Title
Large Divergence of Projected High Latitude Vegetation Composition and Productivity Due To Functional Trait Uncertainty

Journal
Earth's Future, 12(8)

ISSN
2328-4277

Authors
Liu, Yanlan
Holm, Jennifer A
Koven, Charles D
et al.

Publication Date
2024-08-01

DOI
10.1029/2024ef004563

Integrating state data assimilation and innovative model parameterization reduces simulated carbon uptake in the Arctic and Boreal region

Xueli Huo[1,2], Andrew M. Fox[3,4], Hamid Dashti[5], Charles Devine[1], William Gallery[1], William K. Smith[1], Brett Raczka[6], Jeffrey L. Anderson[6], Alistair Rogers[7,8], David J. P. Moore[1]

[1]School of Natural Resources and the Environment, University of Arizona, Tucson, AZ

[2]Department of Atmospheric Sciences, University of Utah, Salt Lake City, UT

[3]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD

[4]GESTAR-II, Morgan State University, Baltimore, MD

[5]Global Change Research Laboratory, University of Wisconsin-Madison, Madison, WI

[6]Data Assimilation and Research Section, NCAR, Boulder, CO

[7]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA

[8]Environmental and Climate Sciences Department, Brookhaven National Laboratory, Upton, NY, USA

Corresponding author: David JP Moore davidjpmoore@arizona.edu

Key Points:

- Assimilating leaf area index and aboveground biomass observations into CLM reduced model bias in estimating them
- Data assimilation significantly improved CLM's performance in carbon and hydrologic cycles, as well as the functional relationships
- Implementation of a new parameterization of photosynthesis in CLM further reduced model bias in estimating the gross primary productivity

Abstract

Model representation of carbon uptake and storage is essential for accurate projection of the response of the arctic-boreal zone to a rapidly changing climate. Land model estimates of LAI and aboveground biomass that can have a marked influence on model projections of carbon uptake and storage vary substantially in the arctic and boreal zone, making it challenging to correctly evaluate model estimates of Gross Primary Productivity (GPP). To understand and correct bias of LAI and aboveground biomass in the Community Land Model (CLM), we assimilated the 8-day Moderate Resolution Imaging Spectroradiometer (MODIS) LAI observation and a machine learning product of annual aboveground biomass into CLM using an Ensemble Adjustment Kalman Filter (EAKF) in an experimental region including Alaska and Western Canada. Assimilating LAI and aboveground biomass reduced these model estimates by 58% and 72%, respectively. The change of aboveground biomass was consistent with independent estimates of canopy top height at both regional and site levels. The International Land Model Benchmarking system assesement showed that data assimilation significantly improved CLM's performance in simulating the carbon and hydrological cycles, as well as in representing the functional relationships between LAI and other variables. To further reduce the remaining bias in GPP after LAI bias correction, we re-parameterized CLM to account for low temperature suppression of photosynthesis. The LAI bias corrected model that included the new parameterization showed the best agreement with model benchmarks. Combining data assimilation with model parameterization provides a useful framework to assess photosynthetic processes in LSMs.

Plain Language Summary

The arctic-boreal zone is warming rapidly, impacting regional and global carbon cycles. The Community Land Model (CLM) can be used to project future carbon uptake and storage in this region. However, CLM is biased in estimating leave area index (LAI) and aboveground biomass that can significantly affect model projections of carbon uptake and storage. We forced the model estimates of LAI and the aboveground biomass to be consistent with satellite-derived LAI observations and a high-quality machine learning product of aboveground biomass in Alaska and Western Canada using data assimilation. The change of aboveground biomass resulted in model estimates of vegetation height consistent with independent estimates at regional and site levels. The assessment using the International Land Model Benchmarking System showed that CLM's performance in simulating carbon and hydrologic cycles was improved. Fixing the model bias in LAI only removed partial bias in carbon uptake, and a new parameterization allowing two key parameters in photosynthesis to vary with leaf temperature was introduced into CLM, to further remove the remaining bias in carbon uptake. Combining data assimilation with this new parameterization yielded more accurate model estimates of carbon uptake.

1 Introduction

The arctic-boreal zone is warming rapidly and the impact of this warming on the carbon cycle will have substantial and globally significant effects, with the region projected to become a source for carbon to the atmosphere in the coming century (Braghiere et al., 2023). Land surface models (LSMs) do not provide consistent estimates of carbon uptake in the arctic-boreal zone (Song et al., 2021; Birch et al., 2021; Fox et al., 2022; Braghiere et al., 2023). The Community Land Model (CLM 5.0; a component of the Community Earth System Model) tends to overestimate GPP in the arctic-boreal zone (Wieder et al., 2019). A recent benchmarking and model development study identified several potential problems with CLM5.0 in the arctic, including errors in where the vegetation is distributed, problems with leaf phenology and the seasonality of GPP and differential bias in GPP of different plant functional types (PFTs, Birch et al., 2021). Moreover, compared to satellite benchmarks, the peak season of leaf area in the arctic-boreal zone was delayed on average by one to two months across 27 LSMs participating in the 6th Coupled Model Intercomparison Project (CMIP6) (Song et al., 2021)

Allocation schemes in CLM 5.0 are empirical and relatively simple (Oleson et al., 2013). The model allocates carbon between leaf, stem (live and dead stem), coarse root (live and dead coarse root), and fine root based on four allometric parameters: 1) ratio of new fine root to new leaf carbon allocation, 2) ratio of new coarse root to new stem carbon allocation, 3) ratio of new stem to new leaf carbon allocation, and 4) ratio of new live wood to new total wood allocation. It is challenging to observe allocation to different pools at large scales, so we infer allocation from studies of biomass. Data to parameterize dynamic allocation schemes are rare and typically include only estimates of the average biomass within the leaf, wood and root pool (Caspersen et al., 2000; Gower et al., 2001; Brown, 2002; Houghton, 2005; Litton et al., 2007; Luyssaert et al., 2007; Keith et al., 2009; Franklin et al., 2012; Oleson et al., 2013; Montané et al., 2017). Decadal and centennial carbon storage depends on how the product of photosynthesis is allocated. Different plant pools (leaf, stem, and root) have different functions and residence times (Delbart et al., 2010) and modeling studies that investigate the influence of allocation on biomass accumulation show that this poorly constrained process exerts huge control over long term carbon storage (Friend et al., 2014; Montané et al., 2017).

