

RESEARCH ARTICLE

10.1029/2018MS001362

**Key Points:**

- Data assimilation was used to initialize biomass and leaf area in the Community Land Model
- Adaptive inflation was needed to give more weight to observations due to substantial discrepancies between model forecast and observations
- Data assimilation reduces forecast error in a land surface model

**Supporting Information:**

- Supporting Information S1
- Data Set S1

**Correspondence to:**

A. M. Fox,  
andrewfox@email.arizona.edu

**Citation:**

Fox, A. M., Hoar, T. J., Anderson, J. L., Arellano, 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, 2471–2494. <https://doi.org/10.1002/2018MS001362>

Received 30 APR 2018

Accepted 25 SEP 2018

Accepted article online 27 SEP 2018

Published online 21 OCT 2018

## Evaluation of a Data Assimilation System for Land Surface Models Using CLM4.5

Andrew M. Fox<sup>1</sup> , Timothy J. Hoar<sup>2</sup> , Jeffrey L. Anderson<sup>2</sup>, Avelino F. Arellano<sup>3</sup> , William K. Smith<sup>1</sup> , Marcy E. Litvak<sup>4</sup> , Natasha MacBean<sup>1</sup> , David S. Schimel<sup>5</sup>, and David J. P. Moore<sup>1</sup>

<sup>1</sup>School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA, <sup>2</sup>National Center for Atmospheric Research, Boulder, CO, USA, <sup>3</sup>Hydrological and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA, <sup>4</sup>Department of Biology, University of New Mexico, Albuquerque, NM, USA, <sup>5</sup>Jet Propulsion Laboratory, Pasadena, CA, USA

**Abstract** The magnitude and persistence of land carbon (C) pools influence long-term climate feedbacks. Interactive ecological processes influence land C pools and our understanding of these processes is imperfect so land surface models have errors and biases when compared to each other and to real observations. Here we implement an Ensemble Adjustment Kalman Filter (EAKF), a sequential state data assimilation technique to reduce these errors and biases. We implement the EAKF using the Data Assimilation Research Testbed coupled with the Community Land Model (CLM 4.5 in CESM 1.2). We assimilated simulated and real satellite observations for a site in central New Mexico, United States. A series of observing system simulation experiments allowed assessment of the data assimilation system without model error. This showed that assimilating biomass and leaf area index observations decreased model error in C dynamics forecasts (29% using biomass observations and 40% using leaf area index observations) and that assimilation in combination shows greater improvement (51% using both observation streams). Assimilating real observations highlighted likely model structural errors and we implemented an adaptive model-variance-inflation technique to allow the model to track the observations. Monthly and longer model forecasts using real observations were improved relative to forecasts without data assimilation. The reliable forecast lead-time varied by model pool and is dependent on how tightly the C pool is coupled to meteorologically driven processes. The EAKF and similar state data assimilation techniques could reduce errors in projections of the land C sink and provide more robust forecasts of C pools and land-atmosphere exchanges.

**Plain Language Summary** The amount of carbon stored in vegetation and soils is an important control on how much carbon dioxide is in the atmosphere, and that influences future climate. Land surface models are used to simulate where this carbon is, but they are imperfect and there are often differences between model predictions and observations of the carbon stores. Here we describe a system that combines model predictions and observations and updates the modeled carbon stores so they are closer to the observations, considering uncertainty in both the model and the observations. We test our system at a location in New Mexico, United States, where we use observations from satellites of the amount of leaves on the vegetation and the amount of carbon stored in the vegetation. When we combine these observations with our land surface model there are large changes in the predicted amounts of stored carbon and the times of the year when the vegetation has the most leaves. These changes persist in the model after we stop updating it with observations, improving the model forecast.

### 1. Introduction

Forecasting how ecosystems respond to environmental drivers is of particular importance in periods of rapid climate and land use change (Y. Luo et al., 2011). The need to predict ecological processes for resource management at local scales (Clark et al., 2001) is mirrored by the need to predict how ecological processes feedback to climate at global scales (Friedlingstein et al., 2006, 2014; Pitman, 2003). Large-scale ecological processes influence surface albedo, carbon (C) cycling, and hydrology, which in turn influence how climate evolves through time (Bonan, 2008; Luyssaert et al., 2014). Land surface models (LSMs) are the components of Earth system models (ESMs) that calculate the exchange of C, water, and energy between the land and the

©2018. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

atmosphere (Hurrell et al., 2013). The development of a robust forecasting system for LSMs is a key component of achieving robust climate forecasts.

A key challenge of forecasting the behavior of ecological systems is that the current state of the ecosystem exerts considerable control on its function. For example, the gross primary productivity (GPP) is controlled by amount of leaf area and the photosynthetic rate per unit area; knowing the present state, in this case the amount of leaf area, allows a more constrained forecast of GPP. At local scales, ecosystem models are often initialized with local data using a variety of techniques (e.g., Aber & Fédérer, 1992; Antonarakis et al., 2011; Moore et al., 2008). A different approach is typically taken at the larger scales relevant to ESMs; C pools in LSMs need to be *spun up* by running the model for many centuries (e.g., Thornton & Rosenbloom, 2005) or analytically solved (e.g., J. Y. Xia et al., 2012) to achieve reasonable pool sizes before the model can be used to address scientific questions. However, complexities in the distribution of geology, climate and historical disturbance and land use, lead to biases in LSMs relative to the real world. In the LSMs, historical bias in C pools is correlated with projected bias (Friedlingstein et al., 2014), leading to unreasonable variation in future model states, C sequestration, and climate feedbacks.

One way to reduce bias in projections of future C sequestration is to create an initialized forecast—ensuring that the initial conditions of the model correspond to observations in the real world. Data assimilation (DA) can be used to improve the coherence of LSMs with observations. DA is a general term for methods that systematically combine information from observations and a model to achieve a better understanding of the system than either models or observations independently. DA has long been widespread in atmospheric sciences and numerical weather prediction (Daley, 1993; Davies & Turner, 1977; Kalnay, 2003; Lynch, 2006) and researchers in hydrological, climate, and ocean modeling have adopted similar techniques (e.g., Beven & Freer, 2001; Karspeck et al., 2013; Raeder et al., 2012). DA has also proved useful in understanding the complexity of processes controlling the exchange of C, water, and energy between the land surface and other components of the Earth system (Fox et al., 2009; Raupach et al., 2005; Trudinger et al., 2007; Wang et al., 2009; Williams et al., 2009). The usefulness of DA in improving the performance of LSMs has resulted in the development of a number of Land Data Assimilation Systems (LDASs) over the last decade. Earlier examples include the North American LDAS (K. E. Mitchell, 2004; Y. Xia et al., 2012) and the Global LDAS (Rodell et al., 2004). More recently, drawing upon the same underlying NASA Land Information System frame as NLDAS and GLDAS, the Famine Early Warning Systems Network LDAS (McNally et al., 2017) and the National Climate Assessment LDAS (Kumar et al., 2018) have been developed. The focus of these systems has been on the assimilation of observations related to soil moisture, with overall goal of improving representation of the water cycle and energy flux. A different approach has been taken by Sawada (Sawada et al., 2015; Sawada & Koike, 2014), who developed a system that uses microwave observations to jointly constrain model estimates of surface soil moisture and vegetation biomass, which they evaluate against independent observations of leaf area index (LAI). Finally, the LDAS-Monde of Albergel et al. (2017) uses a simplified extended Kalman Filter to assimilate surface soil moisture and LAI observations into the interactions between soil-biosphere-atmosphere model across the Mediterranean region of Europe, building upon the earlier site level work of Barbu et al. (2011). Since we cannot observe everything, any method proposed for initialization must simultaneously allow for changes in observed and unobserved quantities. DA relies on using real observations and while many model states, or their proxies, can be readily observed, others are either beyond current technology or investment in labor. We are fortunate that in ecological systems many C pools and fluxes are constrained by growth forms established by evolutionary or ecological processes (Raunkiaer, 1934) and allometric relationships (Chojnacki et al., 2014; Jenkins et al., 2003). For example, the maximum amount of leaf biomass is in part determined by the corresponding stem biomass; leaf biomass in turn limits the amount of C taken up through GPP. This logic can be extended to states associated with other linked processes in a LSM. DA has been successfully applied to simple ecosystem models with modest amounts of data (e.g., Bloom & Williams, 2015; Keenan et al., 2012; Moore et al., 2008; Williams et al., 2005; Zobitz et al., 2014) and encounters computational challenges when it is applied to LSMs embedded in ESMs using multiple streams of heterogeneous data due to both the complexity of the models and magnitude of the data sets (Y. Luo et al., 2016). Some of these challenges may be addressed by the coupling of LSMs to an Ensemble Adjustment Kalman Filter (EAKF). The EAKF is one of several sequential techniques that reduce the computational load of comparing observations with model predictions within a predefined assimilation window until all of the data have

been considered (Evensen, 1994). This analysis is carried out using an ensemble (or sets) of model realizations which are weeded out, nudged, or filtered by another method (Burgers et al., 1998; Houtekamer & Mitchell, 1998). The technique allows for different data types to be combined simultaneously and allows unobserved states to be updated.

When developing a DA system, it is useful to test the system using both model experiments or observing system simulation experiments (OSSEs) and real-world experiments (Atlas, 1997). OSSEs allow us to circumvent model limitations and examine how statistical updates to models using information in the observations update observed and unobserved states and improve forecasts (Masutani et al., 2010). In OSSEs, the observations used as *truth* in the DA system are consistent with the model structure because they are derived from the model itself. Incorrect model predictions are therefore not caused by model structural problems, but instead by the failure of the DA system to shift the model to match *true* states. Pairing OSSEs with DA experiments with real data then allows model inaccuracies to be identified.

In this study we evaluate the performance of a DA system composed of the Community Land Model (CLM) 4.5 (CESM 1.2 release) LSM and an implementation of the EAKF using the Data Assimilation Research Testbed (DART; J. Anderson et al., 2009). The goals are (1) to establish whether the EAKF can be used to create an initialized forecast for CLM that will mitigate against variation in the initial state of C pools and (2) whether the forecast ability of the model is improved. We achieve this by carrying out first an OSSE and then a case study using real data and evaluate the ability of the system to assimilate and correctly replicate LAI and above-ground biomass estimates and evaluate the behavior of related pools and fluxes. We discuss the implications for approaches like this one for improving consistency between LSMs and expanding observational records.

## 2. Methods

### 2.1. Community Land Model

The CLM is a well-established and open source mathematical model of land surface processes (D. Lawrence & Fisher, 2013), developed for use as the land component in the fully coupled CESM (Hurrell et al., 2013). The CLM bibliography currently lists over 450 publications using the model since 2001 (NCAR, 2018). It is a fully prognostic model, and at each time step (30 min), the model calculates the cycling of energy, water, C, and nitrogen and updates state variables for each grid cell. In common with other LSMs, CLM represents the many thousands of plant species as a limited set of plant functional types (PFTs) at the subgrid level (Bonan et al., 2002), allowing the model to prescribe parameters that control land-atmosphere interactions and internal dynamics of the model for each point in space. The climate sensitivities of ecosystem processes are controlled by the initial states and parameter sets of the model.

CLM is driven by a limited set of atmospheric variables (air temperature, shortwave and longwave radiation, specific humidity, precipitation, wind speed, and barometric pressure), which may come from site observations, reanalysis data sets, or a coupled atmospheric model. In this case we use output from an 80 ensemble member Community Atmosphere Model (CAM) DA reanalysis (Raeder et al., 2012), available for years 1998–2010, down sampled for the grid cell containing our location. This introduces some additional uncertainty compared to running the model with observed site-specific meteorology but we rely on this to generate ensemble spread in the DA approach described below.

The PoinT CLM tool is used to generate location-specific surface input data for all our experiments so we can run the model in single point mode. It is a Python script that first reads an ASCII text file defining PFT distribution (60% Evergreen Needleleaf Temperate, 20% C4 grassland, and 20% bare ground) and soil properties (40% sand and 20% clay fractions) at our location. Second, PoinT CLM runs additional CLM tools to subset global surface data sets for additional required characteristics, and then combines this with the user provided information. Finally, it copies the new input files to a disk location using a standard naming convention. The 80-member CAM ensemble is subsetting separately for the site location. As there is no communication between grid cells in CLM, the model will perform exactly as it will in a global run, but at a fraction of the computational cost. Although CLM is normally run globally, using the model in single point mode is common for development and testing (e.g., Bonan et al., 2012; Duarte et al., 2017; Mao et al., 2016; Wieder et al., 2017). A multi-instance CLM version 4.5 with active biogeochemistry and prognostic leaf area case (an ICLM45BGC component set—see technical description in Oleson et al., 2013) with 80 ensemble members was spun up

by simply cycling through the 12 years of CAM output for 1,000 years using 20th century atmospheric CO<sub>2</sub> concentrations and nitrogen deposition rates. We did not use more detailed transient drivers because previous experiments indicated they had only a very minor impact on spun-up model state. Here we focus on the dynamics of C pools within the model. The model is prognostic with respect to C (and nitrogen) state variables in the vegetation, litter, and soil organic matter. Vegetation phenology and canopy heights are also prognostic. Briefly, C is taken up through photosynthesis and, once autotrophic respiration is deducted, C is allocated to either leaves, wood, or roots depending on the phenological phase. A fixed proportion of each pool turns over during each time step and is added to litter C. A detailed description of the biogeochemical component can be found in Koven et al. (2013) and Oleson et al. (2013).

## 2.2. EnKF Implementation

We have coupled CLM with the DART, which is an open-source community software facility for ensemble DA developed and maintained by a group at the National Center for Atmospheric Research (J. Anderson et al., 2009). DART has been coupled with many large, highly complex geophysical models, including the two other components of the CESM, the CAM (Raeder et al., 2012) and the ocean component, Parallel Ocean Project 2 (Karspeck et al., 2013), as well as regional models such as the weather research and forecasting model (Schwartz et al., 2015).

