Ensemble Adjustment Kalman Filter. A random 40-member atmospheric forcing ensemble is used to drive the CLM4CN to provide background error covariance. The results show that assimilating GLASS LAI and updating both LAI and leaf C/N is an effective method to provide a high-accuracy estimate of LAI. The simulations always systematically overestimate LAI, especially in low-latitude regions, with the largest bias up to 5 m²/m², which are effectively corrected in the analyzed LAI, with the bias reduced to ±1 m²/m². Significantly improved regions are located in central Africa, Amazonia, southern Eurasia, northeastern China, and western Europe, where evergreen/deciduous forests and mixed forests are dominant. Except for the temperate zone in the Southern Hemisphere, the analyzed LAI can well represent seasonal variations. The most pronounced assimilation impact in low-latitude regions is attributed to large initial forecast error covariance and sufficient background errors. The MOD 16 evapotranspiration estimates and upscaled gross primary production have been used to evaluate the assimilation impact, which highlight neutral to highly positive improvement.

1. Introduction

Leaf phenology influences the terrestrial energy balance, water budget and carbon cycles (Moore et al., 1995; Richardson et al., 2010). As a critical leaf phenology parameter, leaf area index (LAI) measures the total one-sided area of all leaves in the canopy within a defined region (Chen & Black, 1992). LAI is involved in radiative and ecological processes and hence affects evapotranspiration (ET), runoff, and energy fluxes (Bonan et al., 1992; Monteith & Unsworth, 2014; Pitman, 2003; Richardson et al., 2013).

The fourth-generation land surface models (LSMs) tend to incorporate dynamic ecological, physical, and biogeochemical processes (Bloom & Williams, 2015; Feddema et al., 2005; Hoffman et al., 2017; Ji, 1995; Lawrence et al., 2011; Lawrence & Chase, 2010; Luo et al., 2016; Matheny et al., 2014; Niu et al., 2011; Pitman et al., 2009; Shao & Zeng, 2011; Yang et al., 2011). LAI is prescribed in the second and third generation LSMs, such as Biosphere-atmosphere Transfer Scheme (Dickinson et al., 1993), Simple Biosphere Model (Sellers, 1997; Sellers et al., 1986), Simplified Simple Biosphere Model (Xue et al., 1991), while more advanced models predict LAI (e.g., Calvet et al., 1998, 2004; Dickinson et al., 1998; Cox, 2001; Sitch et al., 2003; Gibelin et al., 2006; Thornton & Zimmermann, 2007). Because of the complexity of the underlying surface, large uncertainties still exist for both the variation amplitude and phase of simulated LAI. Forced by gridded atmospheric data at the regional or global scale, LSMs always require corresponding phenology data (such as LAI and land surface type) as boundary conditions or input parameters.

In situ observations or field experiments can provide highly precise LAI data, but such data are only available at limited locations, which are not sufficient for understanding the interaction of vegetation...
and climate change at regional and global scales. Satellite-derived LAI products have become available on different temporal and spatial scales to represent the global distribution of plant characteristics (Baret et al., 2007, 2013; Knyazikhin et al., 1998; Myneni et al., 2002; Xiao et al., 2014; Yuan et al., 2011).

Replacing default LAI values with remotely sensed data in LSMs can improve the simulation ability, especially for momentum and trace gas exchanges, in various regions, such as East Asia, the western monsoon region of Africa and the temperate forest region of North America [Tian et al., 2004; Kang et al., 2007; Schwartz & Hanes, 2010; Ghilain et al., 2012]. However, satellite-derived data sets cannot be directly incorporated into most LSMs due to the variety of data formats used in various models for different plant functional types. In addition, the imbalance caused by “compulsive” adjustment always causes the model to stop running.

Data assimilation (DA) can merge complementary information from observations and models and can hence obtain optimal estimates of variables. Notably, globally satellite-derived data could provide the data basis for land DA.

Recent studies have shown that assimilating observed or remotely sensed data into LSMs to constrain the vegetation characteristics can improve the simulation ability for terrestrial flux exchanges (Bao et al., 2015; Viskari et al., 2015; Zhao et al., 2013). By assimilating land surface variables (such as albedo and ground temperature) into dynamical models (Jarlan et al., 2008; Li et al., 2014; Liu et al., 2008; Quaife et al., 2008; Wang et al., 2010; Xu et al., 2009), DA has been found to be an effective method for obtaining a time series of high-accuracy LAI data. By assimilating passive microwave observations, Sawada et al. (2015) developed a new land data assimilation system (LDAS) to improve the model skill for surface soil moisture and vegetation dynamics. Liu and Wang (2012) combined observed LAI with dynamical LAI predictions to reconstruct a new advanced LAI data set. By assimilating observed/remotely sensed LAI into the Biome-BGC (BioGeochemical Cycles) model in the Harvard forest region, considerable improvements in the simulation of water and carbon fluxes have been found (Zhang et al., 2013). Rüdiger et al. (2010), Albergel et al. (2010, 2017, 2019), and Barbu et al. (2011) have assimilated satellite-derived vegetation and soil moisture products based on a Simplified Extended Kalman Filter and have found a strong impact on LAI itself, as well as significant improvements for river discharge, land ET, and gross primary production (GPP). Similar studies were conducted in the fields of agriculture (Bao et al., 2015; Dong et al., 2013) and hydrology, and most were conducted at a single site or on regional scales (Fox et al., 2018; Li et al., 2017; Montzka et al., 2012; Pauwels et al., 2007; Sabater et al., 2008).

Aiming at the global scale, the present study attempts to assimilate remotely sensed LAI into the Community Land Model version 4 (CLM4) with explicit carbon and nitrogen components (CLM4CN) using the Data Assimilation Research Testbed (DART) to determine the optimal assimilation scheme and its effectiveness. Detailed descriptions of the dynamic physical model for LAI simulation and of DART are provided in section 2. All data sets utilized for both assimilation and estimation of the analysis impact are also introduced. The experimental design and spin-up process are presented in section 3. Section 4 describes the evaluation and effectiveness of assimilation for a 1-year period. Section 5 provides the main conclusion and future research directions.