In contrast, LSMs, represent short term biophysical processes using well-tested and more mechanistic equations and represent long term ecological or biogeographic processes using less tested and more empirical equations (Bonan, 2019). Our study focuses on the Community Land Model (CLM5.0). In CLM5.0, photosynthesis is represented by the mostly mechanistic Farquhar et al. (1980) model where the response to irradiance is represented by an empirical, non-rectangular hyperbola where key parameterization is associated with the initial slope (quantum yield) and curvature of that relationship. In CLM5.0 the quantum yield approaches the theoretical maximum which has been commonly observed in unstressed dark-adapted plants (Long et al., 1993; Singsaas et al., 2001; Kromdijk et al., 2016) but which is rarely observed in nature, particularly in plants experiencing stress such as drought or low temperature (Rogers et al. 2019, Bolharnordenkampf et al., 1991; Groom & Baker, 1992; Ogren & Evans, 1992; Long et al., 1994). In contrast, quantum yield and convexity measured in arctic plants were reduced significantly at low leaf temperatures (Rogers et al. 2019). This suggests the potential to overestimate GPP in the arctic-boreal zone.

Modelling carbon uptake and storage remains a challenge, especially for the arctic-boreal zone. A CMIP6 analysis showed that tree height was, on average, overestimated in the arctic-boreal zone (Song et al., 2021); an error consistent with poor parameterization of carbon allocation. Biases in GPP, as reported by Bonan et al. (2011), arise from model parameter uncertainties and from model structural/parameterization errors entailing radiative transfer, leaf photosynthesis and stomatal conductance, and canopy scaling of leaf processes. While the sophistication of physiological processes in LSMs has increased steadily over the last few decades (Blyth et al., 2021), there is evidence that GPP is not always represented using accepted photosynthetic parameterizations (Rogers et al., 2017, 2019). If models are parameterized to match benchmarks of GPP without first ensuring that biomass and LAI are correctly modeled, there is a risk of introducing compensating errors in allocation and photosynthetic processes.

Data assimilation can be applied to improve performance of LSMs and circumvent the lack of understanding in processes controlling GPP and allocation. Data assimilation of leaf area index (LAI) is an increasingly common method to reduce errors in allocation of leaf carbon in LSMs; e.g. CLM (Stöckli et al., 2008; Fox et al., 2018; Ling et al., 2019; Raczka et al., 2021), ISBA (Albergel et al., 2010, 2017), ORCHIDEE (Demarty et al., 2007; Bacour et al., 2015; MacBean et al., 2015), CHTESSEL (Boussetta et al., 2015) and Noah-MP (Kumar et al., 2019). For example, an ensemble Kalman filter was used to update the prognostic estimate of LAI from CLM5.0 to more faithfully match the LAI3g (Zhu et al., 2013) satellite data product (Fox et al., 2022). Model estimates of GPP are dependent on LAI magnitude and duration and processes controlling photosynthetic rates, but LSM errors in prognostic LAI are significant (Montané et al., 2017). In Fox et al. (2022), assimilating LAI into CLM5.0 resulted in a globally averaged decline in modelled GPP of 18%, and in the arctic-boreal zone, the decrease was up to 50%.

In this study we use an ensemble data assimilation approach to constrain leaf area and biomass, aiming at reducing biases in GPP and allocation processes in CLM in a subset of the arctic-boreal zone, the Arctic-Boreal Vulnerability Experiment (ABoVE) region which includes Alaska and Western Canada. We verify the change of biomass by comparing modeled and measured tree height at the regional and site levels. To further reduce bias in GPP when the error in phenology is fixed through data assimilation, we implement a new parameterization allowing the variation of maximum quantum yield and curvature of the response of photosynthesis to irradiance with leaf temperature which was developed based on the findings in Rogers et al. (2019). Then, we compare CLM runs with and without data assimilation and with and without the modified photosynthetic process to see the effect of data assimilation and the new parameterization on reducing biases in GPP.

2 Materials and Methodology

We constrained LAI and aboveground biomass state variables estimated by CLM5.0 to satellite and derived data estimates (hereafter observations) by implementing the Ensemble Adjustment Kalman Filter (EAKF) (Anderson, 2001). LAI data assimilation aims to improve GPP by correcting bias in LAI, and biomass data assimilation focuses on correcting wood (stem and root) carbon pools and decomposition (litter and soil) carbon pools with the aim to improve respiration fluxes and vegetation structure such as tree height. We then compared the model output to a suite of independent datasets to verify the adjustment of states was successful. Removal of bias in model states allowed us to develop and implement a new parameterization of photosynthesis in CLM5.0 based on *in situ* data collections (Rogers et al., 2019) to improve model fluxes. We evaluated all the model runs against established land surface benchmarks.

2.1 CLM-DART

The Community Land Model (CLM) is capable of simulating complex biophysical and biogeochemical processes on land (Lawrence et al., 2019). It was run in the biogeochemistry and crop (BGC-Crop) mode in which the carbon and nitrogen in the natural vegetation, litter and soil are prognostic at each time step, and the prognostic crop model is turned on. The land cover and land use are constant in the model run.