DART provides access to a number of different ensemble filters, including the traditional Ensemble Kalman Filter (EnKF; Evensen, 1994), the EAKF (J. L. Anderson, 2001), the Rank Histogram Filter, and the Particle Filter. Here we use the EAKF, which unlike the EnKF does not include a stochastic component, but is a fully deterministic algorithm for estimating model forecast error statistics based on observation uncertainty (J. L. Anderson, 2001).

Key to the effectiveness of DART is the ease with which enhancements to the basic filter algorithms can be implemented. Notably, localization algorithms can be implemented to 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 (J. L. Anderson, 2007b, 2012). This ability is not utilized in this study as we have a relatively large ensemble and are running the model at a single location in space. However, we do investigate the utility of another key enhancement available with DART, namely adaptive inflation (J. L. Anderson, 2007a, 2009). This is an additional approach for combating insufficient forecast error variance (as described by ensemble spread) which can lead to filter divergence when the ensemble moves away from the observations and the DA system consequently rejects them. Insufficient forecast error variance may result from sampling error caused by small ensemble size, but can also be caused by a wide variety of other sources of imperfection in the system, in particular how ensemble spread is generated and maintained and how model error is accounted for. The nature of complex LSMs such as CLM that contains many empirically derived relationships and parameterizations, and processes interacting over very different timescales, results in an expectation of model error that is relatively high when compared to other geophysical models typically used with DART that rely on more fundamental physical equations. Therefore, we investigate through a series of experiments the use of an adaptive inflation algorithm. This applies Bayesian estimation theory to the probability density function of the inflation factors by first associating normally distributed inflation random variables with each element of the state vector, and these are then incrementally updated during the assimilation of observations.

Over the last 5 years the capability to assimilate many land model specific observation types has been added to the general DART framework, including snow cover fraction (Zhang et al., 2014), soil moisture from cosmic ray neutron intensities (Rosolem et al., 2014), total water storage anomalies from Gravity Recovery and Climate Experiment, albedo, and microwave brightness temperatures. Here we describe efforts focused on the C cycle in LSMs, assimilating aboveground biomass, leaf area, and flux of C dioxide (CO<sub>2</sub>).

Assimilating observations using the EAKF in DART results in updates of CLM in three ways. First, and most directly, the filter calculates the Kalman gain to update *observed* model variables. These may be parts of the model that are measured directly, or that are related to the measurement through an observations operator, that can be a simple transfer function. In this case we use observations of (1) aboveground biomass, that is calculated as the summation of three model state variables—leaf C, live stem C, and dead stem C; and (2) leaf area, that is calculated from leaf C and specific leaf area parameters.

Second are a group of *unobserved* model variables that we extract from the restart files and include in the DART state vector (Full list of all state vector variables are listed on Supporting Information Table S1). Although we do not have observations of these variables, they are updated by the EAKF through their ensemble correlation with the variables we do observe, with the magnitude of this update being dependent on the strength of this correlation. This is one of the most powerful aspects of using an ensemble filter as it allows us to transfer information we have about observed variables to update unobserved variables in a statistically rigorous way, assuming a linear relationship. This approach does not require physical relationships between pools to be held constant; C:N ratios or the relative amount of C in soil versus live biomass may vary. We include all the large C and nitrogen pools in the model in the state vector and update them simultaneously. Due to simple partitioning and allocation routines within the model, the correlations between the large C and N pools are strong so that if they were modified, they were modified in concert. Therefore, large disequilibria between pools did not result from the DA and consequently we did not observe model instability and spurious dynamics. These changes to states and fluxes influence many other calculations in the model. There is a third set of model variables which is modified by the DA process but only indirectly; updates to observed or unobserved model components can be transferred to other model variables through biophysical processes and feedbacks in the model. These may be relatively fast effects, or may only become apparent through feedbacks over long time periods. We refer to this set of components as *downstream* updates which includes all model components to a greater or lesser degree but we will focus our discussion on those which are not in the first two groups, for example, energy balance components.

### 2.3. Coupling CLM and DART

Working with CLM 4.5, we use the *multi-instance* capability developed for CESM that enables many instances of CESM component models to run within a single executable. At a predetermined time interval based on the availability of observations, monthly in the case of the experiments described here, the model ensemble is stopped at 00 UTC and restart files describing the model state in each ensemble member are written, along with history files of the half hourly flux output for the previous 24 hr. A subset of model variables and flux outputs (the *state vector*—Table S1 lists the variables included for these experiments) are then passed to DART, along with observations for the previous 24 hr. If required, model variables are converted into observation space through the application of an appropriate observation operator. The ensemble of model variables in the state vector represents the model *forecast*, or prior, that is then updated by DART applying the EAKF to produce an *analysis*, or posterior. These observation space values (i.e., LAI) are then linearly regressed back into model state space (i.e., leaf C), and the updated values are inserted back into the ensemble of restart files, that are then used to initialize a restart of the CLM ensemble at the next time step. There is a relatively high computation cost to stopping and restarting the model in this way, but it has the great advantage that the whole DA system requires minimal adaptation of CLM, which as a community model is undergoing constant development and is frequently updated. This allows DART to stay current with the most up-to-date version of the model without requiring frequent updating. The only DA-specific modifications required for CLM source code were to remove energy, water, and C balance checks for the first time step of the model upon restart following the DA update, and this will be included by default in future releases of the model. All the code required to set up CLM for use with DART, to preprocess observations, and to implement DART and apply the EAKF, along with detailed documentation, are freely available from the DART website (DART, 2017).

The translation between model space and observation space is carried out by an *observation operator* that can vary in complexity. For these experiments, the observation operators involved are very straightforward. To determine a modeled LAI value that can be compared with observations, modeled leaf C is multiplied by a PFT specific leaf area (leaf area per mass of C), and fractional PFT coverage is used to weight a grid cell average value. Similarly, above ground biomass observations are calculated by summing modeled leaf C, live stem C, and dead stem C, again weighted by fractional coverage of PFT.

### 2.4. Case Study Location

Although set up and run using the approach developed for *site-level* runs at an AmeriFlux tower site in central New Mexico, United States (42.425°N, 105.862°W), the spatial scale of both the input climate variables and the satellite derived leaf area and above-ground biomass observations used for assimilation are representative of a much larger area, being 2°, 0.5°, and 0.25°, respectively. This in keeping with the role of CLM as a global model, and therefore the experiments should be thought of being broadly representative of the semiarid



ecosystems in the region, including piñon-juniper woodland and C4 grasslands. The precipitation regime is driven by irregular snow-dominated winter events (from January to March), dry springs (April to June), and summer monsoonal events (from July to September).

## 2.5. Observations

### 2.5.1. Biomass Observations

When identifying observations to use with CLM-DART, we have initially concentrated on remote sensing products with global coverage, relatively coarse spatial resolution, and long temporal coverage, as these are most commensurate with the model's typical usage. The biomass product presented by Liu et al. (2015) is one of the very few such products that meets these requirements and comes with well-quantified uncertainties. The product provides global annual above-ground C estimates for two decades (1993–2012), based on satellite passive microwave observations. Satellite microwave radiation observations can be used to derive estimates of vegetation optical depth (VOD), which is a function of the amount of water in above-ground biomass (Liu et al., 2011; Owe et al., 2001). VOD has the advantage of not saturating at high biomass values, but due to its relatively coarse resolution, it cannot be calibrated or validated directly with plot measurements of biomass. Liu et al. (2015) instead indirectly relate VOD to biomass using the benchmark map of Saatchi et al. (2011), itself a fusion of multiple types of satellite observations and plot measurements. The resultant VOD-derived biomass product is available at 0.25° annually, with an uncertainty characterized by upper and lower bounds for each pixel in each year. For the assimilation, we treated these as 95% confidence limits and estimated an observation standard deviation as 25% of this range. Here we use a subset of the full data set for our location in New Mexico between 2001 and 2010.

### 2.5.2. Leaf Area Observations

Unlike biomass, a wide variety of remotely sensed LAI data products are available. In keeping with our desire to use observations commensurate with typical CLM usage, we elected to use the LAI product described by P. J. Lawrence and Chase (2007) that is input to CLM when used in the diagnostic *Satellite phenology* mode, which forces the model using a phenological climatology. This product is derived from the Moderate Resolution Imaging Spectrometer (MODIS) and reproduces the physical properties of the MODIS land surface products while maintaining the PFT representation used in CLM. Here we use the product aggregated to 0.5° spatial resolution and monthly temporal resolution, which is typical for usage with CLM in Satellite phenology mode. Currently, uncertainty is not well quantified for this product, so we obtained a time series of MODIS LAI products (MOD15A2) downloaded as 8-day composites for our location in New Mexico using the ORNL DAAC MODIS collection 5 Land Products Global Subsetting tool (ORNL DAAC, 2008). This product includes a quality control flag which we used to select high quality observations (QC value <248) and an observation error estimate. We found the high quality observations typically had an observation error standard deviation of 0.1 m<sup>2</sup>/m<sup>2</sup>, with some higher values associated with higher values of LAI. For these experiments we assumed a constant error standard deviation for the 0.5° product of 0.2 m<sup>2</sup>/m<sup>2</sup> for our location from 2001, the first year available, to 2010.

### 2.5.3. Synthetic Data for OSSE

In order to generate synthetic observations to use in the OSSE, one ensemble member from the 80-member CAM DA reanalysis, down sampled for the grid cell containing our location, was selected at random and used to force an offline CLM4.5 with active biogeochemistry run (described by an ICLM45BGC component setting) for years 1998–2010. Synthetic observations are generated from this model run to mirror the real data products described above. Therefore, 13 synthetic observations of biomass were generated from the ensemble member representing the true state of the system each year from 1998–2010 on 1 February by applying a random perturbation to the truth with zero mean and variance of 158 g C/m<sup>2</sup>, equivalent to the mean observation error standard deviation for the VOD-derived biomass product at this location over this time period. The 156 synthetic observations of LAI were generated every month in a similar manner using an observation error standard deviation of 0.2 m<sup>2</sup>/m<sup>2</sup>.

## 2.6. Experimental Setup

We carried out eight experiments and a free-run of the model to test and develop the performance of CLM and DART assimilating above-ground biomass and leaf area. For each experiment we employed an ensemble ICLM45BGC compset CLM run forced with the complete 80-member CAM ensemble files, using years 1998–2010 for the OSSE and years 2001–2010 for the real data runs. This means that the single CAM ensemble

member used to force the model run that was later sampled to generate synthetic observations for the OSSE was also used as one of the 80 CAM ensemble members in the assimilation part of the experiment. We do not believe this impacted the results as it represents only 1 of the 80 ensemble members, and this ensemble member did not behave in a qualitatively different way from all the others. The amount of observation error added when generating the synthetic observations was also large when compared with the ensemble spread (Figure 2).

#### 2.6.1. Initial Conditions Versus Forced Boundary Conditions Experiments

First, we conducted two simulation experiments without any DA to assess the relative impact of initial condition uncertainty versus forced boundary conditions (climate driver) uncertainty. By defining 1 of the 80 ensemble members as truth we estimated the relative importance of accurate model states or accurate climate drivers. Halfway through the 13-year model ensemble run we begin two experiments. The first continues the model ensemble run, but each member subsequently sees identical climate forcing from the single ensemble member defined as truth. The second uses the single ensemble member defined as truth as the initial conditions for all 80 ensemble members, which are then driven forward in time with their specific climate forcing. We calculate root-mean-square error (RMSE) as the ensemble-average of the RMSE between each ensemble member and the ensemble member treated as the truth.

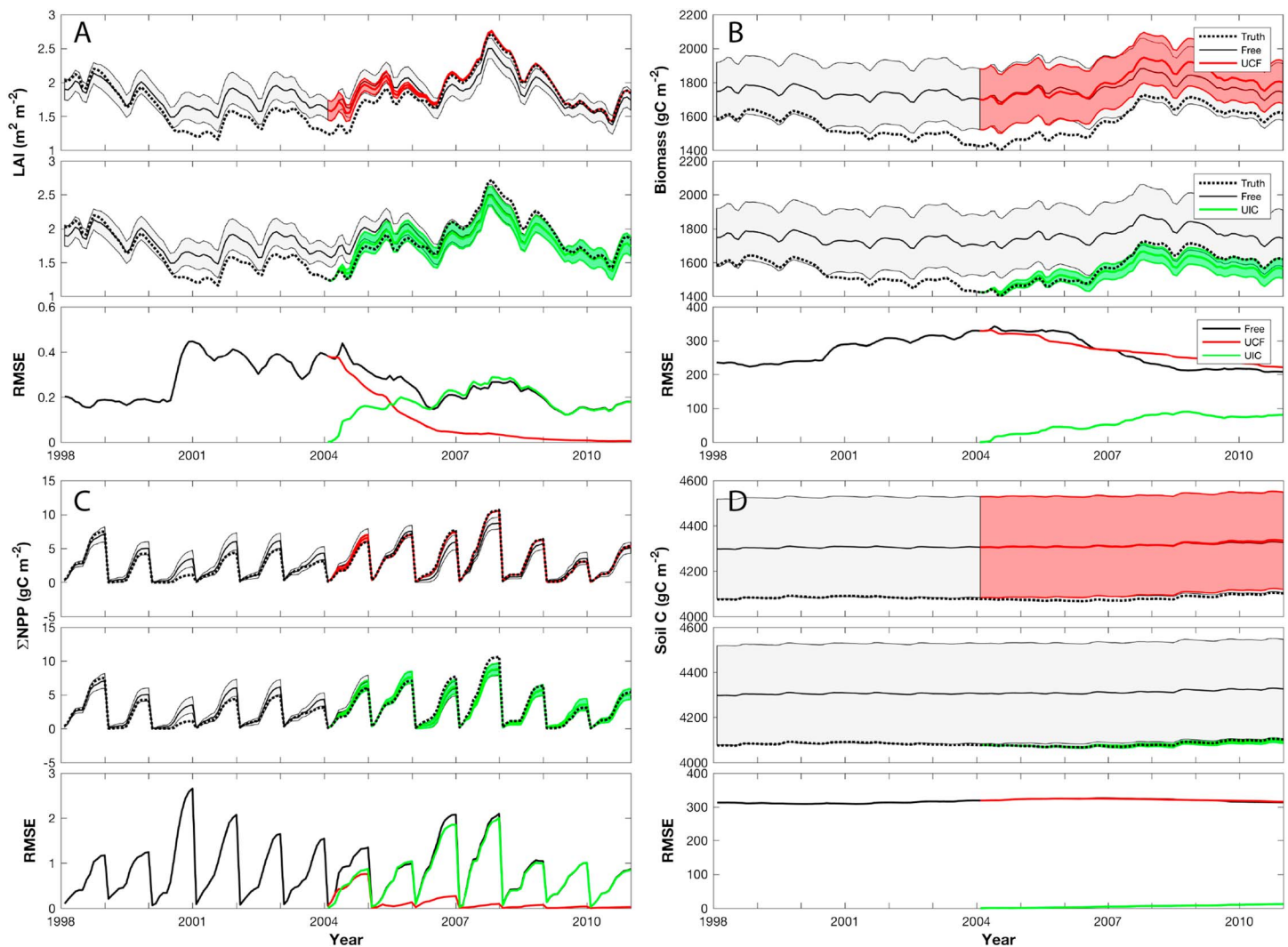
#### 2.6.2. Observing System Simulation Experiments

We implemented three OSSEs. For the OSSEs we again treat one of the ensemble members as truth, and then sample this to provide synthetic observations (section 2.5.3). The first OSSE studies the impacts of assimilating an annual observation of biomass, a second studies the effects of assimilating monthly LAI values, and a third studies the effects of simultaneously assimilating both. In each experiment we compared the reduction in RMSE with respect to C pools and fluxes following assimilation, relative to a free-run of the model ensemble.