2. Model, Data, and Methodology

2.1. CLM4 and CLM4CN

CLM4 (Bonan et al., 2011) is the land component of the Community Earth System Model (CESM1.1; see website http://www.cesm.ucar.edu/models/cesm1.1/; Oleson et al., 2008; Hurrell et al., 2013; Lawrence & Fisher, 2013). The CLM4CN offline mode is used in this study, in which all carbon and nitrogen state variables in vegetation, litter, and soil organic matter are prognostic based on prescribed vegetation phenology. LAI is calculated using the leaf carbon pool and an assumed vertical gradient of specific leaf area (SLA; Thornton & Zimmermann, 2007). Carbon and nitrogen are obtained using plant storage pools in one growing season and are then retained and distributed in the following years.

The prognostic LAI is calculated in CLM4CN as follows:
\[
L = \frac{\text{SLA}_0 \exp(mC_L) - 1}{m}
\]  

where \(L\) is the LAI (\(\text{m}^2/\text{m}^2\)); \(\text{SLA}_0\) is the specific leaf area index at the top of the canopy (\(\text{m}^2\) leaf area g\(^{-1}\) C), which is the ratio of leaf area to leaf mass; \(m\) is a linear slope coefficient based on prescribed plant functional types (PFTs); and \(C_L\) is the leaf carbon (g C m\(^{-2}\) ground area). Parameter \(\text{SLA}_0\) not only links \(L\) and \(C_L\) by providing a structural parameterization but also links \(L\) to GPP for the biochemical parameterization of area-based photosynthetic enzyme concentrations (Reich et al., 1998). As shown in Figure 1, the LAI feedback loop in the prognostic model is represented as follows: LAI influences the GPP by influencing the canopy radiation environment and depends on the canopy-level pool of leaf carbon and nitrogen, while leaf carbon depends on GPP. During the process of allocation, carbon and nitrogen state variables in vegetation, litter, and soil organic material are considered to better simulate the mass and energy exchanges between terrestrial regions and the atmosphere.

2.2. DART and the Ensemble Adjust Kalman Filter (EAKF)

Developed and maintained by the National Center for Atmospheric Research (NCAR), DART is a powerful tool for researchers to explore various DA methods and observations with different high-order numerical models. DART has been coupled with the Community Atmosphere Model (CAM; Raeder et al., 2012), CLM (Kwon et al., 2016; Zhang et al., 2014; Zhao et al., 2016), and many other “high-order” models (Anderson et al., 2009). Furthermore, DART has also been used to improve model prediction through better parameter estimation. For example, Lee et al. (2017) used the Weather Research and Forecasting Model single-column model and DART EnKF to determine the surface roughness length and found that it could reduce the 1-hr forecast errors of wind speed.

The DART system has incorporated several assimilation algorithms, such as the Kalman Filter [KF, Kalman, 1960], EnKF (Evensen, 1994, 2007), EAKF (Anderson, 2001), and particle filter (Moradkhani et al., 2005). The EAKF algorithms are utilized in this paper following Zhang et al. (2014) and Zhao et al. (2016). Compared with traditional EnKF, the EAKF adjusts the gain matrix to avoid the filtering divergence problem by increasing the premise of the analysis error covariance (Anderson, 2001; Talagrand, 1997).

2.3. Data Sets

2.3.1. Ensemble Meteorological Forcing From DART/CAM

Uncertainties in the LAI assimilation with DART/CLM4CN are determined based on atmospheric ensemble reanalysis products, which are obtained by running DART and CAM (DART/CAM4; Raeder et al., 2012). The DART/CAM4 system assimilates observations used in the NCEP-NCAR reanalysis plus radio occultation observations from the Constellation Observing System for Meteorology I onosphere and Climate (Anthes et al., 2008). The CAM4 ensemble reanalysis has been used to force version 2 of the Parallel Ocean Program (Danabasoglu et al., 2012) or CLM4 (Zhang et al., 2014; Zhao et al., 2016), which generated a reasonable ensemble spread and significantly improved the Parallel Ocean Program version 2/CLM analysis.

The DART/CAM system has produced 80 atmospheric forcings, spanning from 1998 to 2010, and contains 6-hr time intervals of air temperature, atmospheric pressure, humidity, wind speed, incoming shortwave/longwave radiation, and solid/liquid precipitation. In this study, 40 ensemble forcings are randomly chosen to drive the DART/CLM4CN because of consideration of both the computational feasibility and EAKF performance (e.g., Reichle et al., 2002; Zhang et al., 2014).

2.3.2. Global Land Surface Satellite (GLASS) LAI

Vast amounts of remotely sensed data have been collected from more than a dozen satellites and ground measurements. From these data, a suite of new inversion algorithms and five GLASS products including LAI have been developed and released (Zhao et al., 2013). The original spatial resolution of 0.05° is linearly interpolated to 0.9° latitude by 1.25° longitude, which is consistent with the resolution of the CLM4CN ensemble simulation. The GLASS LAI data set was provided every 8 days. Recent studies found that GLASS LAI can reasonably represent the global vegetation characteristics and their seasonal variability (Xiao et al., 2017; Yan et al., 2016).
2.3.3. GEOV Version 2 (GEOV2) LAI Data Set for Estimation

The GEOV2 LAI data set from the Copernicus Global Land Service derived from SPOT/VEGETATION and PROBA-V data was used to validate the assimilation result (Verger et al., 2014). The LAI data products are provided every 10 days, and the resolution of the grid is 1/112°. The GEOV2 LAI data in the original resolution are linearly interpolated to a regular grid of 0.9° latitude by 1.25° longitude. Compared with GEOV1 LAI data set, the version 2 data set is close to true value by using multistep filtering method to eliminate the effects of atmosphere and snow (Camacho et al., 2013). Global GEOV2 LAI data set with near real time estimates every 10 days are available at Copernicus portal (http://land.copernicus.eu).

2.3.4. Evaluation Data Sets

The varying 0.5° grids of GPP for the period 1980–2013 are generated by machine learning algorithms from gridded daily air temperature, water availability, radiation, and so on (Tramontana et al., 2016). Six sets of GPP estimates are generated by three machine learning algorithms trained on FLUXNET (Baldocchi, 2008) combined with two partitioning methods (Lasslop et al., 2010; Reichstein, 2005). The data are available in the Max Planck Institute for Biogeochemistry Data Portal (https://www.bgc-jena.mpg.de/geodb/projects/Data.php).

The monthly MOD16 global evapotranspiration (ET) data sets are regular 1-km² covering the time period 2000–2010. The Collection 5 1-km² MOD16 ET algorithms include global land cover type 2 (MOD12Q1; Friedl et al., 2002), FPAR/LAI (MOD15A1; Myneni et al., 2002), as well as MCD53B2 and MCD43B3 albedo (Jin, 2003; Lucht et al., 2000). The MOD16 ET data sets have been estimated using Mu et al. (2007, 2011) and have played critical role in the research of global terrestrial water and energy exchanges. The data sets are available on the website (http://www.ntsg.umd.edu/project/modis/modis16.php).