The Data Assimilation Research Testbed (DART) is open-source community software for ensemble data assimilation (Anderson et al., 2009a). We used CLM-DART, a coupled system of DART and CLM to carry out the model experiments described in this study. We configured CLM-DART similarly to previous studies (Fox et al., 2018, 2022; Raczka et al., 2021) using the Ensemble Kalman Adjustment Kalman Filter (EAKF), a fully deterministic and computationally efficient algorithm (Anderson, 2001). The CLM-DART settings used in this study are provided in Table 1.

The assimilation time step is set to every 8 days to match the frequency that the leaf area index observations are available. The annual biomass observations do not have an assigned observation date; thus, we prescribed the biomass observation during the month of September to align with the leaf area index observations (e.g. Sep 5th, 2012, Sep 6th, 2011, 2013, 2014). We implemented minimal additional quality control given the highly processed nature of the observations but did use the outlier rejection to reject observations that have accurate values but are so far away from the model ensemble mean. If the difference between the observation and the prior ensemble mean is more than N standard deviations from the square root of the sum of the prior ensemble and observation error variance, the observation will be rejected. The number of standard deviations is called the outlier threshold, and the value of the outlier threshold can be found in Table 1. Note that outlier rejection was applied only to LAI observations and turned off for biomass observations due to the scarcity of biomass observations. Otherwise, almost all of the biomass observations would be rejected, resulting in biomass DA having no impact. After assimilation, the ensemble spread decreases consistently. It is crucial to increase and maintain it to prevent insufficient forecast error variance (i.e., ensemble spread), which can lead to excessive rejection of observations. Inflation can achieve this by increasing the ensemble spread without changing the ensemble mean. We used the time- and space-adaptive state-space inflation (El Gharamti et al., 2019) that is spatially distributed and evolve with time as observation changes. The damping parameter is used to reduce the inflation when the frequency or density of observations declines. For more details of damping and inflation, see DART tutorial (https://docs.dart.ucar.edu/en/latest/guide/inflation.html).

We found the inflation generated with the default parameter settings within this approach overly inflated the ensemble spread at grid cells where the observation density was low, leading to unrealistic spatial heterogeneity. We kept the default settings of the inflation standard deviation and its lower bound (both set to be 0.6) which control how quickly the inflation responds to new observations since these settings have been demonstrated to yield good results for large geophysical models (El Gharamti et al., 2019). Based on this, we tuned the damping parameter to reduce the inflation. The inflation applied to the prior state is 1+damping × (current inflation - 1.0), i.e., the sum of 1.0 and the difference between the current inflation value and 1.0 multiplied by the damping value (Anderson, 2007; Anderson, 2009b). Two different damping values, 0.9 and 0.4, were used to account for the varying seasonal spatial coverage of the leaf area index observation. A large damping value, 0.9, was used to damp inflation slowly to increase prior ensemble spread when the availability of data is greater and a small damping value, 0.4, was used to damp the inflation quickly to minimize prior ensemble spread where data availability is lower.

To limit the influence of the observations to specific regions of the DART state and reduce the likelihood of applying spurious updates during the assimilation update step, we used localization. First, we impose a horizontal spatial localization function (Gaspari & Cohn, 1999) with a halfwidth value of 0.015 radians to limit the influence of an observation for prognostic variables to near the observation location. Second, we limit the influence of the observations to specific CLM prognostic variables. When biomass observations are assimilated, six vegetation carbon pools (leaf, live stem, dead stem, fine root, live coarse root and dead coarse root) and decomposition pools (coarse woody debris, litter and soil) are updated, however, when leaf area index observations are assimilated, only the leaf carbon pool is updated. When both biomass and leaf area index observations are available, only biomass observations are assimilated.

Table 1 Summary of CLM-DART settings for the data assimilation (DA) run

Experiment	Observation	Simulation period	Damping value	Outlier threshold	Prognostic variables to update in restart files
Assimilation	leaf area index and aboveground biomass	2011-2019	0.9 in summer (from June 9th to September 5th in 2012 and 2016, and from June 10th to September 6th in other years), 0.4 in other seasons	3 when leaf area index observation is assimilated; -1 when biomass observation is assimilated (outlier rejection turned off)	Six displayed vegetation carbon pools and all decomposition carbon pools when biomass observation is assimilated; Leaf carbon pool only when leaf area index observation is assimilated. Note when both leaf area index and biomass observations are available, only biomass observations are assimilated.

Briefly, the way how CLM-DART works is: CLM provides a forecast simulation until the time when an observation is available (every 8 days in this case). At this time inflation is applied to increase the ensemble spread. An observation operator is then applied to the CLM output that calculates the model estimate of the observation, which we call the observed variable (Text S1 in Supporting Information S1). The observed variable is then adjusted with increments which are calculated using the information of the observation likelihood and the prior distribution (Text S2 in Supporting Information S1). Increments to unobserved variables are calculated based on the covariance between the observed variable and unobserved variables. Increments are applied to the prognostic variables of CLM stored in the restart file. The updated restart file serves as the initial condition for the next forecast. All these steps are repeated in the subsequent assimilation cycles.