There are two important differences between the OSSEs and the real observations experiments that follow which make the OSSE a powerful tool. First, model structural error and parameter uncertainty is removed from the system. These are often large, but difficult to accurately quantify. This makes it difficult to assess and test many key aspects of DA system performance that can be treated independently from dealing with model error. Second, the true state in the OSSE is known without error. Therefore, the impact of the assimilation can be accurately quantified by comparing how the free-run and the DA-run ensemble perform with respect to the ensemble member treated as truth. We compare the change in RMSE between the model ensemble in the free-run and the ensemble member treated as truth, versus the RMSE of the model ensemble in the DA run with the ensemble member treated as truth, with an anticipated reduction in ensemble spread.

#### 2.6.3. Real Observations Experiments

We conducted three experiments assimilating real observations of both biomass and LAI simultaneously over the time period when they were both available. The first experiment did not use inflation, and the second used one of the adaptive inflation algorithms in DART. We then carried out a third experiment to investigate the impacts of the DA on longer-term forecasting. We replicated the previous experiment using biomass and LAI observations, and adaptive inflation, for years 2001–2006 and then continued without assimilating any observations over the following 5 years. In the first two experiments we assimilate observations, update the model state to generate a new initial condition, and then make a month-long forecast until the next observation is available. The mismatch between this forecast and the new observation is used to calculate our RMSE skill metric. The assimilation cycle then starts again as we then assimilate this new observation, before making the next month-long forecast. In order to see if this updated initial condition has an impact on a forecast for longer than 1 month, we then carried out a third experiment to investigate the impacts of the DA on longer-term forecasting. For each of the experiments, we assess the impact of the assimilation against a free-run of the model ensemble. It is important to note that during these experiments using real observations, the true state of the system is never known, and we can only calculate the RMSE between the model ensemble and the observations themselves just prior to their assimilation. So, in all cases we do these calculations for the forecast or prior ensemble, rather than for the analysis, or posterior ensemble. Put differently, the RMSEs provided here quantify the errors in the observation-minus-forecast residuals.



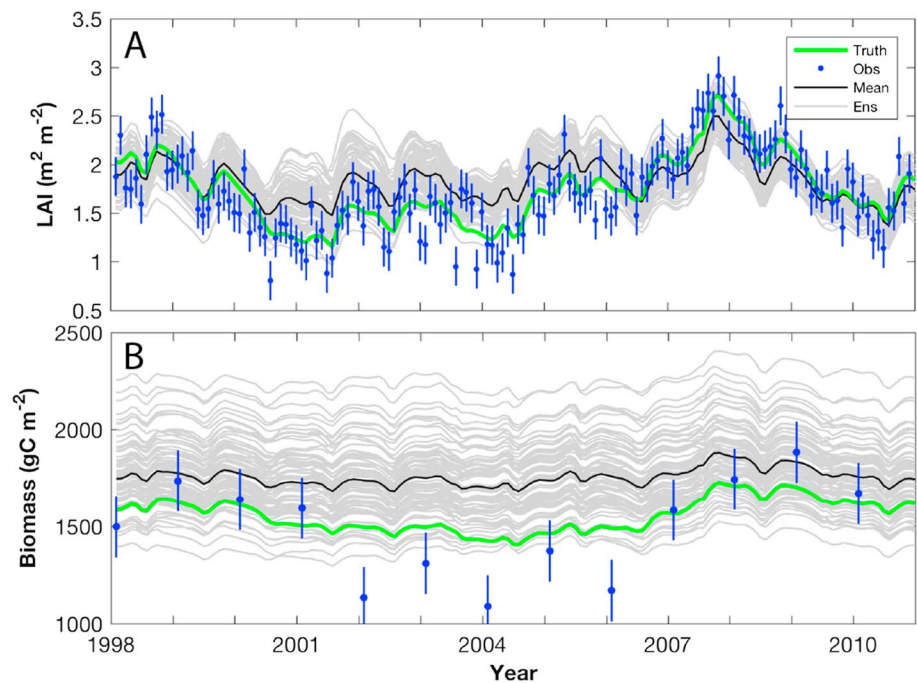
**Figure 1.** Ensemble estimates of C pools and fluxes derived from CLM4.5BGC over 13 years from 1998 through 2010. There are three ensemble estimates shown from 2004 onward. The spread in the free-run (free, gray) and the UIC (red) is caused by variation in meteorology derived from an 80-member Community Atmosphere Model ensemble. The spread in the UCF (green) is derived from different 2004 initial conditions defined by the spread in the free-run model conditions at that time, with no variance between ensemble members in the meteorological forcing after 2004. RMSE is expressed as the ensemble-average of the RMSE between each ensemble member and the ensemble member treated as the truth, which is also used for starting the UIC runs. The shaded areas represent +1 standard deviation around the ensemble mean. LAI = leaf area index; RMSE = root-mean-square error; NPP = Net Primary Production; C = carbon; UIC = Uniform Initial Condition; UCF = Uniform Climate Forcing.

### 3. Results and Discussion

#### 3.1. Initial Conditions Versus Forced Boundary Conditions Experiments

Our first set of experiments was designed to evaluate the relative importance of uncertainty in the model states versus climate drivers before we carried out formal DA. This experiment indicates that accurate representation of the initial conditions of the model reduces model errors on a decadal scale, but the persistence of these reductions is dependent on how tightly related to meteorology the states are. For each model component evaluated (soil C, biomass, LAI, and Net Primary Production [NPP]), the ensemble predicted by a single atmospheric forcing converged, albeit slightly in some cases, while the ensembles driven by ensemble atmospheric forcings diverged (Figure 1). The degree of divergence indicates the timescale over which model forecasts are sensitive to initial conditions, and how closely the model attribute is tied to meteorological drivers; we illustrate this by showing time series of soil C, biomass, LAI, and NPP which have increasingly rapid responses to short-term change in meteorology. The benefits of initialization of model states last longer





**Figure 2.** LAI and above-ground biomass modeled for 80 ensemble members (gray lines) over 13 years. The range of this ensemble indicates the uncertainty introduced into the model from uncertain climate forcing during the spin-up process and is 40% and 60% of the ensemble mean for LAI and above-ground biomass, respectively. Synthetic observations (blue) show their uncertainty, and their spread around the ensemble member treated as truth (green) results from the random application of this uncertainty when generating observations to use in the observing system simulation experiments. LAI = leaf area index.

for model attributes that are a large proportion of total ecosystem stocks and less tightly related to changes in meteorology. For example, removing bias in the soil C stock reduced RMSE by over 95% and this persists nearly unchanged for 6 years (Figure 1), in contrast the benefit of correcting the NPP estimates disappears after 1 year while the LAI and Biomass (Figure 1) patterns lay somewhere in between. We infer from this experiment that adjusting the state (using DA or by another means) would likely reduce ensemble variance and improve the model forecast capability.

### 3.2. DA Using an OSSE

In this section we first illustrate the synthetic observations, then how the distribution of LAI derived from the ensemble changes when DA is implemented, and then explain three experiments that illustrate the effect of assimilating biomass or LAI independently or simultaneously under idealize model conditions.

#### 3.2.1. Temporal Dynamics in Synthetic Observations

While ensemble variability emerges from uncertainty in the meteorological drivers, variability in the synthetic observations derives from uncertainty in real observations added to the ensemble member defined as truth. Figure 2 shows all 80 ensemble members from the free-run of the model (gray curves) and synthetic observations (blue points). These observations were generated around the ensemble member that we defined as truth and variance was added by applying observation error estimated from real observations. All the synthetic observations do not lie on the ensemble member treated as truth due to the random noise added to represent observation error. There are approximately equal number of observations greater than and less than the truth as we are not attempting to simulate observation bias in these experiments, but we see at times that the introduced error is sufficient to result in synthetic observations outside the range of the free-run ensemble. The initial uncertainty in the model prediction of LAI and above-ground biomass resulting from uncertain climate forcing during spin-up is indicated by the range of the ensemble on 1 January 1998. For LAI this is  $0.75 \text{ m}^2/\text{m}^2$  (ranging from  $1.50$  to  $2.25 \text{ m}^2/\text{m}^2$ ), or  $\sim 40\%$  of the ensemble mean value. For above-ground biomass this is, or  $850 \text{ g C}/\text{m}^2$  (ranging from  $1,400$  to  $2,250 \text{ g C}/\text{m}^2$ ), which is over 60% of the ensemble

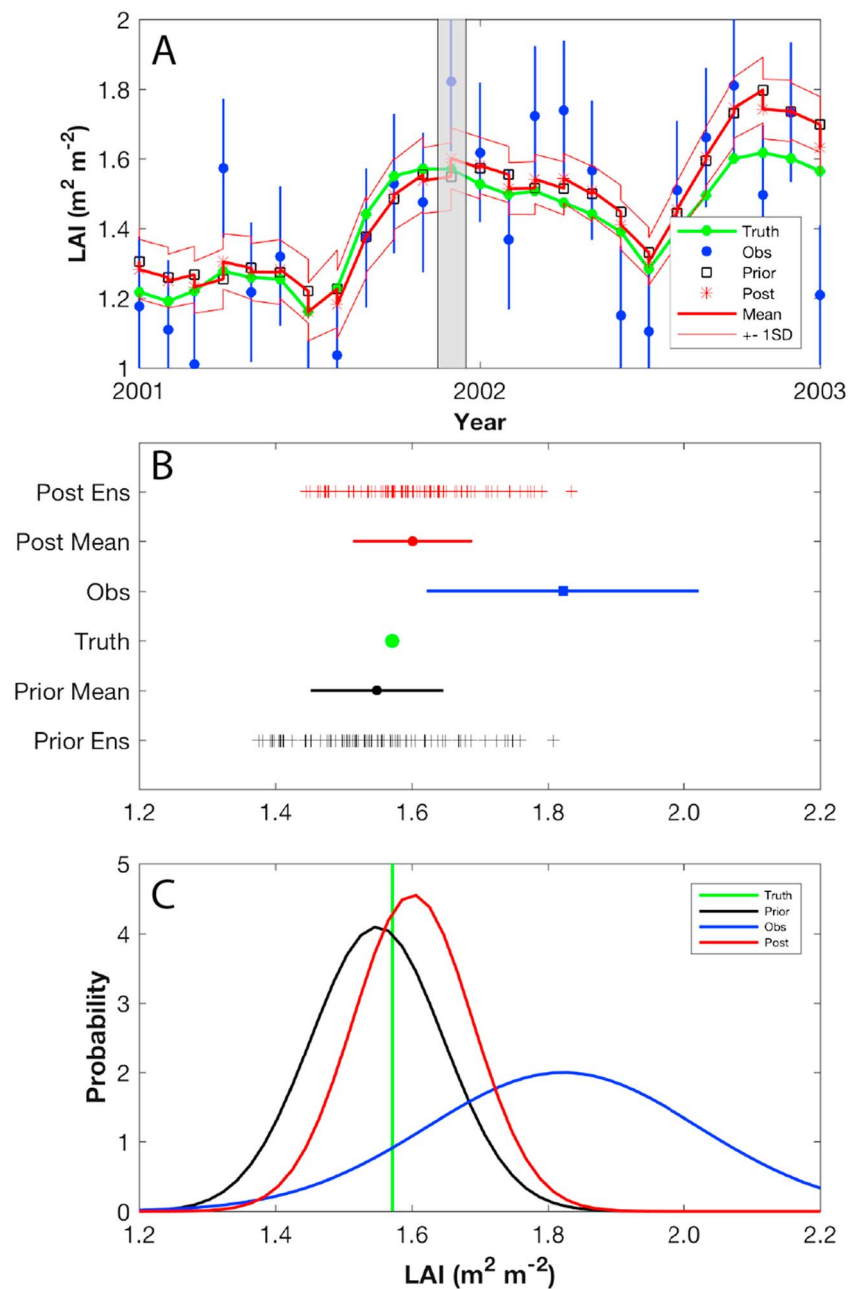
mean value. Over the next 13 years there is limited variation in these total ranges, although as demonstrated by the truth ensemble member, individual ensemble members vary in their relative position in the ensemble much more for LAI than for biomass (Figure 2). There is considerable interannual variation in modeled LAI, with the ensemble mean varying between 1.5 and 2.5  $\text{m}^2/\text{m}^2$ , and this makes it difficult to detect patterns of intra-annual phenological variation, although typically LAI tends to peak in the Fall, and then decline over the following 10 months. Interannual variability in above-ground biomass over this time period is small, and the small intra-annual variability in biomass largely reflects variation in leaf C described above.