The Self-Calibrating Palmer Drought Severity Index (SC-PDSI) was utilized (Wells et al., 2004) to study the influence of drought on LAI simulation or assimilation. The SC-PDSI is found to be more spatially comparable than the PDSI for climate division and has been widely used in recent decades. A negative value indicates drought conditions, and a positive value indicates a wet spell. The source code to the SC-PDSI can be downloaded via the National Agricultural Decision Support System (NADSS; online at http://nadss.unl.edu/).

Figure 1. The flowchart of LAI assimilation using coupled DART and CLM4CN (DART/CLM4CN). The black solid line is the Openloop experiment. The black solid line, blue solid line, and red dashed line are the leaf data assimilation experiment which will be described below.
2.4. Incorporating DART Into CLM4CN

In this study, a global LDAS is developed by coupling CLM4CN and DART (DART/CLM4CN). The procedure of LAI assimilation using DART/CLM4CN is illustrated in Figure 1. Forced by reanalysis from DART/CAM, ensemble CLM4CN members stop and write restart files at an interval of 8 days. If there is an observation available, DART will calculate the increments based on model variables, namely, the forecast (or prior), by the EAKF algorithm, and generate an analysis (or posterior). Subsequently, LAI is linearly regressed back into model state space (i.e., leaf C and leaf N), and the updated variables will be sent back to the CLM4CN ensemble restart file, which is then used as the initial state for the next time step integration. This procedure is repeated during the assimilation period.

2.5. Evaluation Strategies

2.5.1. Diagnostic Analysis of Background/Analysis Departures

For diagnostic purposes, the background/analysis departures are utilized as equations (2) and (3) (Barbu et al., 2011):

1. innovations (observations-minus-background), which are calculated as

   \[ d_o^f = y^o - H(x^f) \]  

   where \( d_o^f \) is the background departures, \( y^o \) is the observation, and \( H(x^f) \) is the model simulation;

2. residuals (observations-minus-analysis), which are calculated as

   \[ d_o^a = y^o - H(x^a) \]  

   where \( d_o^a \) is the analysis departures, \( y^o \) is the observation, and \( H(x^a) \) is the assimilated LAI. Albergel et al. (2017) concluded that when the residuals are small compared to the innovations, the LDAS system is working well.

Furthermore, the analysis increments (analysis-minus-background) are calculated as \( d_o^f = H(x^a) - H(x^f) \) to determine the assimilation effectiveness.

2.5.2. Statistical Methods

The impact on LAI itself is evaluated using the relative root-mean-square error (rRMSE) of the ensemble simulation/assimilation versus GEOV2 LAI. Because of large differences in LAI values for different PFTs, the rRMSE are employed to evaluate the performance of both simulation and assimilation. rRMSE is defined as

\[ rRMSE = \frac{RMSE}{\text{Mean} (\text{Observation})} \]  

where Mean (Observation) is the average of GEOV2 LAI, RMSE is the root-mean-square error of the ensemble simulation/assimilation versus GEOV2 LAI.

3. Experimental Setup and the Spin-Up Process

3.1. Experimental Design

Two experiments are designed in this study. The experiment using the default CLM4CN (Openloop) is conducted first. The leaf DA (LDA) experiment are conducted with the GLASS LAI assimilated. DART extracts the state vector and calculates the increments by the EAKF algorithm at a frequency of 8 days. Adjusted LAI, leaf C, and leaf N are sent back to the CLM restart files as initial conditions for further simulation in LDA experiment. LAI is assimilated at the grid level for all vegetation types, as described in Albergel et al. (2017).

During assimilation, no bias correction on the observations is conducted, because the observations are more confidential than model. DART will reject the observation when the bias of prior mean and observation is larger than three times of the expected value, which is defined as \( \sqrt{\sigma_{prior}^2 + \sigma_{obs}^2} \) (in which \( \sigma_{prior} \) and \( \sigma_{obs} \) are standard deviations of prior probability density function (PDF) and observation PDF, respectively). In this study, the EAKF is applied to solve this three-times-limit problem by increasing ensemble model...
errors. Furthermore, all the observations are accepted by DART in the LDA experiments. Considering the computational feasibility and model performance, the CLM4CN is run at a spatial resolution of 0.9° latitude by 1.25° longitude.

### 3.2. Spin-Up Process

Terrestrial modeling with ecosystem components commonly requires a steady-state (SS) solution for all state variables as an initial condition (Thornton et al., 2007; Thornton & Rosenbloom, 2005; Yang & Dickinson, 1995). In the CLM4CN, a long-time spin-up process is also required to allow all pools in the carbon and nitrogen models to reach their steady states (Oleson et al., 2010; Shi et al., 2013). Furthermore, properly adding initial forecast-error covariance can significantly improve the quality of analysis, especially during the initialization phase of assimilation for ensemble filter algorithms (Houtekamer & Houtekamer & Mitchell, 1998; Xu et al., 2006).

To achieve the ensemble initial SS condition with a large forecast-error covariance, a spin-up process that includes three steps is conducted (Figure 2). In the first step, Qian’s forcing (Qian et al., 2006) is used to drive the model to run for 4,000 years. This initial condition is obtained by native dynamics at a spatial resolution of 1.9° latitude by 2.5° longitude (Shi et al., 2013). Next, the ensemble average of 40 atmospheric forcing members for 1998 is used to drive the single CLM4CN to run for 1,000 years. The aim of this step is to make all variables adjust to a relative SS condition, according to the ensemble forcing data produced by DART/CAM. For most biogeochemistry models, the total ecosystem carbon (C_{TOT}) is used as a primary diagnostic variable for the native dynamics simulation. The threshold values of dC_{TOT}/dt in this study is 1.0 g C/m²/y. In the last step, 40 ensemble atmospheric forcing members are used to drive ensemble CLM4CN members to run from 1998 to 2001, and 40 ensemble initial conditions were achieved.