2.2 Observations used in data assimilation

2.2.1 Leaf Area Index (LAI)

The MCD15A2H version 6 Moderate Resolution Imaging Spectroradiometer (MODIS) LAI product is an 8-day, 500-meter, global satellite data product available from 2011 to 2019 and was obtained using NASA Application for Extracting and Exploring Analysis Ready Samples (AρρEEARS; https://appeears.earthdatacloud.nasa.gov/). AρρEEARS categorized MODIS LAI flags into several categories, and only MODIS LAI pixels within the very good or better category are used in the assimilation. The 500 m LAI and the associated uncertainty denoted by its standard deviation were re-gridded to the model resolution (~25km) using spatial averaging. MODIS LAI covers most of the ABoVE region (64.6%~77.6%) from mid-June to mid-

September and decreases with time from mid-September to mid-November as snow covers LAI first in higher and then lower altitudes. From late November to early January, no LAI observations are available in the region, and from mid-January to early June, the coverage of LAI expands from south to north with time.

2.2.2 Aboveground biomass

The annual, regional 30-m, aboveground biomass is a machine learning product specifically developed for the boreal forest biome portion in the ABoVE domain (Wang et al., 2021). It upscales spaceborne lidar-based estimates of aboveground biomass with satellite surface reflectance, climate and topographic data based on a machine learning model, and it overlaps the model simulation period from 2011 to 2014. The original data in standard "B" grid tiles (106 aboveground biomass and 106 standard error) were re-projected from Albers equal area conic to WGS84 projection and aggregated to the approximate model grid resolution (~25km) from 30m using spatial averaging. Finally, tiles were mosaiced annually and adjusted to the precise model grid using nearest neighbor pixel matching. Aboveground biomass in CLM was calculated as the sum of leaf carbon, live and dead stem carbon. Biomass in Wang et al. (2021) was assumed to be 50% carbon. All of the biomass observations are considered to be of satisfactory quality and are assimilated into CLM.

2.3 A new parameterization of GPP in arctic plants

To investigate whether cold temperature inhibition of photosynthetic capacity could account for overestimates of GPP, we implemented a new parameterization of the photosynthesis module in CLM based on *in situ* observations collected at a site (71.28°N, 156.65°W) near Barrow (now Utqiagvik) in Alaska (Rogers et al., 2019). We mapped the field estimated maximum quantum yield and convexity (Rogers et al., 2019) to the CLM parameters for maximum quantum yield and curvature, respectively (Text S3 in Supporting Information S1). We updated CLM maximum quantum yield and curvature values to match the *in-situ* measurements (Table 2) and applied linear and nonlinear regression to estimate temperature responses for the parameters (Text S4 and Figure S1 in Supporting Information S1). Values were not extrapolated above 25 °C or below 5 °C.

Table 2 Default and updated values of maximum quantum yield and curvature at three different leaf temperatures

Leaf temperature (°C)	Maximum quantum yield denoted by $0.5\boldsymbol{\Phi}_{PSII}$ (mol CO_2 mol^{-1}absorbed quanta)		Curvature denoted by $\boldsymbol{\Theta}_{PSII}$ (unitless)	
	Default	Updated	Default	Updated
5	0.425	0.132	0.7	0.44
15		0.217		0.5
25		0.316		0.65

2.4 Model Simulations

We carried out two 40-member ensemble CLM runs: free run (no assimilation) and data assimilation (DA) run. In the DA run, both LAI and aboveground biomass observations assimilated serially. LAI observations are assimilated every 8 days and biomass observations are assimilated once a year on the specific date we assigned due to the fact that the frequency of LAI observations is 8 days and biomass observations are annual. These runs were used as the starting point for single-member CLM model runs that included (or did not include) the reparametrized photosynthesis module.

All simulations were run at a spatial resolution of 0.25 × 0.25 degrees (~25 × 25 kms) in the ABoVE region. The surface and domain data of this resolution were generated using the CLM mkmap tool from the default input datasets. In both free and DA runs, CLM was driven by 40 ensemble members of the CAM6 reanalysis forcing data (Raeder et al., 2021; Ds345.0, 2020) which is an atmospheric ensemble generated by assimilating atmosphere observations into version 6 of the Community Atmosphere Model (CAM6) from 2011 to 2019 with a spatial resolution of 0.9 × 1.25 degrees. Atmospheric data were interpolated onto the 0.25 × 0.25 degrees land grid automatically by the default bilinear interpolation within CLM.

Initial conditions to perform the free and DA runs from 2011 to 2019 were estimated using a single-member model spin-up, initialized from CLM default present-day condition using atmospheric data from the first member of the CAM6 ensemble from 2011 to 2019 cycled 120 times (1080 years total) to equilibrium. Initializing each ensemble member in this way was too computationally costly, so the initial ensemble spread was created by running CLM with 40 CAM6 forcing ensemble members (2011-2019) four times (36 years total) from the initial condition generated by the single-member spin-up.

To evaluate the impacts of either the new parameterization or data assimilation or both of these approaches on reducing biases in GPP, three additional model experiments are performed: parameterization, initialization, and parameterization + initialization runs. Running all three model experiments with 40 ensemble members for the entire simulation period (2011-2019) would be computationally expensive and unaffordable. Due to limited computational resources, in all three model runs, CLM was driven by the first member of the CAM6 ensemble and ran for one year (2015). These additional model runs are like the free run in that no observations were assimilated, so they are model forecasts. First, the *parameterization simulation* includes the new parameterization but is initialized with the free run model state on 1 January 2015, Second, the *initialization simulation* doesn't have the new parameterization but is initialized from the updated DA run model state on 1 January 2015. Third, the *parameterization + initialization simulation* includes both the new parameterization and is initialized from the DA run model state on 1 January 2015.

2.5 Model evaluation data sets

2.5.1 Canopy top height

Improved aboveground biomass in CLM should result in a more realistic estimate of canopy height assuming the tree allometry is broadly correct. We evaluated canopy height using a global canopy top height dataset and local airborne light detection and ranging (lidar).