### 3.2.2. Illustration of the Effect of State DA on the Distribution of States Predicted by the Ensemble

The EAKF adjusts the ensemble toward the synthetic observations thus reducing the error in LAI on average (Figure 3). Occasionally, because we have added observation error to the synthetic observations the RMSE increases when the synthetic observations happen to fall a long way from the truth. The time series of model states exhibits the characteristic *sawtooth* pattern of sequential state DA systems (e.g., H. L. Mitchell et al., 2002). Similar patterns have also been seen in other studies assimilating LAI specifically (e.g., Albergel et al., 2010; Barbu et al., 2011). Each time an observation is available the ensemble forecast (prior) mean is adjusted toward the observation, as determined by the Kalman gain, to give a new analysis (posterior) value that is then used to initialize a forecast over the next time period. We see that in the illustrated time point, in December 2001, that the posterior distribution has a mean closer to the observation (and the ensemble member treated as true) and has a reduced uncertainty (Figures 3b and 3c).

### 3.2.3. Reductions of Model Errors by Assimilating Biomass and LAI

We found that assimilating two data-streams jointly was more effective than assimilating either independently (Table 1). We conducted three experiments to assess the relative importance of different observations. First, we assimilated just the annual biomass observations, then just the monthly LAI observations, and finally both simultaneously. In each experiment we can compare the reduction in RMSE following assimilation relative to the free-run ensemble. Considering the improvement in fit for LAI, assimilating biomass alone initially has a limited impact, until 2002 when there is 50% reduction in error (Figure 4a). This then persists for the next 4 years, after which assimilating biomass alone does not improve LAI forecasts relative to the free-run (Figure 4a). Assimilating the more frequent LAI observations results in a relatively constant, and low, RMSE of 0.1–0.2  $\text{m}^2/\text{m}^2$  for LAI throughout the assimilation period. This reflects the uncertainty in the observations. The fact that the ensemble does not *collapse* onto the true ensemble member is desirable and indicates the DA system is performing as expected. Assimilating both biomass and LAI together does not improve estimates of LAI beyond assimilating LAI alone. Given how effective the LAI only assimilation is compared to the biomass only assimilation, perhaps this is not surprising. The picture is different when we consider estimates of biomass. As expected, assimilating biomass observations leads to an immediate reduction in RMSE (Figure 4b). However, assimilation of LAI observations alone also leads to a gradual improvement in estimates of biomass (Figure 4b), due to the strong correlation between LAI and biomass within the model. Indeed, for portions of the record model estimates of biomass are better when assimilating LAI alone as the biomass observations fall outside the ensemble (purely by chance) because of observation error. Throughout most of the simulated data experiment, as theory would predict, the best estimates of biomass come from assimilating both biomass and LAI observations (red curve in Figures 4a and 4b), with the additional LAI observations compensating for uncertainty introduced through the biomass observation error. Assimilating synthetic observations of both LAI and biomass into CLM improves model predictions of these biomass pools by moving the ensemble mean toward the synthetic observations (Figure 3) but the effect depends on whether they are assimilated individually or simultaneously (Figure 4). The difference in the effect of assimilating either LAI or biomass is apparently caused by differences in the magnitude of the pools and the frequency at which they are assimilated. In addition to improving forecasts of LAI and biomass, we found that assimilating both types of observations was also optimal for improving forecasts of the other C pools and fluxes we considered in these experiments (Table 1). In addition to improving forecasts of LAI and biomass, we found that assimilating both types of observations was also optimal for improving forecasts of the other C pools and fluxes we considered in these experiments (Table 1). Indeed, the large reductions in RMSE relative to the Free ensemble for LAI (58%) and Biomass (69%) are matched with equivalent reductions associated with the unobserved litter, roots, and soil C pools (58%, 72%, and 71%, respectively). This highlights the strong correlation found between observed and unobserved variables due to the relatively



**Figure 3.** Illustration of how assimilation of sequential LAI data (blue dots) leads to a *sawtooth pattern* in the ensemble mean and spread. Expansion of the ensemble spread occurs between observations and narrowing of the spread occurs when observations are assimilated (panel a). The lower panels illustrate how the increment is calculated on 11/01/2001 (shaded area in top panel); first a normal distribution is estimated for the prior ensemble using its mean and variance (black dot; panel b) and the observation using the square of the known error standard deviation (blue dot; panel b). The posterior distribution is calculated as the Gaussian product of the prior and the observation likelihood. The prior ensemble is then updated to have an equivalent mean and variance as this posterior distribution, which in most cases narrows the ensemble spread (compare red and black curves in panel c). LAI = leaf area index.

simple, deterministic allocation and partitioning routines found within CLM. Therefore, in the absence of model error in the OSSE, these unobserved variables are updated very effectively. Improvements in simulated gross fluxes of GPP (42%) and ecosystem respiration (ER, 27%) are less than for the pools, but still very substantial. This can be explained by considering that the fluxes are functions of not only the model states that are updated by the DA system, but also the atmospheric forcing that varies between

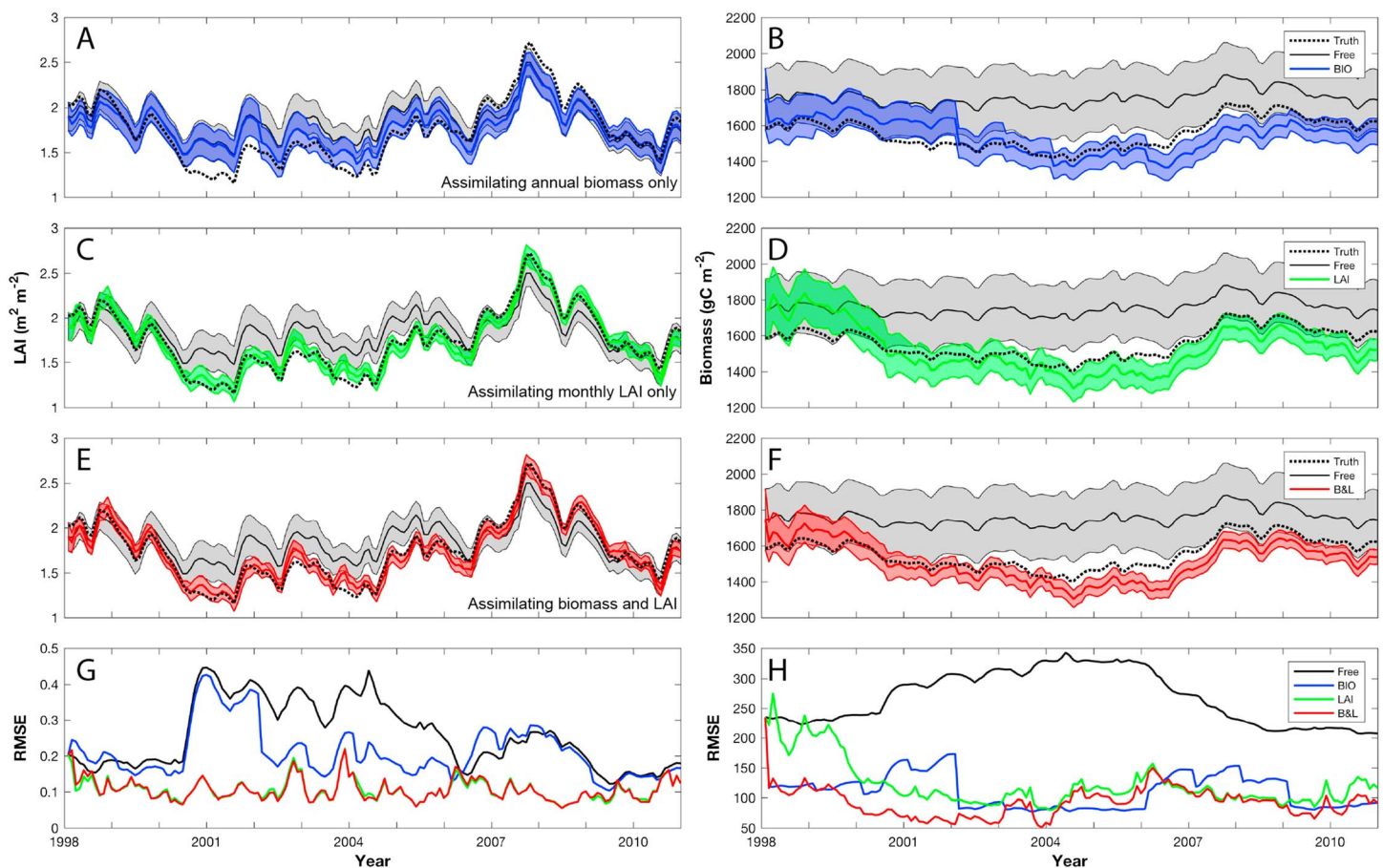
**Table 1**

*RMSE Between Model Forecast Ensemble and Ensemble Member Treated as Truth During the Observing System Simulation Experiments for Selected Carbon Stocks and Fluxes*

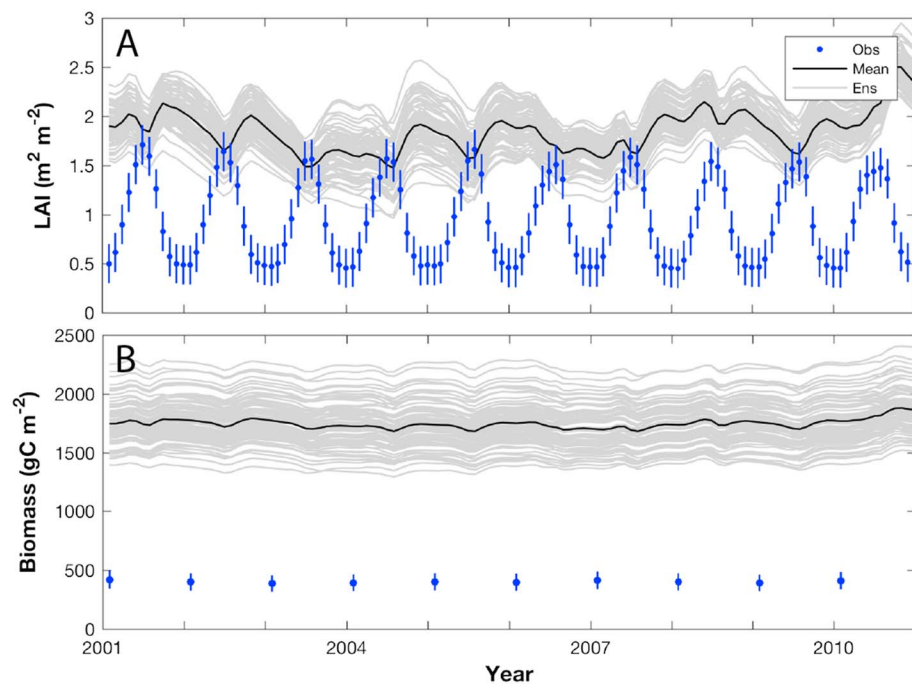
| RMSE                                     | Experiment |         |       |       |
|------------------------------------------|------------|---------|-------|-------|
|                                          | Free       | Biomass | LAI   | B & L |
| LAI ( $\text{m}^2/\text{m}^2$ )          | 0.26       | 0.23    | 0.11  | 0.11  |
| Biomass ( $\text{g C}/\text{m}^2$ )      | 268.3      | 126.3   | 120.8 | 82.6  |
| Litter ( $\text{g C}/\text{m}^2$ )       | 25.4       | 17.5    | 12.5  | 10.7  |
| Roots ( $\text{g C}/\text{m}^2$ )        | 116.9      | 53.5    | 50.8  | 33.1  |
| Soil ( $\text{g C}/\text{m}^2$ )         | 317.2      | 115.5   | 149.1 | 90.9  |
| NEP ( $\text{g C m}^{-2}$ per month)     | 6.78       | 6.66    | 6.26  | 6.27  |
| GPP ( $\text{g C}/\text{m}^2$ per month) | 12.6       | 11.1    | 11.0  | 7.3   |
| ER ( $\text{g C}/\text{m}^2$ per month)  | 7.30       | 6.79    | 5.37  | 5.30  |

*Note.* RMSE = root-mean-square error; LAI = leaf area index; NEP = net ecosystem production; GPP = gross primary productivity; ER = ecosystem respiration.

ensemble members, and which is not corrected for by DA. This result suggests that GPP is more sensitive to model states, specifically LAI, than atmospheric forcing when compared with ER as the relative reduction in error is larger after DA. The remaining errors in the gross fluxes are compounded in the net flux, net ecosystem production, which experiences the smallest reduction in error.



**Figure 4.** LAI and biomass estimates from three experiments assimilating just annual biomass observations (panels a and b), just monthly LAI observations (panels c and d) and both biomass and LAI observations together (panels e and f). Ensemble mean and +1 standard deviation describing ensemble spread is plotted for free-run (no assimilation) and experiment. RMSE between the ensemble member treated as truth and each experiment ensemble, and the free-run ensemble when no assimilation was carried out is shown in panels (g) and (h). LAI = leaf area index; RMSE = root-mean-square error.