### 4. Results and Discussion

#### 4.1. Impact of the Analysis on Control Variables

The spatial distributions of global LAI in July 2002 for (a) GEOV2 LAI, (b) simulation, and (c) assimilation with C-N constraint (LDA) are shown in Figure 3. These results are verified against the upscaled LAI at grid levels. There are two latitudinal belts of high LAI values located around the tropical area and the 50–65°N region, where BET tropical and NET boreal forests are dominant. Due to the presence of deserts, plateaus and bare ground, LAI is low in western North America, western Australia, southern Africa, and southern South America, where shrubs and/or grasses are dominant. Globally, the CLM4CN can simulate the LAI distribution characteristics (Figure 3b), but it systematically overestimates the global LAI, especially at low latitudes. If both LAI and leaf carbon/nitrogen are updated, the analyzed LAI data are more similar to the GEOV2 LAI data, especially in the low-latitude regions (Figure 3c). As a result, in the LDA experiment, LAI is used to analyze leaf C and N to produce more consistent LAI estimates, and the assimilation results are better in low-latitude regions than in high-latitude regions.

The spatial distributions of global LAI in November 2002 (not shown) exhibit the same characteristics as those shown in Figure 3. The global LAI is systematically lower than that in July. There are three high LAI regions located in the tropics, including the Amazon, central Africa, and some islands in Southeast Asia. The region with large LAI values in the high latitudes of the Northern Hemisphere in July does not
appear in November because November is not the growing season for most of the vegetation in the Northern Hemisphere. Similar to July, the Openloop experiment overestimate global LAI, especially in the low-latitude regions. Globally, the LDA experiment provides a more reasonable LAI distribution.

The global LAI and LAIs in five latitudinal bands have been studied in this section based on climatic background divisions. The five bands are the boreal zone (40°–65°N), the temperate zones in the northern and southern hemisphere (23°–45°N and 23°–90°S), and the northern and southern equatorial zones (0°–23°N and 0°–23°S).

There are three ways for improving the model performance by assimilating observations into the CLM (Fox et al., 2018). The first and most direct way is to update modeled LAI based on the calculated Kalman gain. According to equation (1), the LAI value in the previous period will not directly affect LAI changes in the next period. The second way is to update the other model variables (in this study the relative variables are leaf C and leaf N) extracted from the restart files. In the ecological system, many variables and linking process are constrained by growth forms established by evolutionary or ecological process and allometric relationships, which is also embodied in land surface models (for CLM4 can refer to equation (1)). Leaf C and leaf N are updated by the EAKF based on their correlation with the LAI, which is the key step to transfer information from LAI to other relative variables via statistical method. The model allocation rules of ecological process take part in the LDA experiments, indicating that we add a carbon and nitrogen constrain. The strong correlation between the LAI and leaf C and leaf N lead to an updated ensemble initial conditions for the next time step. The LAI RMSE for Openloop run at global scale is 0.21 and reduced to 0.08 after assimilation (see Table 1). The most obviously improved regions for LAI RMSE are the Northern and Southern Equatorial regions, with the RMSEs of 0.55/0.65 reduced to 0.12/0.20, respectively. With the change of LAI, leaf C and leaf N vary accordingly. At global scale, −37.5% reduction for LAI relative to Openloop experiment will result in equivalent reductions associated with leaf C and leaf N (33% and 38%, respectively). The variation of leaf N respond to LAI change a little stronger than leaf C. The third way is indirect that the other modelled variables may also change after assimilation through biophysical processes and feedbacks in the model. Take land surface energy balance for example, the analyzed GPP and ET will be discussed in section 4.4.

Figure 4 illustrates the averaged LAI analysis increments (analysis—Openloop) for the months of (a) January, (b) April, (c) July, and (d)
October 2002. At the very beginning of the year (Figure 4a), assimilation is not very active in the higher latitude regions of both the Northern and Southern Hemispheres. The analyzed LAIs tend to increase over the western United States, Southeast Asia, the eastern Amazon, and western Australia. The analyzed LAIs decrease significantly in the lower regions, especially in the northern and central Amazon, central Africa, and southern Asia. With more assimilation steps, the assimilation is more efficient, especially in the 55–65°N regions, which are covered with deciduous needleleaf or mixed forests. The marked impact of assimilation for the higher-latitude region of the Northern Hemisphere shows a strong seasonal dependency.

Probability density function of innovation and residuals of LAI for global and all subregions during July 2002 are illustrated in Figure 5. The distribution of the residuals is more centered on 0 than that of the innovations. Table 2 also displays the proportion of the background/analysis departures within 50% of the averaged observations. The percentage of analysis departures is improved from background departures by 4.2%, 2.6%, 0.8%, 18.6%, 4.2%, and 1.7% in global, boreal, northern temperate, northern equatorial, southern equatorial, and southern temperate regions, respectively. Both innovations and residuals dominantly exhibit large negative biases, indicating that the model always highly overestimates LAI, although assimilation can correct this bias. The similarity between innovations and residuals in the northern temperate region (Figure 5c) in July implies that the assimilation is not very efficient. The analysis is more efficient for the left-tail distribution than for the right, indicating that overestimation of LAI is beneficial for assimilation. The same issues apply in November (not shown).

To determine where assimilated LAI is better than the simulation and the corresponding reason, Figure 6 shows the absolute differences in innovation and residuals for (a) improved regions and (b) not improved regions and GEOV2 LAI with values (c) in the range of 0–1.5 m²/m² and (d) larger than 1.5 m²/m². The regions improved by LAI analysis are located in the Amazon, southeastern South America, central Africa, southeastern North America, southern and eastern China, and western Europe. The improved regions are also the regions with LAI values larger than 1.5 m²/m². Conversely, the not improved regions are southwestern North America, central South America, southwestern Europe, and western Australia (Figure 6b). Compared with Figure 6b, the regions with LAI values less than 1.5 m²/m² (Figure 6c) show the same distribution characteristics. In conclusion, the assimilation effect is not satisfactory when the simulated LAI is less than 1.5 m²/m², where the dominant vegetation is moist tundra or temperate savanna, which have averaged LAI values of 0.82 and 1.37 m²/m², respectively (Asner et al., 2003). The low standard deviation of such vegetation types with low LAI values provides a small background error matrix, which may be the reason for
the poor assimilation performance. Furthermore, the good assimilation performance associated with LAI values larger than 1.5 m²/m² may be attributed to the larger standard deviation from the simulation, implying poor modeling performance for dense plants.