We extracted canopy height data for the ABoVE region from the Geoscience Laser Altimeter System (GLAS) aboard ICESat (Ice, Cloud, and land Elevation Satellite) 1km × 1km global dataset (Simard et al., 2011). These data were re-gridded to the model resolution (~25km × ~25km) for regional comparisons across grid cells that were dominated by NEBT (needleleaf evergreen boreal tree, as shown in Figure 3a). To evaluate height of other PFTs, we used the National Ecological Observatory Network (NEON) airborne observation platform (AOP) estimates of canopy top height (NEON, 2023) derived from the airborne lidar data. These estimates have a 1m × 1m spatial resolution and are distributed in 1km × 1km tiles. The NEON AOP canopy top height estimates agree well with ground measurements at two NEON sites in Alaska: Healy and Delta Junction (Figure S2 in Supporting Information S1) and were used as the benchmark for comparison at the two sites. For each site, we collected all tiles of NEON AOP data within the model gridcell, covering 25% and 40% of area for Healy and Delta Junction, respectively. We compare the model height to the distribution of canopy top height from NEON AOP considering the abundance of each PFT.

2.5.2 ILAMB (International Land Model Benchmarking)

The International Land Model Benchmarking system (ILAMB, Collier et al., 2018) is an open-source land model evaluation tool that compares model simulations to benchmark datasets including global-, regional-, and site-level data and calculates scores to represent model performance. ILAMBv2.6 was used to assess whether assimilating LAI and aboveground biomass observations into CLM improves model performance for the terrestrial carbon and water cycles in the ABoVE region. The assessment integrated analysis for 12 variables in the carbon and water cycles utilizing 22 benchmark datasets which were downloaded from the ILAMB data archive (https://www.ilamb.org/ILAMB-Data/DATA/). Note that data from the global benchmark datasets in regions other than the ABoVE region were masked out during the evaluation. For each variable, ILAMB produces maps, time series, statistics, assessment of variable-to-variable relationship, scores for bias, RMSE, seasonal cycle, interannual variability, spatial distribution and an overall score ($S_{overall}$) representing the overall performance of the model (Collier et al., 2018).

Note the default CLM5.0 simulations driven by the Global Soil Wetness Project (GSWP3v1) forcing which scores the best compared to other forcing data sets (Lawrence et al., 2019) is also included in the ILAMB assessment to evaluate the impact of the alternative CAM reanalysis forcing as well as data assimilation on the performance of CLM.

2.5.3 FLUXCOM gross primary productivity

The FLUXCOM GPP product used as the benchmark in comparing the seasonal cycle of GPP from different model runs is identical to the GPP benchmark in ILAMB, and it was downloaded from the ILAMB data archive. This product overlaps the simulation period from 2011 to 2013 and is one of the 0.5 × 0.5 degrees, monthly, global gridded FLUXCOM GPP ensemble products. It was generated using an artificial neural networks machine learning approach with CRUNCEPv6 meteorological data and mean seasonal cycles of several MODIS based variables (Tramontana et al., 2016; Jung et al., 2019). The seasonal values of the GPP product were calculated within ILAMB and stored in its output files. Only data in the ABoVE region and during the overlapped time period were used in the assessment.

3 Results

3.1 State data assimilation reduced LAI and aboveground biomass to match remote sensing data products

Assimilating leaf area index and aboveground biomass observations into CLM5.0 significantly improves the model's estimates of LAI and aboveground biomass both temporally and spatially in the ABoVE region. The free run in which CLM was run without assimilation significantly overestimates LAI and aboveground biomass (Table 3, Figure 1). The DA run corrects a significant amount of the LAI bias in the free run, with modeled LAI reduced by 58% and phenology aligning with the observations (Figure 1a). The aboveground biomass is reduced by 72% through data assimilation. The DA run represents the spatial variability of LAI and aboveground biomass more closely to the observations compared to the free run (Figure 1c, 1d).

The impact of DA varies across the ABoVE domain in proportion to the bias in the free run (Figure 2a, 2b). LAI is overestimated in the free run across 83% of the area. Although the average bias is 1.27 m^2/m^2, in large portions of the domain the bias is as high as 4 to over 6 m^2/m^2 (Figure 2a). In the DA run, LAI bias relative to the satellite estimate is reduced to 0.019 m^2/m^2. The model over- and under- estimates the satellite data but the extent of extreme errors is significantly reduced (Figure 2b, e). Similarly, aboveground biomass is overestimated in the free run across 95% of the area; the average model bias is 4222 gC/m^2 but in some southern portions of the domain the bias is much higher than 15,000 gC/m^2 (Figure 2c, f). In the DA run, aboveground biomass bias is reduced significantly, resulting in relatively small positive and negative differences with the aboveground biomass data product (Figure 2d, f). The magnitude of initial bias and consequently the size of the adjustment required was surprisingly high, indicating a significant misrepresentation of either cumulative carbon uptake, allocation or turnover.

Table 3 Statistics of LAI (from 2012 to 2019) and aboveground biomass (in 2014). Mean $\pm$ standard deviation (RMSE, bias).