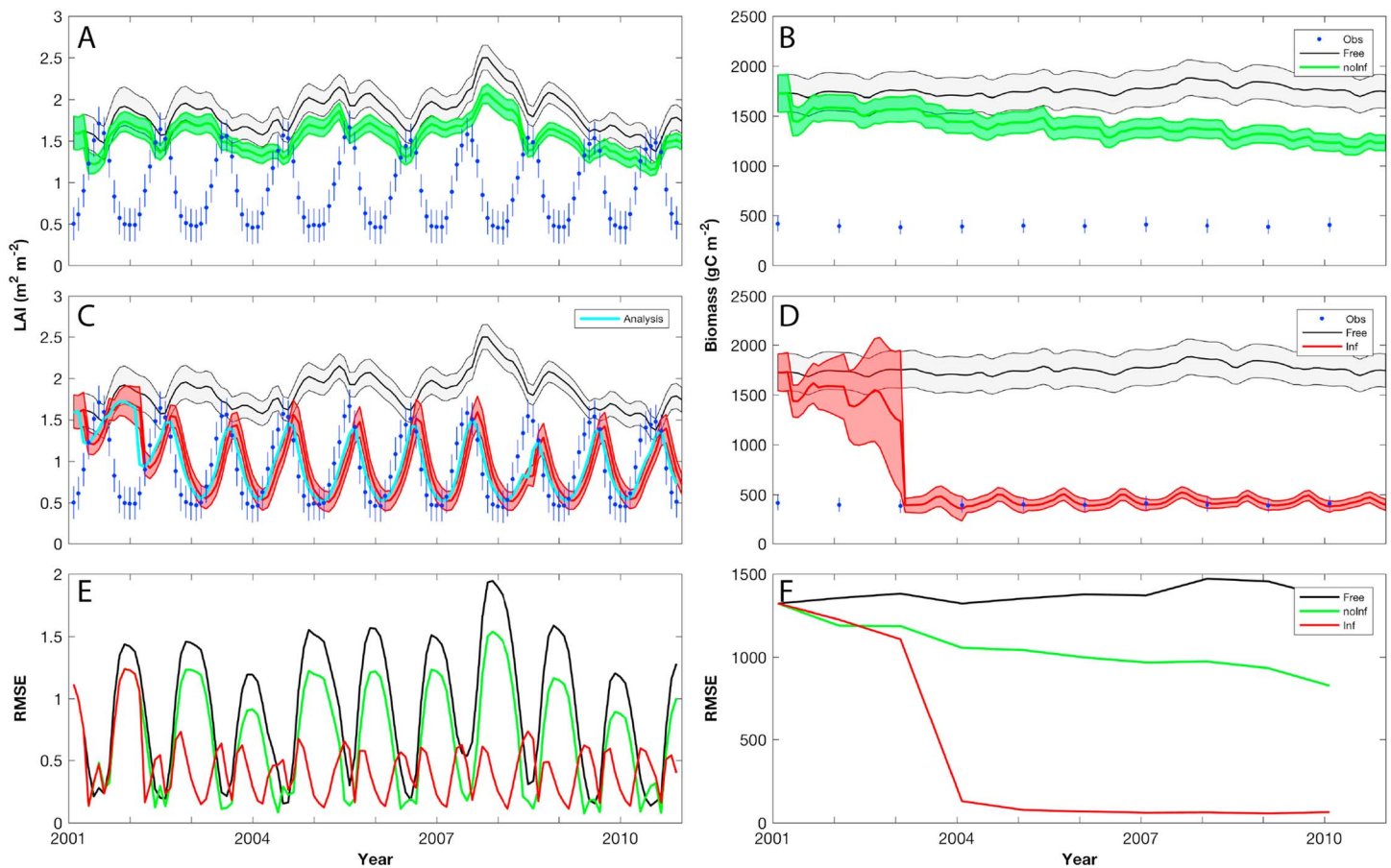
**Figure 5.** Time series of observations of LAI available at monthly time intervals, and biomass available at annual time intervals assimilated by Community Land Model- Data Assimilation Research Testbed. The sources of these derived products are detailed in section 2.5. LAI = leaf area index.

### 3.3. Model Comparison With Real Observations

In the real world we never know the true state of any part of the system of interest, but instead we only have uncertain observations. We find that the prior model does not compare favorably with the observations from our location in central NM as described in section 2.5 (Figure 5). Instead, we find systematic bias in magnitude of both LAI and biomass (Figure 5) and an apparent failure to correctly represent the phenology and environmental sensitivity of LAI (Figure 5a). For both satellite products we see potential mismatches between the fractions of PFT observed and those modeled. LAI observations exhibit a very smoothly varying annual cycle that shows very little interannual variability. In part this is a function of smoothing due to spatial averaging over multiple ecosystem types contained within the  $0.5^\circ$  LAI product and the relatively large observation error. Nonetheless, the lack of interannual variability, and an annual cycle that appears disconnected from known temporal patterns of moisture availability, to some extent reflects known limitations of remote sensing in these semiarid ecosystems (Biederman et al., 2016; Verma et al., 2015). Modeled biomass overestimates VOD-derived product threefold. As expected, there is little variation in biomass over this time period, and modeled biomass values are indicative of the spin-up process, and represent *steady-state* values calculated for the model under this climatic regime. As with the LAI product, there is potentially a discrepancy between the relative fractions of ecosystem-types in the  $0.25^\circ$  pixel size of the biomass product and prescribed PFT coverage in the model, with total biomass particularly sensitive to woody plant cover. For example, Anderson-Teixeira et al. (2010) note that above ground biomass for grassland, shrubland, juniper savanna, and piñon-juniper woodlands varies between 28, 137, 1,570, and 2,940  $\text{g C/m}^2$ , respectively.

The representation of LAI in the model is clearly poor (Figure 5). The CLM is not unique in failing to capture the magnitude and seasonality of LAI in semiarid ecosystems; several vegetation models compare unfavorably to satellite-based vegetation indices and LAI (e.g., Dahlin et al., 2015; MacBean et al., 2015; Traore et al., 2014). The reasons for this are likely varied and could include (1) failure to represent the sensitivity of vegetation to moisture (Biederman et al., 2017; Smith et al., 2018), (2) poor representation of hydrological processes and moisture availability to vegetation (Barron-Gafford et al., 2017; Matheny et al., 2014), and (3) failure to represent phenological processes (MacBean et al., 2015). DA may provide a temporary solution to this problem (Figure 6). While solving model structural errors are beyond the scope of this manuscript, it is clear that substantial model development is required for semiarid systems.





**Figure 6.** Forecast LAI and biomass for a free-run of the model ensemble with no assimilation (gray); model forecasts for an experiment without using adaptive inflation (green, panels a and b); and using adaptive inflation (red, panels c and d), and RMSE between the forecast ensemble and observations (panels e and f). Ensemble mean for the analysis of LAI is also shown for illustrative purposes in panel (c) as discussed in the text. LAI = leaf area index; RMSE = root-mean-square error.

For the purposes of this study that focuses on the technical development of the DA system, we treat these observations and their associated uncertainties as being accurate and ascribe the model-data mismatch to model error. That is to say, we have more confidence in the observations than in the model at this location and do not attempt any bias correction on the observations. We ask whether the EAKF can overcome these large data-model mismatches.

### 3.3.1. Assimilation Using Adaptive Inflation

The EAKF is predicated on the idea that the prior (forecast) ensemble has sufficient spread to sample the forecast uncertainty. However, the forecasts are all impacted by systematic model deficiencies; none of them sample the uncertainty associated with the unknown model errors. In addition, even with a perfect model, sampling error due to finite ensemble sizes also leads to insufficient ensemble spread. For this reason, DART includes algorithms to increase the forecast ensemble spread by inflating the ensemble of each state variable, increasing the distance between each ensemble member and the ensemble mean by an inflation factor. Here we use an adaptive inflation algorithm (J. L. Anderson, 2009) that uses the statistics of the differences between prior ensemble mean estimates of observations and the observed values to automatically select a time-varying inflation factor for each state variable.

DART also includes tools to ignore observations that are too far away from the ensemble mean. In our case, if the distance between the prior ensemble mean estimate of an observation and the observed value was more than three times the expected value (the square root of the sum of the prior spread and the observations error standard deviation), the observation was not used.

We find that the EAKF can overcome the large bias and apparent model error if we implement a technique to inflate the forecast error variance, which we define as the spread of the ensemble. In experiments where we

**Table 2**  
RMSE Between Model Forecast Ensemble LAI and Biomass and LAI and Biomass Observations

| RMSE                                |           | Experiment |         |           |          |
|-------------------------------------|-----------|------------|---------|-----------|----------|
|                                     |           | Free       | No Inf. | Inflation | Forecast |
| LAI ( $\text{m}^2/\text{m}^2$ )     | 2001–2010 | 0.93       | 0.70    | 0.44      | -        |
|                                     | 2006–2010 | 0.96       | 0.69    | 0.39      | 0.33     |
| Biomass ( $\text{g C}/\text{m}^2$ ) | 2001–2010 | 1,376.2    | 1,049.9 | 417.7     | -        |
|                                     | 2006–2010 | 1,406.3    | 940.29  | 62.8      | 51.4     |

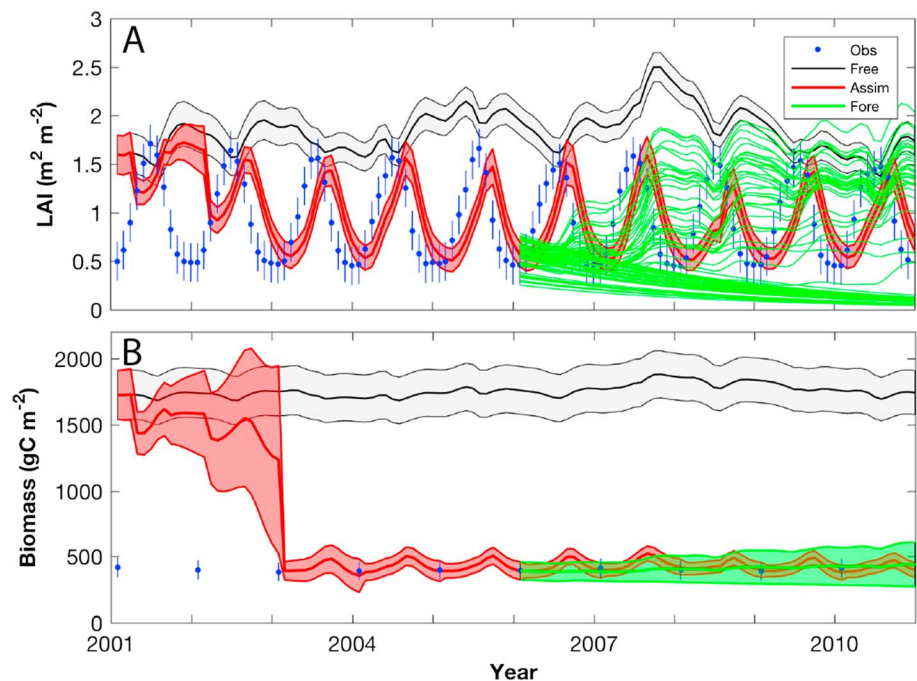
Note. RMSE = root-mean-square error; LAI = leaf area index. Forecast RMSE was only calculated for the time period 2006–2010.

limited the allowable mismatch to three standard deviations with no inflation we found that only the highest LAI observations were used in the assimilations (Figure 6a) and that all the biomass observations were ignored (Figure 6b). In contrast, when we used the DART adaptive inflation (with an inflation standard deviation of 0.6 and inflation damping of 0.9; see J. L. Anderson, 2009), the model error is increased in response to the relative model-data mismatch (section 2.2). This results in the observations lying within the outlier threshold cut-off of three standard deviations and the filter responds in a similar way to the OSSE (Figures 6c and 6d middle panels). There is large reduction in the RMSE in LAI and biomass for the adaptive inflation experiment (red curve), both relative to the free-run (black curve) and no-adaptive inflation experiment (green curves in Figures 6e and 6f bottom panels). For LAI, RMSE over the 10-year period of the assimilation is reduced by over 50%, from 0.93 to 0.44  $\text{m}^2/\text{m}^2$  and by 70% for biomass from 1,376 to 418  $\text{g C}/\text{m}^2$  (Table 2). It is important to remember that although the RMSE is calculated against observations that are used in the assimilation, as the true state of the system is unknown, this is the RMSE for the model *forecasts*, not the *analyses*. Put a different way, the RMSE describes the error between the model and the observations before the observations are assimilated and the model ensemble is updated. For biomass, this overall error includes the first 3 years of the assimilation period when observations lay outside the outlier threshold and the adaptive inflation algorithm was increasing forecast error variance. If this was treated as a *burn-in* period for the DA system (or techniques were implemented to increase forecast error variance prior to the time with the first observations), this error would be reduced still further. The apparent systematic shift in Figure 6 is symptomatic of (1) plotting the forecast and (2) the relatively infrequent observations relative to the time step of the model, combined with the modeled trajectory of leaf phenology being different from observed. It is important that we plot the forecast and calculate the RMSE based on this so as not to be assimilating observations we are using for our skill metric. However, to demonstrate the effectiveness of the assimilation we have also plotted the analysis on Figure 6c. Note that although the analysis is much closer to the observations, it is still lagged behind the observations. This reflects that it is always being moved toward the observations, but due to observation error and error in the forward model they never correspond exactly. This is common to all sequential DA systems, but the effect would be less apparent if more frequent observations were being assimilated.

The success of the EAKF illustrates a clear application to solve biases in initial conditions resulting from incorrect model spin-up. In this example, we are explicitly trusting the data more than the model, however we see great potential for diagnostic studies; we can use the EAKF to ensure that the model will estimate C, water, and energy budgets which are consistent with available data. This could help constrain changes in the historical global land C budget due to land use and land cover change processes not yet robustly implemented in models given the wealth of global satellite observations and other LSM benchmarks over the last several decades (Hoffman et al., 2017; Y. Q. Luo et al., 2012).