As showed in Figure 7, the CLM4CN fail to capture the magnitude and seasonality of LAI globally, especially in the Boreal and Southern Temperate regions. The seasonal patterns of assimilation are more consistent with the simulation than the observations, implying that the dynamic forwarding model has more of an impact on the variation characteristics. In the boreal zone, the annual variation in the observed LAI is much larger than that in the simulation/assimilation, and the minimum deviation occurs in July. The biases in the simulation and assimilation results vary within the ranges of 0.4–1.4 and 0.1–1.0 m²/m², respectively (Figure 7b). In the northern temperate zone, the annual LAI patterns from the simulation and assimilation are consistent with those from observations, with averaged biases of approximately 1.0 and 0.5 m²/m², respectively (Figure 7c). The simulated LAIs in the northern and southern equatorial regions are 1.9 and 2.7 m²/m², respectively, both of which are higher than those of the observations. With assimilation, the biases in the two regions can be reduced to 0.4 m²/m² (Figures 7d and 7e). In the southern temperate region, the simulated LAI is approximately 0.5–0.8 m²/m² smaller than that in the observations, and the model cannot reproduce the characteristic seasonal patterns of LAI. Because these regions are largely covered by savanna, shrubs and/or grass, the above results imply that the model performance is poor for simulating the LAI values for these PFTs. This pattern could possibly be because the dynamic vegetation model overestimates carbon fixation and/or allocation of biomass to leaves. Furthermore, DA can also compensate for missing process in the model. For example, the assimilated LAI also showed an improvement in the decaying period, partly due to the added information on harvesting. This result illustrates the utility of satellite data-based products by highlighting an area of research that requires more refined models.

**Figure 5.** Probability density function of innovation (observations—openloop, in blue) and residuals (observations—analysis, in red) for LAI for (a) global, (b) boreal (45–65°N), (c) northern temperate (23–45°N), (d) northern equatorial (0–23°N), (e) southern equatorial (0–23°S), and (f) southern temperate (23–90°S) in July 2002.
To further evaluate the assimilation impacts, Figure 8 shows the evolutions of the monthly forecast (prior)/analyzed (posterior) rRMSE of LAI in the latitudinal band for simulation/assimilation and observations. Because of large differences in LAI values for different PFTs, the rRMSE technique is employed to evaluate the performance of simulation/assimilation. The rRMSE results for the assimilation are lower than those for the simulation except for the southern temperate zone. In the regions with obvious seasonal variation, the rRMSEs change in the opposite direction as that of LAI, that is, the higher the observed LAI is, the lower the RMSE. Globally, the rRMSEs of observations are less than 1.0 (during the growing season), implying that the observations exhibit good agreement with the mean of the observations. The rRMSEs of the simulation results are approximately 1.0 or higher than 1.0, indicating that the RMSE magnitude of simulation is comparable to the RMSE magnitude of the observations. The largest rRMSE appears during the fall leaf senescence period for boreal forests, implying that fall phenology is difficult to predict without actively incorporating information from observations. Better assimilation results can be found in the low-latitude regions with high LAI observations, and the rRMSEs are 40% lower throughout the year.

4.2. Evaluation of Analysis Impact

To analyze the relationship of LAI assimilation with wet spell or drought conditions, Figure 9 shows the spatial distribution of SC-PDSI in July 2002. Only values greater than 3 (implying severe and extreme humidity) and less than −3 (implying severe and extreme aridity) are displayed. Compared with Figure 6, the not improved regions for LAI are consistent with the severe drought and extreme drought regions, especially in the midlatitude region of the North Hemisphere and Australia, which feature open shrubland, barren or sparsely vegetated land, and grassland. In conclusion, LAI assimilation is not sensitive to severe drought and extreme drought, partly because the low corrected LAI values cannot contribute much to the modeling improvement.

GPP from both the Openloop and the LDA experiment are compared to the monthly GPP estimates in July 2002. Simulated GPP tends to overestimate the monthly GPP estimates in the low-latitude regions, especially over the Amazon, central Africa, and the southwest of Asia. Over high-latitude regions, simulated GPP is

| SUBREGIONS          | Innovations | Residuals |
|---------------------|-------------|-----------|
| Global              | 44.09%      | 48.25%    |
| Boreal              | 56.19%      | 58.81%    |
| Northern Temperate  | 41.91%      | 42.74%    |
| Northern Equatorial | 10.46%      | 29.09%    |
| Southern Equatorial | 33.55%      | 37.72%    |
| Southern Temperate  | 21.05%      | 22.74%    |

To further evaluate the assimilation impacts, Figure 8 shows the evolutions of the monthly forecast (prior)/analyzed (posterior) rRMSE of LAI in the latitudinal band for simulation/assimilation and observations. Because of large differences in LAI values for different PFTs, the rRMSE technique is employed to evaluate the performance of simulation/assimilation. The rRMSE results for the assimilation are lower than those for the simulation except for the southern temperate zone. In the regions with obvious seasonal variation, the rRMSEs change in the opposite direction as that of LAI, that is, the higher the observed LAI is, the lower the RMSE. Globally, the rRMSEs of observations are less than 1.0 (during the growing season), implying that the observations exhibit good agreement with the mean of the observations. The rRMSEs of the simulation results are approximately 1.0 or higher than 1.0, indicating that the RMSE magnitude of simulation is comparable to the RMSE magnitude of the observations. The largest rRMSE appears during the fall leaf senescence period for boreal forests, implying that fall phenology is difficult to predict without actively incorporating information from observations. Better assimilation results can be found in the low-latitude regions with high LAI observations, and the rRMSEs are 40% lower throughout the year.

4.2. Evaluation of Analysis Impact

To analyze the relationship of LAI assimilation with wet spell or drought conditions, Figure 9 shows the spatial distribution of SC-PDSI in July 2002. Only values greater than 3 (implying severe and extreme humidity) and less than −3 (implying severe and extreme aridity) are displayed. Compared with Figure 6, the not improved regions for LAI are consistent with the severe drought and extreme drought regions, especially in the midlatitude region of the North Hemisphere and Australia, which feature open shrubland, barren or sparsely vegetated land, and grassland. In conclusion, LAI assimilation is not sensitive to severe drought and extreme drought, partly because the low corrected LAI values cannot contribute much to the modeling improvement.