	Obs	Free	Assim	Mean Change $\frac{Assim-Free}{Free}$ (%)	Reduction in error $\frac{Assim\ RMSD-Free\ RMSD}{Free\ RMSD}$ (%) $\frac{Assim\ bias-Free\ bias}{Free\ bias}$ (%)

LAI (m^2/m^2)	0.87	2.14+0.006 (1.27, 1.27)	0.89+0.001 (0.019, 0.019)	-58.4	-98.5 -98.5
Abovegro und Biomass($gC//m^2$)	1692	5914+16 (4222.2, 4222.1)	1670+1 (22.5, -22.5)	-71.8	-99.5 -100.5

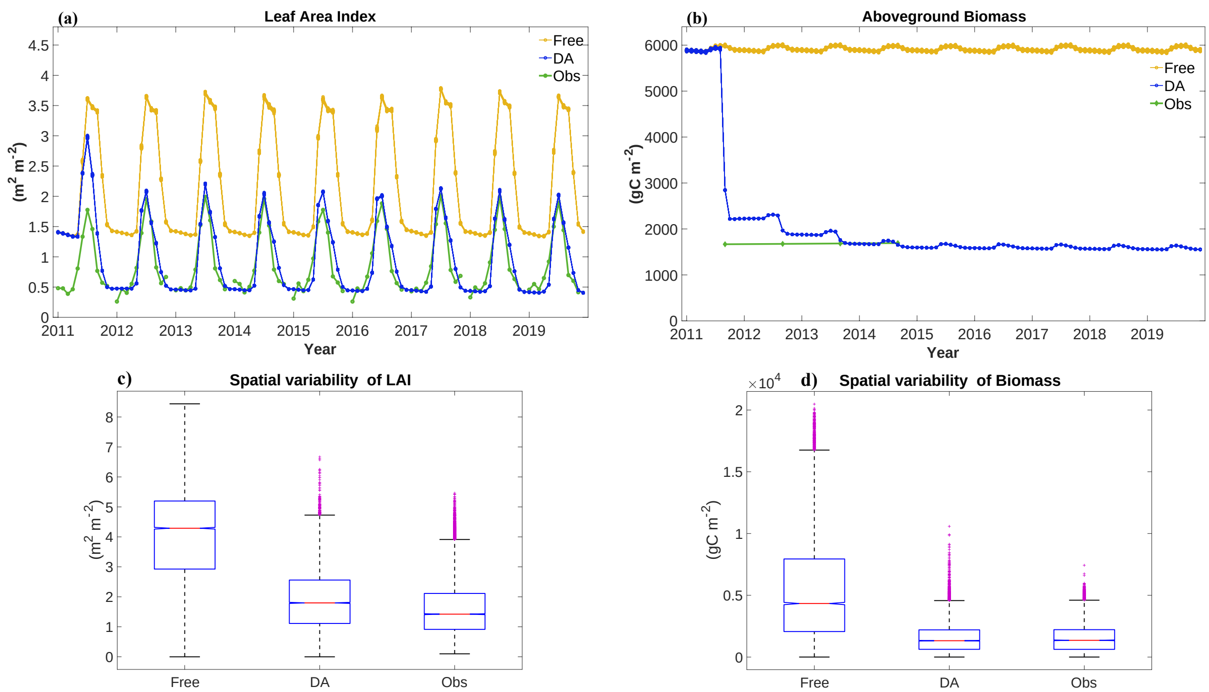
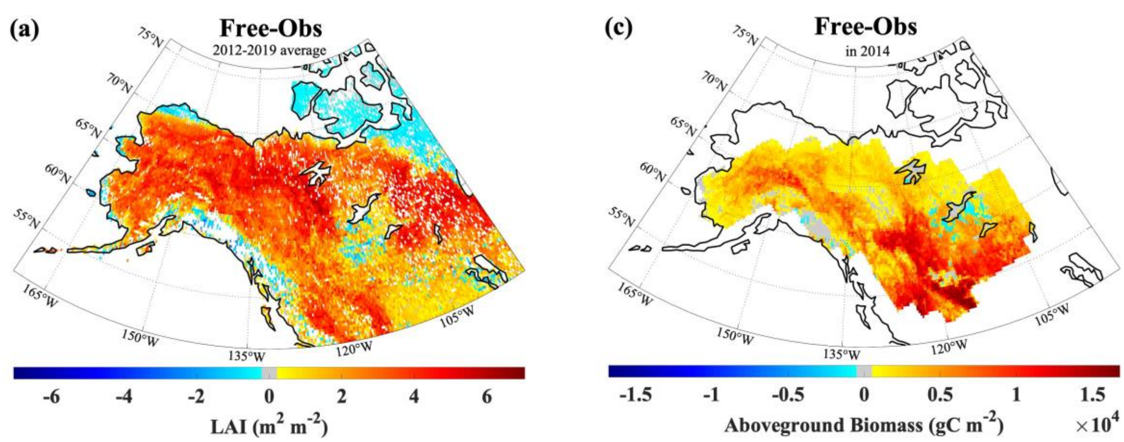


Figure 1. Time series of (a) monthly LAI and (b) aboveground biomass from the free run (orange line), data assimilation (DA) run (blue line) and the observation (green line). LAI is averaged over the ABoVE region and aboveground biomass is averaged over the ABoVE Boreal Forest domain to be consistent with the spatial coverage of aboveground biomass observations. The 8-day MODIS LAI observation is averaged to the monthly time scale, and the aboveground biomass observation is annual. The boxplot shows the spatial variability of (c) LAI averaged from 2012 to 2019 and (d) aboveground biomass in 2014 when LAI and biomass in the DA run are stable.