### 3.3.2. Longer-Term Forecasts

The DA also had a positive effect on model forecasts over longer time periods (Figure 7). We assessed this with our final experiment in which we assimilated biomass and LAI for the first 5 years, and then used the ensemble at the end of this time period to provide the initial conditions for a longer-term forecast over the next 5 years when no observations were assimilated (Figure 7). Given what we learnt in section 3.1 (Figure 1) we might expect the forecast ensemble to trend back toward the free-run ensemble, rapidly in the case of LAI, and somewhat more slowly for biomass. The reality is a more complex picture. For biomass, which has been updated to be distant from the free-run, there is little evidence of a trend in the ensemble

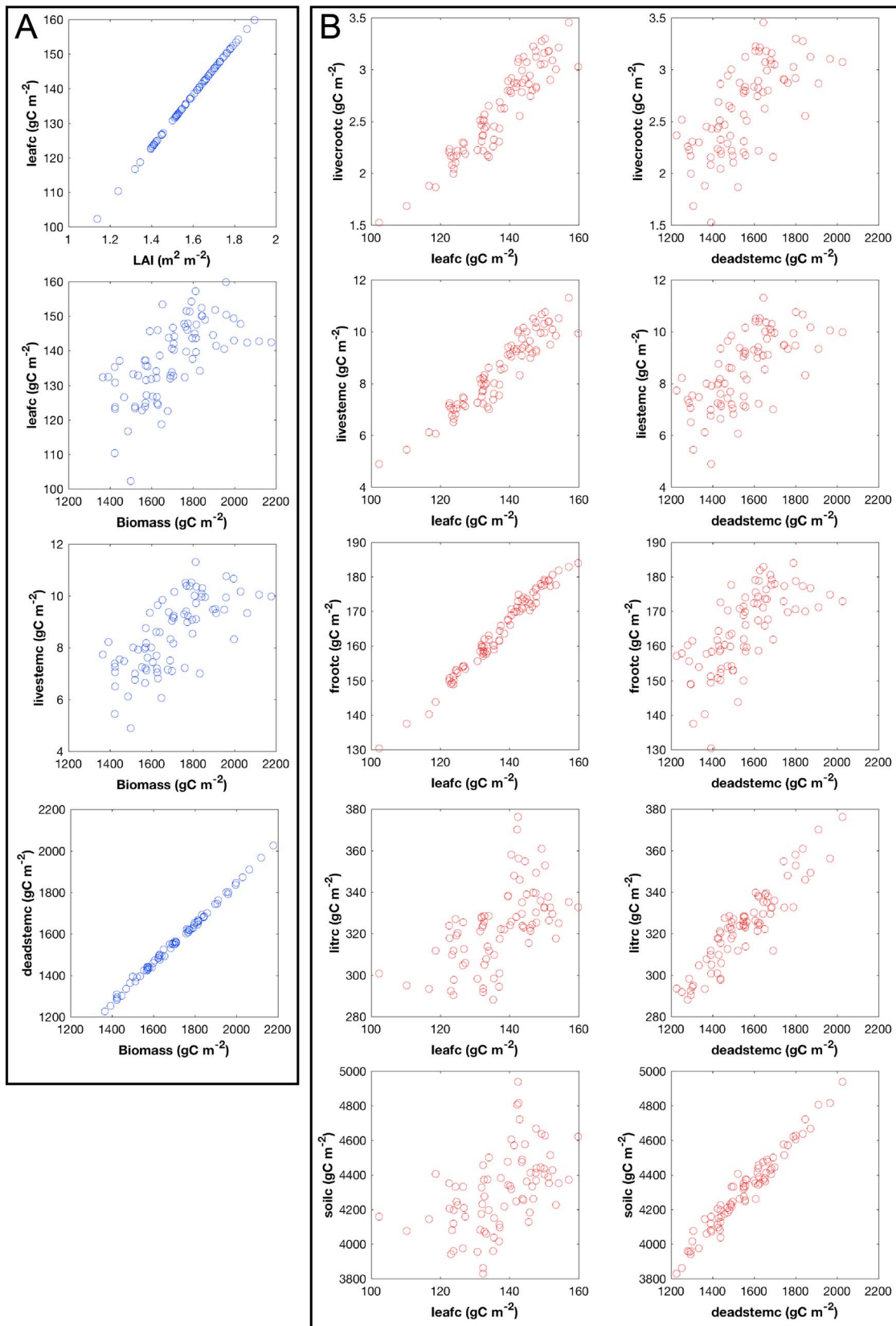


**Figure 7.** Five-year forecasts of LAI and biomass, initialized by 5 years of assimilation of LAI and biomass observations using adaptive inflation. LAI = leaf area index.

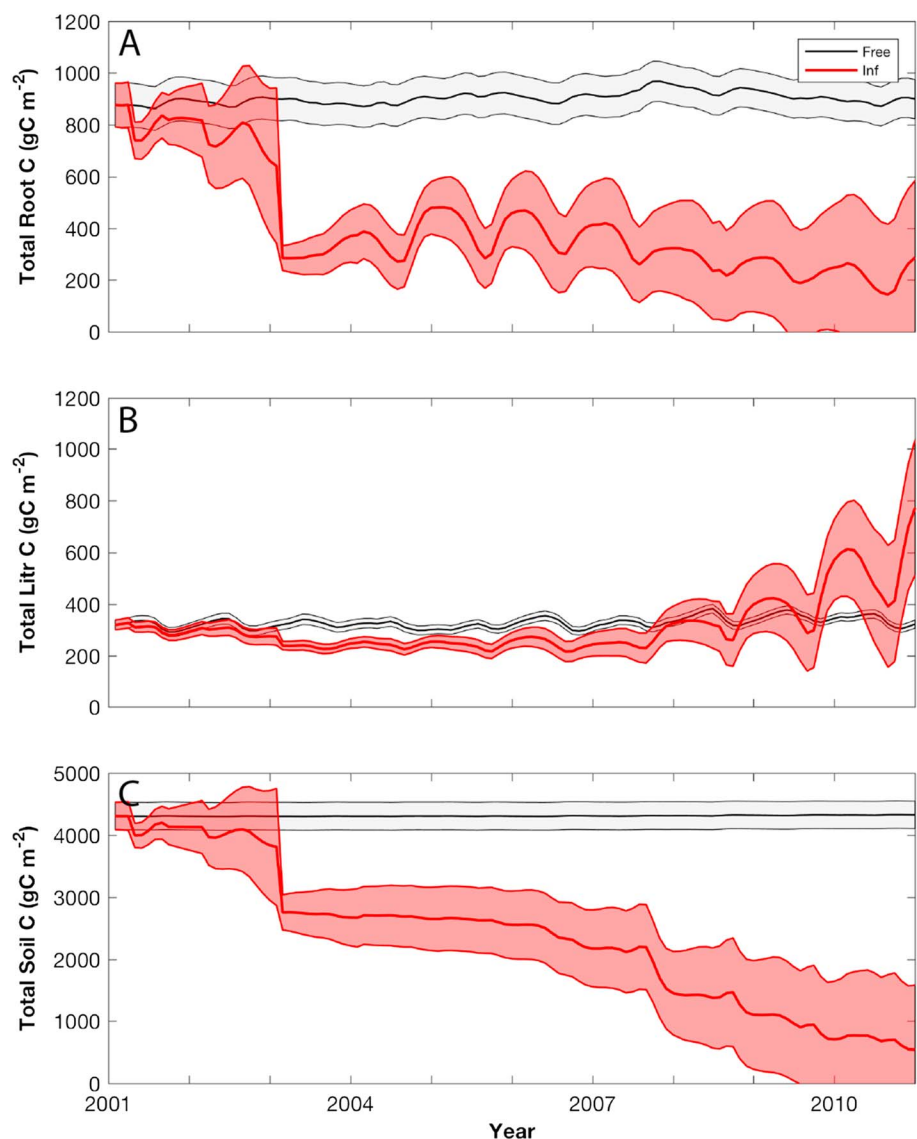
mean, but there is a gradual increase in model forecast error variance and by the end of the experiment this is approximately equivalent to the free-run (Figure 7b). This is good news from a forecasting perspective as it enables very accurate predictions of biomass for 5 years, with an RMSE in this period of only  $51 \text{ g C/m}^2$ , which is less than 5% of the RMSE of the original free-run during this period (Table 2). The longer-term LAI forecast provides further, unexpected additional evidence of the potential importance of initial conditions for forecasting with CLM and how the DA system can update the model state into unusual space. In this case we have to plot individual ensemble members to begin to diagnose what happens during the forecast period. This is because plotting the ensemble mean and variance alone hides that the ensemble splits from very similar initial LAI values to a situation in which approximately half the ensemble gradually decreases toward zero and half trends back toward the free-run although there is considerable variability in the rate of this trend for individual ensemble members. The only clear predictor at the start of the forecast period as to whether an ensemble member's LAI will decline or recover is fine root C. Interestingly, it is the ensemble members with an initial high fine root C for which LAI declines. These ensemble members also have ratios between fine root C and leaf C that are very variable, and deviate away from the linear relationship found between these two variables in the free-run ensemble. We discuss in the following section the implications of updating unobserved variables. Also, it should be noted that as the DA system does not change the model structure directly, the seasonal cycle generated in LAI by assimilating observations with this seasonal cycle is not continued.

### 3.3.3. Unobserved State Variables

A great advantage of ensemble DA techniques, such as the EAKF, is that they enable unobserved model state variables to be updated by using statistical information from the ensemble describing their covariance with the observed variables. Figure 8a shows scatter plots between observation values and model state variables (e.g., leaf C content, *leafc*, see Table S1) used to calculate these observations. The translation between model space and observation space is carried out using an observation operator (section 2.3). The effectiveness of observation operators is in part related to how linear they are, as once increments have been calculated by the filter in observation space, these are translated back into model space through a linear regression between the observations and model state variables. Unsurprisingly, given the simplicity of the observation operator, the correlation between observations and model state variables is very high for leaf C and LAI, and dead stem C and biomass, and so the observations lead to effective updates of these model variables.



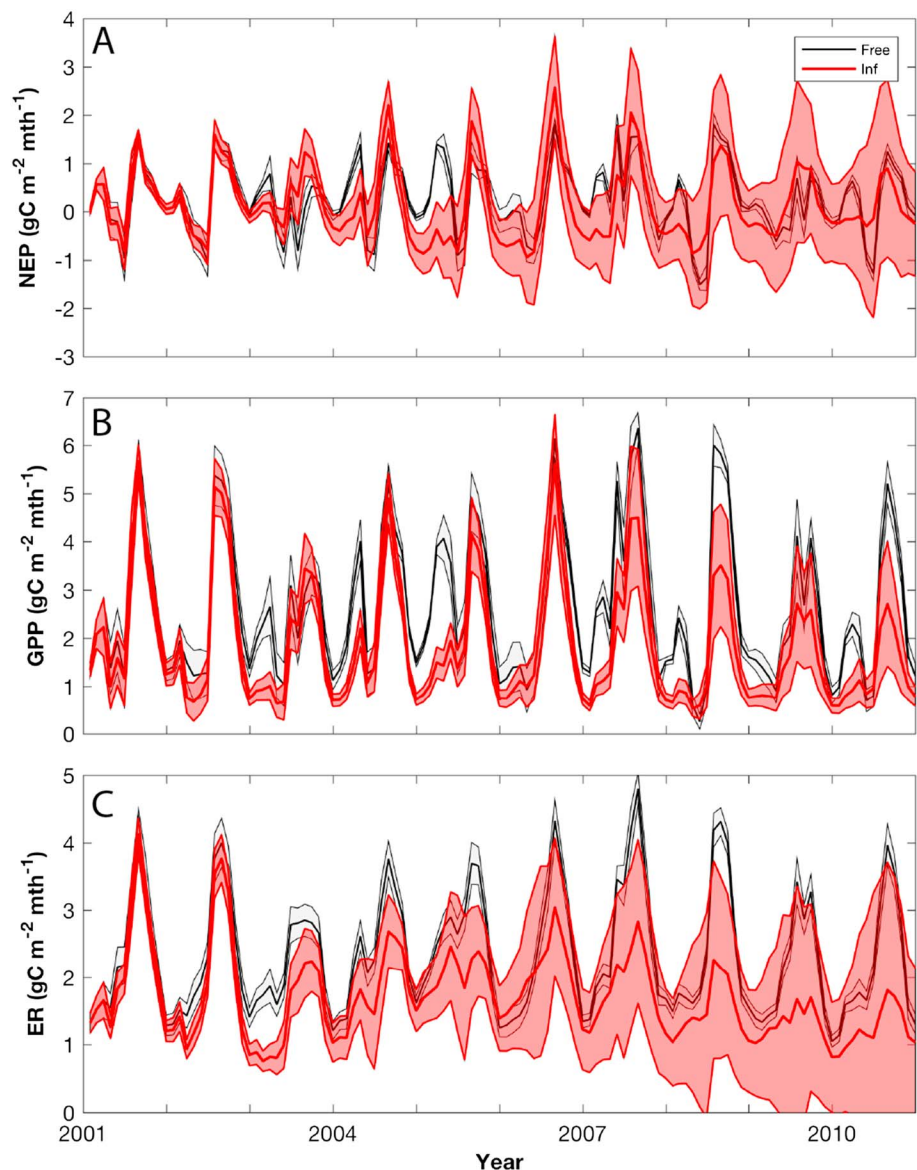
**Figure 8.** (a) Scatter plots of forecast ensemble members between observations and model state variables used to calculate these observed values. (b) Scatter plots between the dominant *observed* model state variables and other unobserved variables. LAI = leaf area index.



**Figure 9.** Unobserved state variables are updated by the Ensemble Adjustment Kalman Filter. Updates to fine root C, live coarse root C, and dead coarse root C causes changes in total root C, while updates in the three vertically resolved litter C pools and three vertically resolved soil C pools causes changes in total litter C and soil C, respectively. C = carbon.