GPP from both the Openloop and the LDA experiment are compared to the monthly GPP estimates in July 2002. Simulated GPP tends to overestimate the monthly GPP estimates in the low-latitude regions, especially over the Amazon, central Africa, and the southwest of Asia. Over high-latitude regions, simulated GPP is
underestimated, particularly over the 55–65°N regions covered by NET forest or mixed forests. The analyzed GPP can reduce these biases as showed in Figure 10b. ET from both the Openloop and the LDA experiment are compared to the MODIS ET estimates in July 2002. The assimilation impact on ET is small comparing with GPP. However, the comparison with the MODIS data is rather positive. Figure 11 shows the ET from (a) MODIS estimates, (b) Difference between LDA and Openloop experiments, (c) Openloop, and (d) analysis. Simulated ET tends to underestimate the MODIS estimates over northern and central Amazon, southern Africa, and northern North America. Analysis is able to reduce this bias as showed in Figure 11b. The improvement for the analyzed GPP and ET suggests that the model should be improved through enhancing key state variables. On the other hand, the bias should be included in the analysis system (De Lannoy et al., 2007).

4.3. Discussion

This study is conducted based on Ensemble Adjust Kalman Filter. Although DART have provided accesses to multiple assimilation algorithms (e.g., the EnKF and particle filter), the EAKF algorithm is the most mature technique developed for land surface DA within DART/CLM. Many studies using DART coupling with CLM have used the EAKF algorithms (Fox et al., 2018; Zhang et al., 2014; Zhao et al., 2016), which

Figure 7. Comparison of seasonal cycles among simulation, assimilation, and observations for (a) global, (b) boreal (45–65°N), (c) northern temperate (23–45°N), (d) northern equatorial (0–23°N), (e) southern equatorial (0–23°S), and (f) southern temperate (23°–90°S).
is a fully deterministic algorithm for estimating model forecast error statistics based on observation uncertainty (Anderson, 2001). Comparison with other algorithms should be conducted in the future work, for example, particle filters may provide a means to capture non-Gaussian errors (Moradkhani et al., 2012).

The effectiveness of LAI assimilation varies with time and space because of different plant functional types. This work is implemented pointwise, indicating that we do not consider spatial covariances here. Localization algorithms have not been implemented in this study. As described in Anderson (2007, 2012), the localization can ameliorate sampling error when using small ensemble sizes to sample the statistical relationship between observations and state variables and to prevent spurious updates when variables are known a priori to be unrelated. We have a relatively large ensemble number in this study, while the ensemble models are run at global scale, so localization should be considered in the near future.

The length of the assimilation interval was chosen to be 8 days, which is the same with the temporal resolution of GLASS LAI. Rüdiger et al. (2010) found that by jointly assimilating LAI and other observations (e.g., soil moisture) at an interval of 1 day would increase Jacobian values and enhance the model’s response at the end of the interval to the initial perturbations of the model states. While the global LAI estimates are

---

Figure 8. The same comparison as in Figure 7, but for the rRMSE of simulation/assimilation versus GEOV2 LAI.
only provided every 8 or 10 days, the assimilation uncertainty should be reflected in the observation error specification. Furthermore, Viskari et al. (2015) found that increasing the frequency of the observations would also increase the observation noise.

The DA performance is mostly influenced by appropriate statistics for background errors. The background error covariance is provided as driving by a random 40-member atmospheric forcing ensemble. Although some studies have found that if perturb to the initial conditions and/or model parameters it will tend to damp the atmosphere variability, many other existed studies have been conducted by only perturbing

Figure 9. Spatial distribution of SC-PDSI in July 2002, in which only values greater than 3 (implying severe and extreme humidity) and less than −3 (implying severe and extreme aridity) are shown.

Figure 10. GPP from the (a) MODIS estimates, (b) difference between LDA and Openloop experiments, (c) Openloop, and the (d) analysis.
the atmospheric forcing (Zhang et al., 2014; Zhao et al., 2016). Furthermore, the spin-up process (especially step 3) was designed to achieve the 40 ensemble initial conditions, providing perturbation for the initial conditions to realize the assimilation.

Ensemble initializations play a remarkable role during assimilation because they provide an isotropic and homogeneous error for the background fields. The high discrete degree for BET tropical, C4 grass, and crop types demonstrates that these vegetation types are sensitive to meteorological forcing. Except for the growing seasons, the variation ranges are also large during the leaf-out and leaf-senescence periods for the BDS temperate forest type. The DA performance is improved when the ensemble models diverge with respond to the meteorological forcing in this study and Fox et al. (2018).

Many other studies have been conducted by assimilating LAI, soil moisture, and biomass into the land surface models at single site or regional scale. Significant improvements were obtained for LAI itself, as well as GPP, ET, and relative hydrologic variables. Our results indicate that the assimilation of LAI can compensate for missing process in the model and then further correct the deficiency in the model. Take the agriculture for example, DA can add harvesting information to the model for the crop type especially for the decaying period. This also suggests that the model performance can be improved through enhanced improvement for the key parameters.

5. Conclusions

In this study, CLM4CN is linked with DART as a new LDAS by assimilating satellite-derived LAI with carbon and nitrogen constraints. The 40 randomly chosen ensemble atmospheric reanalysis data sets generated by DART/CAM are employed to introduce uncertainties. GLASS LAI are assimilated into the CLM4CN at a frequency of 8 days, and the spatial resolution is 0.9° × 1.25°.

The Openloop experiment (without assimilation) and LDA experiment are designed to determine whether carbon and nitrogen constraints impact assimilation. The results show that if both LAI and leaf C/N are updated, the analyzed LAI can be significantly improved, especially in low-latitude regions. The assimilation impacts of the LAI are influenced by different vegetation. In low-latitude regions covered by many forests, assimilation can significantly correct the LAI overestimation. The highly improved low-latitude regions for the analyzed LAI are due to large initial forecast error covariances and large background errors.
namely, the standard deviation of 40 ensemble members. For the regions in middle latitudes covered by grass, shrubs and savanna, the results of the assimilation are still far from satisfactory for the low standard deviation of both the initial conditions and ensemble spread. Less improvement is found in November than in July, implying a seasonal dependence for DA. The significant difference in performance between July and November is found in the 50–65°N region, implying that vegetation growth and deciduous processes are another issue to be considered during LAI DA.

The not improved regions for LAI simulation are consistent with severe drought and extreme drought regions, especially in the mid-latitude areas of the Northern Hemisphere and Australia, which are covered with open shrubland, barren or sparsely vegetated land, and grassland. Thus, LAI assimilation is not sensitive to severe drought and extreme drought, partly because the low corrected LAI values cannot contribute much to modeling improvement. Furthermore, the GPP and ET have been improved after LAI assimilation, especially for the low latitude regions.