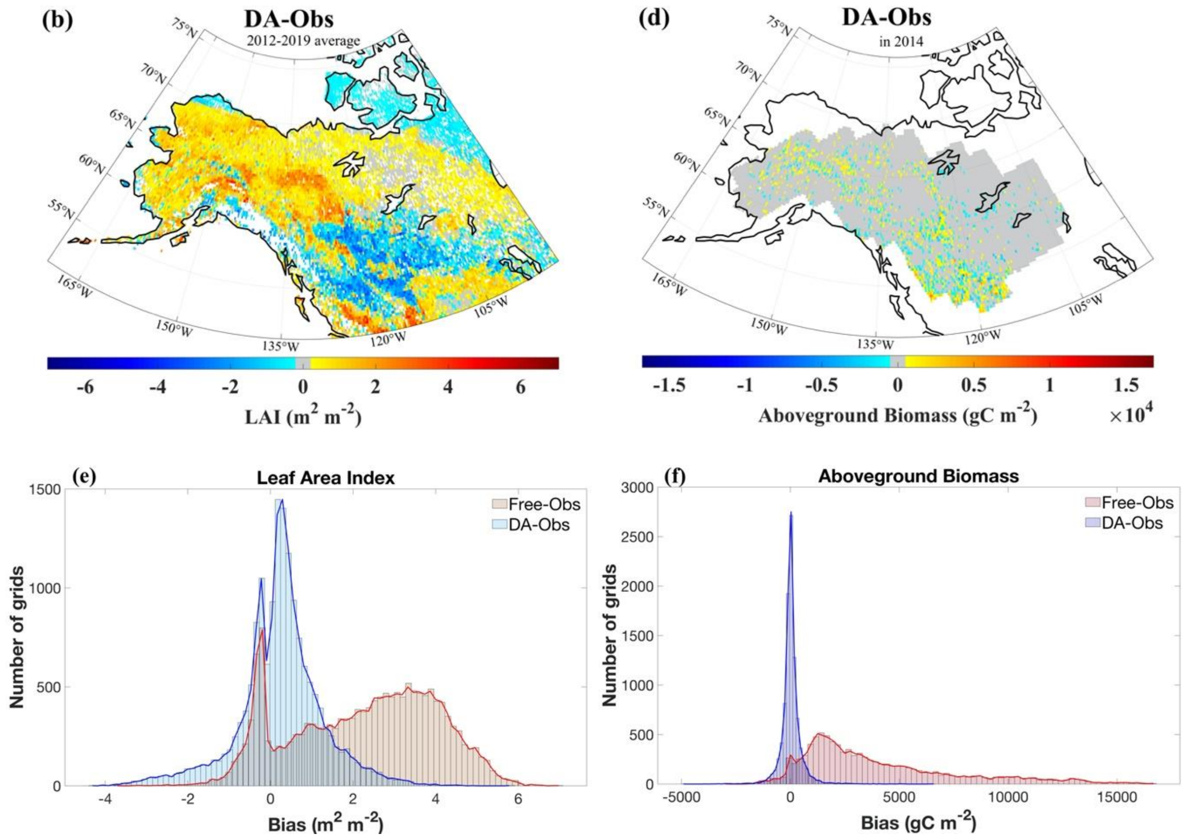


Figure 2. Spatial maps of the difference between modeled LAI (ensemble mean) and MODIS LAI observations averaged over July and August from 2012 to 2019: (a) free run; (b) data assimilation (DA) run. Spatial maps of the difference between modeled aboveground biomass (ensemble mean) and aboveground biomass observation in 2014 using annual values: (c) free run; (d) DA run. Histograms of spatial bias of LAI (e) and of aboveground biomass (f). To avoid bias, comparisons are restricted to the period when observations cover the most of ABoVE region and DA run achieves stability. LAI comparisons are averaged over July and August because MODIS observes most of the region in that time. We also limit LAI comparisons to the time between 2012 to 2019 because LAI in the DA run is stable during that time. Differences in aboveground biomass are restricted to 2014 because that is the last year when observations are available and the modeled biomass in the DA run is stable from then on.

3.2 Independent estimates of vegetation height support the aboveground biomass corrected by the DA system.

Comparing the model's estimates of canopy top height with independent canopy top height estimates provides support for the changes in aboveground biomass. Canopy height correlates well with aboveground biomass (Lefsky et al., 2002; Drake et al., 2002; Lefsky et al., 2005; Takagi et al., 2015). CLM calculates canopy top height from dead stem carbon, the major component of aboveground biomass, using a linear equation. Assuming this relationship is reasonable, independent height data is a proxy of aboveground biomass and can be used to validate the changes in the aboveground carbon stock altered by the DA system. For areas where needleleaf evergreen boreal trees (NEBT) was greater than 95% (dark blue shade in Figure 3a), we found that DA significantly improved model estimates of height compared with satellite lidar-derived canopy height (Figure 3b). NEBT was chosen because it is widespread, often dominates large areas in the ABoVE region, and it is typically above the minimum detection limit of ICESat (5m). The distribution of the height of NEBT (Figure 3b) shows that canopy height is overestimated in the free run compared with the ICESat data, and the canopy height from the DA run is closer to the validation data. One caveat is that some of the mismatch between the distribution of canopy height in the model runs and the validation data might be caused by the mismatch between the spatial distribution of NEBT in CLM and the true spatial distribution of NEBT.