The stationary allometries and turnover rates used to partition C between the different C pools within CLM (see Montané et al., 2017) result in strong correlations between these dominant observed model state variables and other unobserved variables (e.g., fine root C, *frootc*, see Table S1 and Figure 8b). By including these other C pools in the DART state vector, they are updated at the same time as the observed variables using these regression equations. The impact of these updates can be seen on all three of the large C pools for roots, litter, and soil (Figure 9). Root C provides a good illustration of the impact of assimilating both LAI and biomass when compared with Figure 7. Over the first 2 years we see the influence of assimilating LAI, with a modest decrease in root C when the first LAI observations are assimilated, and the appearance of a seasonal cycle in the second year. This then persists over the rest of the assimilation period. The most marked update occurs simultaneously with the large change in biomass in 2003, and is a reduction of a similar, relative magnitude to that seen in biomass. The impacts of assimilating LAI and biomass on litter C are less—note the weaker correlations with litter in Figure 8. Nevertheless, there is a decrease litter C in early 2003, and there is a more apparent seasonal cycle. For the first few years the soil C forecast follows a very similar trajectory to biomass and root C, although the relative reduction in 2003 is less. However, after 2003 there is no enhanced





**Figure 10.** Impact of assimilation on ecosystem level carbon fluxes. NEP = net ecosystem production; GPP = gross primary productivity; ER = ecosystem respiration.

seasonal cycle as we see in the other pools, which reflects three factors: (1) the much weaker correlation between LAI and soil C (Figure 8); (2) the contrasting size of these C pools; and (3) the rather indirect physical relationship between them in the model. Following the large update in these unobserved C pools in 2003, corresponding with the similarly dramatic change in biomass at this time, ensemble spread begins to increase. This is a relatively small increase over the next 4 years, and then increases rapidly. We do not see this increase in ensemble spread in the observed variables, and hypothesize that one reason for this could be an undesirable result of the inflation applied.

For all three of these C pools there is an apparent degradation of the forecast over the last 3 years of the simulation, with ensemble members drifting to the lower bounds of possible values for root C and soil C; litter C also exhibits an unrealistic increase and seasonal cycle. This is not a feature for the observed variables, and so must be introduced through errors in the relationship between the observed and unobserved C pools used by the filter to calculate the updates. This is compounded by potentially too much inflation being added to all variables updated by the EAKF. The introduction of errors is either caused by small sample size, or the limitations of having to use a linear relationship. This is compounded by the very strong correlations between the

variables (itself potentially a function of small ensemble size and resulting sample error), causing unrealistically large updates in the unobserved variables. Comparisons between LSMs and global observational networks often highlight inappropriately high correlations between C pools and fluxes in models (Shao et al., 2013). This problem of overly strong correlations between variables is typically dealt with using localization in an ensemble DA context (section 2.2). Typically, localization is carried out across space, but J. L. Anderson (2012) shows that its use may be appropriate between an observations and colocated variable of a different quantity, and in future studies we will test this possibility with CLM. Such large changes in leaf area, biomass, and other C pools would be expected to result quickly in feedbacks to C fluxes (Figure 10). Here in the desert Southwest United States, net ecosystem production (positive indicates C uptake) is small and shows considerable interannual variability. Patterns of GPP and ER are more easily interpreted, and we begin to see divergence between the assimilation run and free-run by the end of the second year, by which time the adaptive inflation routine has adjusted the model states to more closely reflect the observations. The reduction in LAI during the winter and spring periods following assimilation results in much lower GPP during the first half of the year; however, GPP remains as high as in the free-run in the latter part of each year, when LAI observations, assimilation run and the free-run are more in agreement. In the latter part of the assimilation period from 2008 onward we see a marked reduction in mean GPP for the assimilation run relative to the free-run, but with increasing uncertainty. The overall reduction in productivity causes decreased ER during the first few years of the assimilation period. The mean ER continues to decline throughout the rest of the assimilation period. This is because the filter continues to update the large root, litter, and soil C pools (the principal sources for ER) causing large—and unrealistic—increases in ensemble spread for these pools, that result in the spread of ER increasing. This is possibly a result of overly simplistic allocation schemes within the model causing spuriously high relationships between observed and unobserved C pools (Shao et al., 2013), and unrealistic updates, possibly compounded through too large inflation values being applied to the observed and unobserved variables updated by the EAKF. This will be addressed with localization in future developments of the DA scheme. Nevertheless, it is clear during the first half of the assimilation period, assimilating LAI and biomass observations had large and direct impacts on simulated C fluxes, reducing simulated productivity in this case.

#### 4. Conclusion and Perspective

The temporal autocorrelation inherent in ecosystem processes means that C stocks and fluxes estimated by LSMs like CLM4.5 are sensitive to initial conditions, so improved estimates of current stocks and fluxes allow for better predictions of future stocks and fluxes. This analysis demonstrates that state DA is an important and tractable method for achieving more accurate initial conditions for LSMs. Assimilation of total biomass and LAI led to large reductions in errors in both the assimilation period and well into the forecast period. We also observed improved representation of the seasonal cycle in LAI, however as we do not alter the model itself this improved seasonal cycle is not maintained during the longer-term forecast. Similar large changes to LAI after DA have been reported by other authors. For example, Viskari et al. (2015) when they used an EAKF to assimilate LAI observations into the ED2 model. Barbu et al. (2011) also found that assimilating observations using a Simplified Extended Kalman Filter led to large LAI corrections, particularly during the senescence periods when their model overestimated the LAI values, and also had a delay in the start of the growing period. When using this system at the regional scale, Albergel et al. (2017) found that although stronger corrections were applied to LAI than to soil moisture, high root-mean-square difference remained between observations and analysis attributable to model error. The way we have implemented the DA, stocks, and fluxes that are not observed are altered to reconcile the model logic with the observations, which allows us to infer the dynamics of unobserved stocks and fluxes with more confidence than a conventional (free) model run. However, because some land surface processes are more closely tied to dynamic changes in meteorology, the influence of updating model states through DA is dependent roughly on the mean residence time of the stock and, in the case of fluxes, the effect diminished very quickly. Adaptive inflation techniques previously implemented in meteorological and climate models (Altaf et al., 2013; J. L. Anderson, 2009) can be used in terrestrial models to overcome very large initial mismatches between model predictions and observations even when these mismatches are the apparent result of model structural errors. It is worth noting that DA is not a replacement for model development, however it speeds the identification of deficiencies in model structure that are most consequential.

A number of improvements would help achieve the goal of long-term, global DA. Using the technique presented here requires the development of a multidecade global ensemble climate drivers which would allow us to generate the model ensemble spread. Given the utility of the long-lived terrestrial C pools, investment in developing estimates and uncertainties for these observations would likely yield progress. Also, some development of the model and DA system itself would be useful. While the update of states that were not observed is powerful, some variables are correlated in the model through artifacts; evaluation of the covariance between model states is required and where appropriate minimized using localization. Scale mismatches between grid cells and observations could be minimized by linking observations to particular PFTs rather than to grid cell means.

While we have illustrated the technique in one type of ecosystem, our analysis suggests two important potential uses for DA coupled with LSMs at broader geographic scales. First, DA can be used to improve diagnostic uses of models. Given our understanding of the land surface and our available historical observations, DA constrains the possible model estimates for pools and fluxes that we cannot observe. This feature allows us to improve historical C, water or energy budgets using gridded regional or global observations if uncertainty is estimated appropriately. Second, forecasts of large C pools with the longest residence time are improved for longer forecast lead times. If applied globally, DA could provide a more robust decadal or centennial projection of how much C is stored on land.

#### Acknowledgments

We gratefully acknowledge Yi Liu for providing us access to the VOD biomass data set. Both the Community Land Model and the Data Assimilation Research Testbed are open source software, available from remote repositories. CLM: <https://github.com/ESCOMP/ctsm>; DART: <http://doi.org/10.5065/D6WQ0202> (Direct download address May 2018 <https://www2.cisl.ucar.edu/software/dart/download>). The supporting material includes a ds01.zip file containing the following: An input\_data directory containing:

1. Forcing files directory: Containing the 80 member atmospheric data ensemble used to force CLM in site mode for the study location.
2. NMPJ\_monthly\_obs directory: Containing monthly/annual observations for LAI/biomass used in this study.
3. Surface\_data\_files directory: Containing files required to describe the land surface for the study location.
4. Setup\_script directory: Containing a shell script to set up a multi-instant version of a CLM site level run, and couple that with DART and datm.streams template files required by this script. An output\_data directory containing all the CLM-DART output used in the analysis of this paper. A Matlab script that in conjunction with DART diagnostics Matlab functions will plot the figures used in this manuscript.

#### References

- Aber, J. D., & Fédéré, C. A. (1992). A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production in temperate and boreal forest ecosystems. *Oecologia*, 92(4), 463–474. <https://doi.org/10.1007/BF00317837>
- 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., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, 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>
- Altaf, M. U., Butler, T., Luo, X., Dawson, C., Mayo, T., & Hoteit, I. (2013). Improving short-range ensemble Kalman storm surge forecasting using robust adaptive inflation. *Monthly Weather Review*, 141(8), 2705–2720. <https://doi.org/10.1175/MWR-D-12-00310.1>
- Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). The data assimilation research testbed: A community facility. *Bulletin of the American Meteorological Society*, 90(9), 1283–1296. <https://doi.org/10.1175/2009BAMS2618.1>
- 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:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2)
- Anderson, J. L. (2007a). An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus A*, 59(2), 210–224. <https://doi.org/10.1111/j.1600-0870.2006.00216.x>
- Anderson, J. L. (2007b). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, 230(1–2), 99–111. <https://doi.org/10.1016/j.physd.2006.02.011>
- Anderson, J. L. (2009). Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus A*, 61(1), 72–83. <https://doi.org/10.1111/j.1600-0870.2008.00361.x>
- 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/MWR-D-11-00013.1>
- Anderson-Teixeira, K. J., Delong, J. P., Fox, A. M., Brese, D. A., & Litvak, M. E. (2010). Differential responses of production and respiration to temperature and moisture drive the carbon balance across a climatic gradient in New Mexico. *Global Change Biology*, 17(1), 410–424.
- Antonarakis, A. S., Saatchi, S. S., Chazdon, R. L., & Moorcroft, P. R. (2011). Using Lidar and Radar measurements to constrain predictions of forest ecosystem structure and function. *Ecological Applications: A Publication of the Ecological Society of America*, 21, 1120–1137. <https://doi.org/10.1890/1052-0774.1120>
- Atlas, R. (1997). Atmospheric observations and experiments to assess their usefulness in data assimilation (special issue on data assimilation in meteorology and oceanography: Theory and practice). *Journal of the Meteorological Society of Japan. Ser. II*, 75(1B), 111–130. [https://doi.org/10.2151/jmsj1965.75.1B\\_111](https://doi.org/10.2151/jmsj1965.75.1B_111)
- 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-Ags land surface model: Grassland case study. *Biogeosciences*, 8(7), 1971–1986. <https://doi.org/10.5194/bg-8-1971-2011>
- Barron-Gafford, G. A., Sanchez-Cañete, E. P., Minor, R. L., Hendryx, S. M., Lee, E., Sutter, L. F., et al. (2017). Impacts of hydraulic redistribution on grass-tree competition vs facilitation in a semi-arid savanna. *The New Phytologist*, 215(4), 1451–1461. <https://doi.org/10.1111/nph.14693>
- Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1–4), 11–29. [https://doi.org/10.1016/S0022-1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8)
- Biederman, J. A., Scott, R. L., Arnone, J. A. III, Jasoni, R. L., Litvak, M. E., Moreo, M. T., et al. (2017). Shrubland carbon sink depends upon winter water availability in the warm deserts of North America. *Agric. For. Meteorol.*, 249, 407–419.
- Biederman, J. A., Scott, R. L., Goulden, M. L., Vargas, R., Litvak, M. E., Kolb, T. E., & Burns, S. P. (2016). Terrestrial carbon balance in a drier world: The effects of water availability in southwestern North America. *Global Change Biology*, 22, 1867–1879. <https://doi.org/10.1111/gcb.13222>
- 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. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, 320(5882), 1444–1449. <https://doi.org/10.1126/science.1155121>