The LAI DA in this study is conducted at grid levels for all vegetation types, which may cause poor DA results for grids with multiple vegetation types (e.g., the transitional zone between forests and grass). As described in Barbu et al. (2014), the analysis increments are calculated for each individual vegetation type. The future direction of this research must focus on considering the LAI assimilation for each PFT. Furthermore, improving DA effectiveness in nongrowing seasons for low-growing vegetation is necessary.

References

Albergel, C., Calvet, J. C., Mahfouf, J. F., Rüdiger, C., Barbu, A. L., Lafont, S., et al. (2010). Monitoring of water and carbon fluxes using a land data assimilation system: A case study for southwestern France. Hydrology and Earth System Sciences, 14(6), 1109–1124. https://doi.org/10.5194/hess-14-1109-2010

Albergel, C., Dutra, E., Bonan, B., Zheng, Y., Munier, S., Balsamo, G., et al. (2019). Monitoring and forecasting the impact of the 2018 summer heatwave on vegetation. Remote Sensing, 1(3), 520. https://doi.org/10.3390/rs11050520

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Ector, D., Barbu, A. L., et al. (2017). Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX, v8.0: LDAS-Monde assessment over the Euro-Mediterranean area. Geoscientific Model Development, 10(10), 3889–3912. https://doi.org/10.5194/gmd-10-3889-2017

Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. Monthly Weather Review, 129(12), 2884–2903. https://doi.org/10.1175/1520-0493(2001)129<2884:EAKF>2.0.CO;2

Anderson, J. L. (2007). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. Physica D: Nonlinear Phenomena, 239(2), 99–111. https://doi.org/10.1016/j.physd.2006.02.011

Anderson, J. L. (2012). Localization and sampling error correction in ensemble Kalman filter data assimilation. Monthly Weather Review, 140(7), 2359–2371. https://doi.org/10.1175/MWRD-11-00013.1

Anderson, J. L., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Arrellano, A. (2009). The data assimilation research testbed: A common facility. Bulletin of American Meteorological Society, 90(9), 1283–1296. https://doi.org/10.1175/2009BAMS2618.1

Anesio, A. M., Bernardi, P. A., Chen, Y., Cucurull, L., Dymond, K. F., Ector, D., et al. (2008). The COSMIC/FORMOSAT-3 mission: Early results. Bulletin of American Meteorological Society, 89(3), 313–334. https://doi.org/10.1175/BAMS-89-3-313

Asner, G. P., Scurlock, J. M. O., & Hicke, J. A. (2003). Global synthesis of leaf area index observations: Implications for ecological and remote sensing studies. Global Ecology and Biogeography, 12(3), 191–205. https://doi.org/10.1046/j.1466-822X.2003.00026.x

Baldocchi, D. (2008). Turner Review No. 15. ‘Breathing’ of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems. Australian Journal of Botany, 56(1), 1–26. https://doi.org/10.1071/BT07151

Bao, S. N., Cao, C. X., Huang, J. X., Ma, H. Y., Tian, L. Y., Su, W., & Ni, X. L. (2015). Remote sensing of vegetation coverage using a land surface model: Grassland case study. Remote Sensing of Environment, 170, 377–390. https://doi.org/10.1016/j.rse.2015.07.015

Barbu, A. L., Calvet, J.-C., Mahfouf, J.-F., Albergel, C., & Lafont, S. (2011). Assimilation of Soil Wetness Index and Leaf Area Index into the ISBA-A-gs land surface model: Grassland case study. Biogeosciences, 8(7), 1971–1986. https://doi.org/10.5194/bg-8-1971-2011

Barbu, A. L., Calvet, J.-C., Mahfouf, J.-F., & Lafont, S. (2014). Integrating ascat surface soil moisture and geol leaf area index into the surfex modelling platform: A land data assimilation application over france. Hydrology and Earth System Sciences, 18(1), 173–192. http://doi.org/10.5194/hess-18-173-2014

Baret, F., Hagle, O., Geiger, B., Bicheron, P., Miras, B., Hue, M., et al. (2007). LAI and FAPAR CYCLOPS global products derived from VEGETATION. Part 1: Principles of the algorithm. Remote Sensing of Environment, 110(3), 275–286. https://doi.org/10.1016/j.rse.2007.02.018

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhrma, H., Pacholczyk, P., & Smets, B. (2013). GLOVE: LAI and FAPAR essential variables and FCOVER global time series capitalizing over existing products. Part 1: Principles of development and production. Remote Sensing of Environment, 137, 299–309. https://doi.org/10.1016/j.rse.2012.12.027

Bloom, A. A., & Williams, M. (2015). Constraining ecosystem carbon dynamics in a data-limited world: Integrating ecological “common sense” in a model-data fusion framework. Biogeosciences, 12(5), 1299–1315. https://doi.org/10.5194/bg-12-1299-2015

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., et al. (2011). Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. Journal of Geophysical Research, 116, G02014. https://doi.org/10.1029/2010JG001593

Bonan, G. B., Pollard, D., & Thompson, S. L. (1992). Effects of boreal forest vegetation on global climate. Nature, 359(6397), 716–718. https://doi.org/10.1038/359716a0
Calvet, J.-C., Noilhan, J., Roujean, J.-L., Bessemoulin, P., Cabelguenne, M., Olioso, A., & Wigneron, J.-P. (1998). An interactive vegetation SVAT model tested against data from six 780 contrasting sites. *Agriculture and Forest Meteorology*, 92(2), 73–95. https://doi.org/10.1016/s0168-1923(98)00091-4

Calvet, J.-C., Rivalsand, V., Picon-Cochard, C., & Guehl, J.-M. (2004). Modelling forest transpiration and CO2 fluxes–Response to soil moisture stress. *Agriculture and Forest Meteorology*, 124(3-4), 143–156. https://doi.org/10.1016/j.agrformet.2004.01.007

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & Weiss, M. (2013). GEOVI: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment*, 137, 310–329. https://doi.org/10.1016/j.rse.2013.02.030

Chen, J. M., & Black, T. A. (1992). Defining leaf area index for non-flat leaves. *Plant Cell and Environment*, 15(4), 421–429. https://doi.org/10.1111/j.1365-3040.1992.tb00992.x

Cox, P. M. (2001). Description of the TRIFFID dynamic global vegetation model. Technical Note 24, Hadley Centre, United Kingdom Meteorological Office, Bracknell, UK.

Danabasoglu, G., Bates, S. C., Briegleb, B. P., Jayne, S. R., Jochum, M., Large, W. G., et al. (2012). The CCSM4 ocean component. *Journal of Climate*, 25(5), 1361–1389. https://doi.org/10.1175/JCLI-D-11-00091.1

De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., & Verhoest, N. E. C. (2007). State and bias estimation for soil moisture profiles by an ensemble Kalman filter. *Water Resource Research*, 43, W06401. https://doi.org/10.1029/2006WR005100

Dickinson, R. E., Sellers, A. H., Kennedy, P. J. (1993). Biosphere-atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model. Technical note. Boulder, Colorado: National Center for Atmospheric Research.

Dickinson, R. E., Shakhun, M., Bryant, R., & Graumlich, L. (1998). Interactive canopies for a climate model. *Journal of Climate*, 11, 2823–2836. https://doi.org/10.1175/1520-0442(1998)11<2823:ICFACM>2.0.CO;2

Dong, Y. Y., Zhao, C. J., Yang, G. I., Chen, L. P., Wang, J. H., & Peng, H. K. (2013). Integrating a very fast simulated annealing optimization algorithm for crop leaf area index variational assimilation. *Mathematical and Computer Modelling*, 58(3-4), 877–885. https://doi.org/10.1016/j.mcm.2012.12.013

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5), 10,143–10,162. https://doi.org/10.1029/94jc00572

Evensen, G. (2007). *Data Assimilation, the Ensemble Kalman Filter*. Berlin, Heidelberg: Springer.

Feddema, J. J., Oleson, K. W., Bonan, G. B., Mearns, L. O., Buja, L. E., Meehl, G. A., & Washington, W. M. (2005). The importance of land cover change in simulating future climates. *Science*, 310(5754), 1674–1678. https://doi.org/10.1126/science.1118160

Fox, A. M., Hoar, T. J., Anderson, J. L., Arelano, A. F., Smith, W. K., Litvak, M. E., et al. (2018). Evaluation of a data assimilation system for land surface models using CLM4.5. *Journal of Advances in Modeling Earth Systems*, 10(10), 2471–2494. https://doi.org/10.1002/2018MS001362

Friedl, M. A., McVier, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS. Algorithms and early results. *Remote Sensing of Environment*, 83(1-2), 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0

Ghilain, N., Arboleda, A., Sepulcre-Cantó, G., Batelaan, G., Ardö, J., & Gellens-Meuleman, F. (2012). Improving evapotranspiration in a land surface model using biophysical variables derived from MSG/SEVIRI satellite. *Hydrology and Earth System Science*, 16(8), 2567–2583. https://doi.org/10.5194/hess-16-2567-2012

Gibelin, A.-L., Calvet, J.-C., Roujean, J.-L., Jarlan, L., & Los, S. O. (2006). Ability of the land surface model ISBA-Ags to simulate leaf area index at global scale: Comparison with satellite products. *Journal of Geophysical Research*, 111, D18102. https://doi.org/10.1029/2005JD006691

Hoffman, F. M., Koven, C. D., Kepp-Aleks, G., Lawrence, D. M., Riley, W. J., Randerson, J. T., et al. (2017). 2016 International Land Model Benchmarking (ILAMB) Workshop Report (No. DOE/SC). USDOE Office of Science (SC) (United States). https://doi.org/10.2172/1330803

Houtekamer, P. L., & Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126, 796–811. https://doi.org/10.1175/1520-0493

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The community Earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, 94(9), 1339–1360. https://doi.org/10.1175/BAMS-D-12-00121.1

Jarlan, L., Balsamo, G., Lafont, S., Beljaars, A., Calvet, J.-C., & Mougin, E. (2008). Analysis of leaf area index in the ECMWF land surface model and impact on latent heat and carbon fluxes: application to West Africa. *Journal of Geophysical Research*, 113, D24117. https://doi.org/10.1029/2007JD009370

Ji, J. J. (1995). A climate-vegetation interaction model, simulating physical and biological processes at the surface. *Journal of Biogeography*, 22(2/3), 445–451. https://doi.org/10.2307/2845941

Jin, Y. (2003). Consistency of MODIS surface BRDF/Albedo retrievals: 2. Validation. *Journal of Geophysical Research*, 108(DS), 4159. https://doi.org/10.1029/2002JD002804

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering Transactions*, 82(1), 35–45. https://doi.org/10.1115/1.3662552

Kang, H. S., Xue, Y. K., & Collatz, G. J. (2007). Impact assessment of satellite-derived leaf area index datasets using a general circulation model. *Journal of Climate*, 20(6), 993–1015. https://doi.org/10.1175/JCLI4054.1

Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Diner, D. J., & Running, S. W. (1998). Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research*, 103(D24), 32257–32275. https://doi.org/10.1029/98jd2462

Kwon, Y., Yang, Z. L., Zhao, L., Hoar, T. J., Toure, A. M., & Rodell, M. (2016). Estimating snow water storage in North America Using CLM4, DART and snow radiation data assimilation. *Journal of Hydrometeorology*, 17(11), 2853–2874. https://doi.org/10.1175/JHM-D-16-0028.1

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arndt, A., Barr, A., et al. (2010). Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. *Global Change Biology*, 16(1), 187–208. https://doi.org/10.1111/j.1365-2486.2009.02041.x

Lawrence, D., & Fisher, R. A. (2013). The Community Land Model philosophy: Model development and science applications. *Leaps Newsletter*, 13, 13–16.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., et al. (2011). Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, 3, M03001. https://doi.org/10.1029/2011MS000045
Zhao, L., Yang, Z. L., & Hoar, T. J. (2016). Global soil moisture estimation by assimilating AMSR-E brightness temperatures in a coupled CLM4-RTM-DART system. *Journal of Hydrometeorology*, 17(9), 2431–2454. https://doi.org/10.1175/JHM-D-15-0218.1

Zhao, X., Liang, S., Liu, S., Yuan, W., Xiao, Z., Liu, Q., et al. (2013). The Global Land Surface Satellite (GLASS) remote sensing data processing system and products. *Remote Sensing*, 5(5), 2436–2450. https://doi.org/10.3390/rs5052436

Zhao, Y. X., Chen, S. N., & Shen, S. H. (2013). Assimilating remote sensing information with crop model using ensemble Kalman filter for improving LAI monitoring and yield estimation. *Ecological Modelling*, 270, 30–42. https://doi.org/10.1016/j.ecolmodel.2013.08.016