- Bonan, G. B., Levis, S., Kergoat, L., & Oleson, K. W. (2002). Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles*, 16(2), 1021. <https://doi.org/10.1029/2000GB001360>
- Bonan, G. B., Oleson, K. W., Fisher, R. A., Lasslop, G., & Reichstein, M. (2012). Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the community land model version 4: Community land model canopy scaling. *Journal of Geophysical Research*, 117(G2), G02026. <https://doi.org/10.1029/2011JG001913>
- Burgers, G., Gerrit, B., van Leeuwen, P. J., & Geir, E. (1998). Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126(6), 1719–1724. [https://doi.org/10.1175/1520-0493\(1998\)126<1719:ASITEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2)
- Chojnacki, D. C., Heath, L. S., & Jenkins, J. C. (2014). Updated generalized biomass equations for North American tree species. *Forestry*, 87, 129–151. <https://doi.org/10.1093/forestry/cpt053>
- Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., & Pringle, C. (2001). Ecological forecasts: An emerging imperative. *Science*, 293(5530), 657–660. <https://doi.org/10.1126/science.293.5530.657>
- Dahlin, K. M., Fisher, R. A., & Lawrence, P. J. (2015). Environmental drivers of drought deciduous phenology in the community land model. *Biogeosciences*, 12(16), 5061–5074. <https://doi.org/10.5194/bg-12-5061-2015>
- Daley, R. (1993). *Atmospheric data analysis*. Cambridge: Cambridge University Press.
- DART (2017). The IMPROVED Data Assimilation Research Testbed - DART Manhattan. Retrieved July 20, 2018, from <http://www.image.ucar.edu/DARes/DART/>
- Davies, H. C., & Turner, R. E. (1977). Updating prediction models by dynamical relaxation: An examination of the technique. *Quarterly Journal of the Royal Meteorological Society*, 103(436), 225–245. <https://doi.org/10.1002/qj.49710343602>
- Duarte, H. F., Raczka, B. M., Ricciuto, D. M., Lin, J. C., Koven, C. D., Thornton, P. E., et al. (2017). Evaluating the community land model (CLM4.5) at a coniferous forest site in northwestern United States using flux and carbon-isotope measurements. *Biogeosciences*, 14(18), 4315–4340. <https://doi.org/10.5194/bg-14-4315-2017>
- 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), 10143. <https://doi.org/10.1029/94JC00572>
- Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., et al. (2009). The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. *Agricultural and Forest Meteorology*, 149(10), 1597–1615. <https://doi.org/10.1016/j.agrformet.2009.05.002>
- Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., et al. (2006). Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison. *Journal of Climate*, 19(14), 3337–3353. <https://doi.org/10.1175/JCLI3800.1>
- Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, 27(2), 511–526. <https://doi.org/10.1175/JCLI-D-12-00579.1>
- Hoffman, F. M., Koven, C. D., Keppel-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-0186). 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(3), 796–811. [https://doi.org/10.1175/1520-0493\(1998\)126<0796:DAUAEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2)
- 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>
- Jenkins, J. C., Chojnacki, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, 49(1), 12–35.
- Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge: Cambridge University Press.
- Karspeck, A. R., Yeager, S., Danabasoglu, G., Hoar, T., Collins, N., Raeder, K., et al. (2013). An ensemble adjustment Kalman filter for the CCSM4 ocean component. *Journal of Climate*, 26(19), 7392–7413. <https://doi.org/10.1175/JCLI-D-12-00402.1>
- Keenan, T. F., Davidson, E., & Moffat, A. M. (2012). Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. *Global Change Biology*, 18(8), 2555–2569. <https://doi.org/10.1111/j.1365-2486.2012.02684.x>
- Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., et al. (2013). The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. *Biogeosciences*, 10(11), 7109–7131. <https://doi.org/10.5194/bg-10-7109-2013>
- Kumar, S. V., Jasinski, M., Mocko, D., Rodell, M., Borak, J., Li, B., et al. (2018). NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. *Journal of Hydrometeorology*. <https://doi.org/10.1175/JHM-D-17-0125.1>
- Lawrence, D., & Fisher, R. A. (2013). The community land model philosophy: model development and science applications. *iLeaps Newsletter*, 13, 13–16.
- Lawrence, P. J., & Chase, T. N. (2007). Representing a new MODIS consistent land surface in the community land model (CLM 3.0). *Journal of Geophysical Research*, 112(G1), G01023. <https://doi.org/10.1029/2006JG000168>
- Liu, Y. Y., de Jeu, R. A. M., McCabe, M. F., Evans, J. P., & van Dijk, A. I. (2011). Global long-term passive microwave satellite-based retrievals of vegetation optical depth. *Geophysical Research Letters*, 38, L18402. <https://doi.org/10.1029/2011GL048684>
- Liu, Y. Y., van Dijk, A. I. J. M., de Jeu, R. A. M., Canadell, J. G., McCabe, M. F., Evans, J. P., & et al. (2015). Recent reversal in loss of global terrestrial biomass. *Nature Climate Change*, 5(5), 470–474. <https://doi.org/10.1038/nclimate2581>
- Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., et al. (2016). Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles*, 30, 40–56. <https://doi.org/10.1002/2015GB005239>
- Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., et al. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications: A Publication of the Ecological Society of America*. Retrieved from [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list\\_uids=13678945382857759521related:ls0\\_1lg1b0J](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=13678945382857759521related:ls0_1lg1b0J)
- Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., et al. (2012). A framework for benchmarking land models. *Biogeosciences*, 9(10), 3857–3874. <https://doi.org/10.5194/bg-9-3857-2012>
- Luyssaert, S., Jammot, M., Stoy, P. C., Estel, S., Pongratz, J., Ceschia, E., et al. (2014). Land management and land-cover change have impacts of similar magnitude on surface temperature. *Nature Climate Change*, 4(5), 389–393. <https://doi.org/10.1038/nclimate2196>
- Lynch, P. (2006). *The emergence of numerical weather prediction: Richardson's dream*. Cambridge: Cambridge University Press.
- MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., & Ciais, P. (2015). Using satellite data to improve the leaf phenology of a global terrestrial biosphere model. *Biogeosciences*, 12(23), 7185–7208. <https://doi.org/10.5194/bg-12-7185-2015>



- Mao, J., Ricciuto, D. M., Thornton, P. E., Warren, J. M., King, A. W., Shi, X., et al. (2016). Evaluating the community land model in a pine stand with shading manipulations and  $^{13}\text{C}$  labeling. *Biogeosciences*, 13(3), 641–657. <https://doi.org/10.5194/bg-13-641-2016>
- Masutani, M., Woollen, J. S., Lord, S. J., Emmitt, G. D., Kleespies, T. J., Wood, S. A., & Campana, K. A. (2010). Observing system simulation experiments at the National Centers for Environmental Prediction. *Journal of Geophysical Research*, 115(D7), D07101. <https://doi.org/10.1029/2009JD012528>
- Matheny, A. M., Bohrer, G., Stoy, P. C., Baker, I. T., Black, A. T., Desai, A. R., et al. (2014). Characterizing the diurnal patterns of errors in the prediction of evapotranspiration by several land-surface models: An NACP analysis. *Journal of Geophysical Research: Biogeosciences*, 119, 1458–1473. <https://doi.org/10.1002/2014JG002623>
- McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., et al. (2017). A land data assimilation system for sub-Saharan Africa food and water security applications. *Sci Data*, 4, 170012. <https://doi.org/10.1038/sdata.2017.12>
- Mitchell, H. L., Houtekamer, P. L., & Pellerin, G. (2002). Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Monthly Weather Review*, 130(11), 2791–2808. [https://doi.org/10.1175/1520-0493\(2002\)130<2791:ESBAME>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2791:ESBAME>2.0.CO;2)
- Mitchell, K. E. (2004). The multi-institution North American land data assimilation system (NLDAS): Utilizing multiple GCIIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research*, 109(D7), D07S90. <https://doi.org/10.1029/2003JD003823>
- Montané, F., Fox, A. M., Arellano, A. F., MacBean, N., Alexander, M. R., Dye, A., et al. (2017). Evaluating the effect of alternative carbon allocation schemes in a land surface model (CLM4.5) on carbon fluxes, pools, and turnover in temperate forests. *Geoscientific Model Development*, 10, 3499–3517. <https://doi.org/10.5194/gmd-10-3499-2017>
- Moore, D., Hu, J., Sacks, W. J., Schimel, D. S., & Monson, R. K. (2008). Estimating transpiration and the sensitivity of carbon uptake to water availability in a subalpine forest using a simple ecosystem process model informed by measured net  $\text{CO}_2$  and  $\text{H}_2\text{O}$  fluxes. *Agricultural and Forest Meteorology*, 148(10), 1467–1477. <https://doi.org/10.1016/j.agrformet.2008.04.013>
- NCAR (2018). CLM Bibliography. Retrieved 15 October 2018, from [http://www.cesm.ucar.edu/models/cesm2/land/clm\\_bibliography.htm](http://www.cesm.ucar.edu/models/cesm2/land/clm_bibliography.htm)
- Oleson, K., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Kovan, C. D. et al. (2013). Technical description of version 4.5 of the Community Land Model (CLM) NCAR/TN-503+STR NCAR Technical Note.
- ORNL DAAC (2008). MODIS Collection 5 land products global subsetting and visualization tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed December 03, 2016. Subset obtained for MOD15A2 product at 34.438N, 106.2385W, time period: 2000-02-18 to 2016-11-16, and subset size:  $7 \times 7$  km. <https://doi.org/10.3334/ORNLDAAC/1241>
- Owe, M., de Jeu, R., & Walker, J. (2001). A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, 39(8), 1643–1654. <https://doi.org/10.1109/36.942542>
- Pitman, A. J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, 23(5), 479–510. <https://doi.org/10.1002/joc.893>
- Raeder, K., Anderson, J. L., Collins, N., Hoar, T. J., Kay, J. E., Lauritzen, P. H., & Pincus, R. (2012). DART/CAM: An ensemble data assimilation system for CESM atmospheric models. *Journal of Climate*, 25(18), 6304–6317. <https://doi.org/10.1175/JCLI-D-11-00395.1>
- Raunkiaer, C. (1934). *The life forms of plants and statistical plant geography; being the collected papers of C. Raunkiaer*. Oxford: Clarendon Press.
- Raupach, M. R., Rayner, P. J., Barrett, D. J., Defries, R. S., Heimann, M., Ojima, D. S., et al. (2005). Model-data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications. *Global Change Biology*, 11(3), 378–397. <https://doi.org/10.1111/j.1365-2486.2005.00917.x>
- Rodell, M., Houser, P. R., Jambor, U., Gottschalk, J., Mitchell, K., Meng, C.-J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, 85(3), 381–394. <https://doi.org/10.1175/BAMS-85-3-381>
- Rosolem, R., Hoar, T., Arellano, A., Anderson, J. L., Shuttleworth, W. J., Zeng, X., & Franz, T. E. (2014). Translating aboveground cosmic-ray neutron intensity to high-frequency soil moisture profiles at sub-kilometer scale. *Hydrology and Earth System Sciences*, 18(11), 4363–4379. <https://doi.org/10.5194/hess-18-4363-2014>
- Saatchi, S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., et al. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences of the United States of America*, 108(24), 9899–9904. <https://doi.org/10.1073/pnas.1019576108>
- Sawada, Y., & Koike, T. (2014). Simultaneous estimation of both hydrological and ecological parameters in an ecohydrological model by assimilating microwave signal. *J. Geophys. Res. D: Atmos.*, 119(14), 8839–8857. <https://doi.org/10.1002/2014JD021536>
- Sawada, Y., Koike, T., & Walker, J. P. (2015). A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics: LDAS for ecohydrological model. *Journal of Geophysical Research: Atmospheres*, 120, 5910–5930. <https://doi.org/10.1002/2014JD022895>
- Schwartz, C. S., Romine, G. S., Sobash, R. A., Fossell, K. R., & Weisman, M. L. (2015). NCAR's experimental real-time convection-allowing ensemble prediction system. *Weather and Forecasting*, 30(6), 1645–1654. <https://doi.org/10.1175/WAF-D-15-0103.1>
- Shao, P., Zeng, X., Moore, D. J. P., & Zeng, X. (2013). Soil microbial respiration from observations and Earth system models. *Environmental Research Letters: ERL [Web Site]*, 8(3), 034034. <https://doi.org/10.1088/1748-9326/8/3/034034>
- Smith, W. K., Biederman, J. A., Scott, R. L., Moore, D. J. P., He, M., Kimball, J. S., et al. (2018). Chlorophyll fluorescence better captures seasonal and interannual gross primary productivity dynamics across dryland ecosystems of southwestern North America. *Geophys. Res. Lett.*, 45(2), 748–757. <https://doi.org/10.1002/2017GL075922>
- Thornton, P. E., & Rosenbloom, N. A. (2005). Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecological Modelling*, 189(1–2), 25–48. <https://doi.org/10.1016/j.ecolmodel.2005.04.008>
- Traore, A. K., Ciais, P., Vuichard, N., Poulter, B., Viovy, N., Guimberteau, M., et al. (2014). Evaluation of the ORCHIDEE ecosystem model over Africa against 25 years of satellite-based water and carbon measurements. *Journal of Geophysical Research: Biogeosciences*, 119, 1554–1575. <https://doi.org/10.1002/2014JG002638>
- Trudinger, C. M., Raupach, M. R., Rayner, P. J., Kattge, J., Liu, Q., Pak, B., et al. (2007). OptIC project: An intercomparison of optimization techniques for parameter estimation in terrestrial biogeochemical models. *Journal of Geophysical Research*, 112(G2), G02027. <https://doi.org/10.1029/2006JG000367>
- Verma, M., Friedl, M. A., Law, B. E., Bonal, D., Kiely, G., Black, T. A., et al. (2015). Improving the performance of remote sensing models for capturing intra- and inter-annual variations in daily GPP: An analysis using global FLUXNET tower data. *Agricultural and Forest Meteorology*, 214–215, 416–429. <https://doi.org/10.1016/j.agrformet.2015.09.005>
- Viskari, T., Hardiman, B., Desai, A. R., & Dietze, M. C. (2015). Model-data assimilation of multiple phenological observations to constrain and predict leaf area index. *Ecological Applications*, 25(2), 546–558. <https://doi.org/10.1890/14-0497.1>



- Wang, Y.-P., Trudinger, C. M., & Enting, I. G. (2009). A review of applications of model–data fusion to studies of terrestrial carbon fluxes at different scales. *Agricultural and Forest Meteorology*, 149(11), 1829–1842. <https://doi.org/10.1016/j.agrformet.2009.07.009>
- Wieder, W. R., Knowles, J. F., Blanken, P. D., Swenson, S. C., & Suding, K. N. (2017). Ecosystem function in complex mountain terrain: Combining models and long-term observations to advance process-based understanding. *J. Geophys. Res. Biogeosci.*, 122(4), 825–845. <https://doi.org/10.1002/2016JG003704>
- Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., & Kattge, J. (2009). Improving land surface models with FLUXNET data. *Biogeosciences*, 6(7), 1341–1359. <https://doi.org/10.5194/bg-6-1341-2009>
- Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., & Kurpius, M. R. (2005). An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology*, 11(1), 89–105. <https://doi.org/10.1111/j.1365-2486.2004.00891.x>
- Xia, J. Y., Luo, Y. Q., Wang, Y.-P., Weng, E. S., & Hararuk, O. (2012). A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, 5(5), 1259–1271. <https://doi.org/10.5194/gmd-5-1259-2012>
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research*, 117(D3), D03109. <https://doi.org/10.1029/2011JD016048>
- Zhang, Y.-F., Hoar, T. J., Yang, Z.-L., Anderson, J. L., Toure, A. M., & Rodell, M. (2014). Assimilation of MODIS snow cover through the data assimilation research testbed and the community land model version 4: SNOW DATA ASSIMILATION. *Journal of Geophysical Research, D: Atmospheres*, 119(12), 7091–7103. <https://doi.org/10.1002/2013JD021329>
- Zobitz, J. M., Moore, D. J. P., Quaife, T., Braswell, B. H., Bergeson, A., Anthony, J. A., & Monson, R. K. (2014). Joint data assimilation of satellite reflectance and net ecosystem exchange data constrains ecosystem carbon fluxes at a high-elevation subalpine forest. *Agricultural and Forest Meteorology*, 195–196, 73–88. <https://doi.org/10.1016/j.agrformet.2014.04.011>