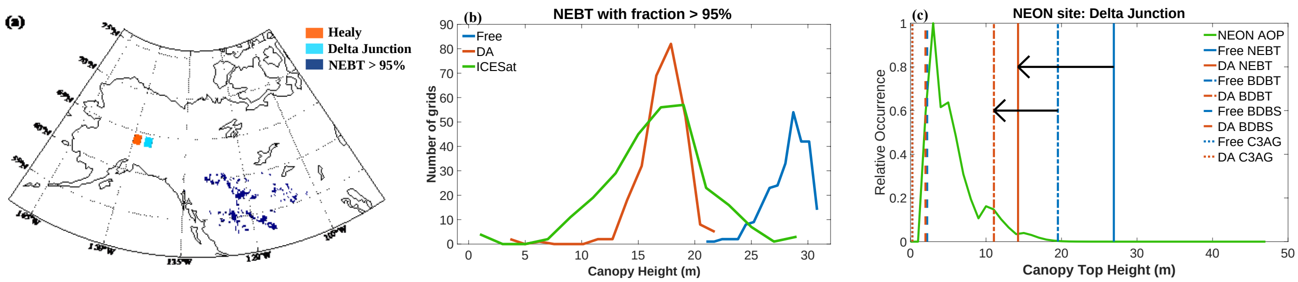


Figure 3. Comparison of canopy top height estimated by CLM5.0 with independent canopy height estimates from ICESat lidar observations at regional and site levels. (a) Locations of the grid cells dominated by NEBT (needleleaf evergreen boreal tree, in dark blue) and two NEON sites (Healy in red, Delta Junction in cyan). (b) The distribution of canopy top height of the widespread NEBT estimated by CLM5.0 free run (in blue), DA run (in red) and derived from ICESat lidar measurements (in green). (c) The distribution of canopy top height at Delta Junction from CLM free run (vertical blue lines), DA run (vertical red lines) and NEON airborne observation platform (AOP, in green). PFTs in CLM within the gridcell where Delta Junction is located are NEBT (needleleaf evergreen boreal tree), BDBT (broadleaf deciduous boreal tree), BDBS (broadleaf deciduous boreal shrub), C3AG (C3 Arctic grass).

Assimilating LAI and biomass observations leads to improved model estimates of canopy height for other PFTs as well. We compared model estimates of canopy height to those from the National Ecological Observatory Network (NEON) airborne observation platform (AOP) at two NEON sites: Healy and Delta Junction. Figure 3c displays the distribution of canopy top height from the NEON AOP, free run and DA run at Delta Junction. Similar results are found at Healy as well. The arctic grass (C3AG) has no change, and the shrub (BDBS) is slightly shorter in the DA run compared to the free run. Notably, both boreal trees, needleleaf evergreen (NEBT) and broadleaf deciduous (BDBT), are much shorter in the DA run. The heights of boreal trees in the free run are near or over 20 meters, whereas in the DA run, the maximum height of the boreal trees is around 15 meters. NEON AOP data suggests that the possibility of a tree taller than 20 meters is extremely low, supporting the realism of canopy top height estimates from the DA run.

3.3 DA results match with most large-scale land model benchmarks better

The DA run showed substantial improvement over both the free run and the default run of CLM when compared to a wide range of independent land model benchmarks (Figure 4). Compared to ILAMB carbon and hydrological benchmarks, the DA run outperformed both the free run and the default CLM5.0 run with GSWP3v1 forcing (Lawrence et al., 2019). Nine of the twelve benchmarks showed improvement with respect to the default model: LAI, aboveground biomass, total biomass, GPP, ecosystem respiration, evapotranspiration, latent heat, sensible heat, and terrestrial water storage. All four functional relationships between LAI and other variables (aboveground biomass, total biomass, GPP, and evapotranspiration) were improved. The snow water equivalent in the data assimilation run was worse compared to the default CLM5.0 run with GSWP3v1 forcing but better than the free run forced with CAM6 forcing, probably due to the degradation in the snowfall or snowmelt rate in the CAM6 forcing which needs further investigation, rather than errors introduced by data assimilation. DA alters other vegetation carbon pools in addition to the aboveground carbon pools (Figure S5 in Supporting Information S1). Consequently, the ratio of each vegetation carbon pool to the total vegetation carbon content changes (Figure S6 in Supporting Information S1). However, the lack of such benchmark data hinders us from verifying the plausibility of the change.

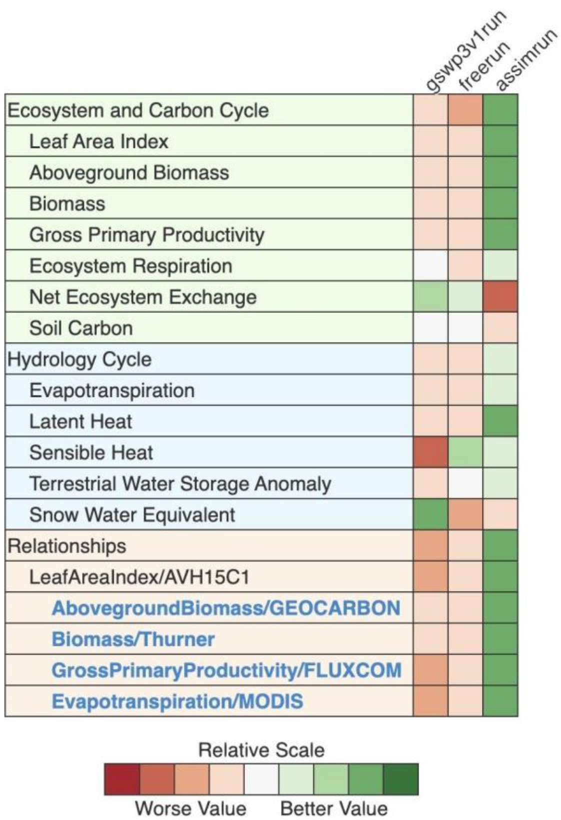


Figure 4. ILAMB summary diagram for the default CLM5 run driven by GSWP3v1 forcing (gswp3v1run), free run and data assimilation run. The color represents the overall score described in the Methodology.

To estimate the overall carbon balance, land surface models calculate net ecosystem exchange as the small remainder between large photosynthetic and respiration fluxes. The ILAMB benchmarking suggests our DA approach improves estimates of GPP and ecosystem respiration but has poorer performance with respect to net ecosystem exchange (Figure 4). While aboveground respiration decreased as the DA reduced aboveground biomass, below ground respiration did not respond similarly. Belowground respiration in the DA run increases in August, September, and October (Figure S4 in Supporting Information S1) because of a slight increase in soil carbon pool updated by data assimilation.

3.4 GPP bias in CLM stems from estimating model LAI states and photosynthetic rates

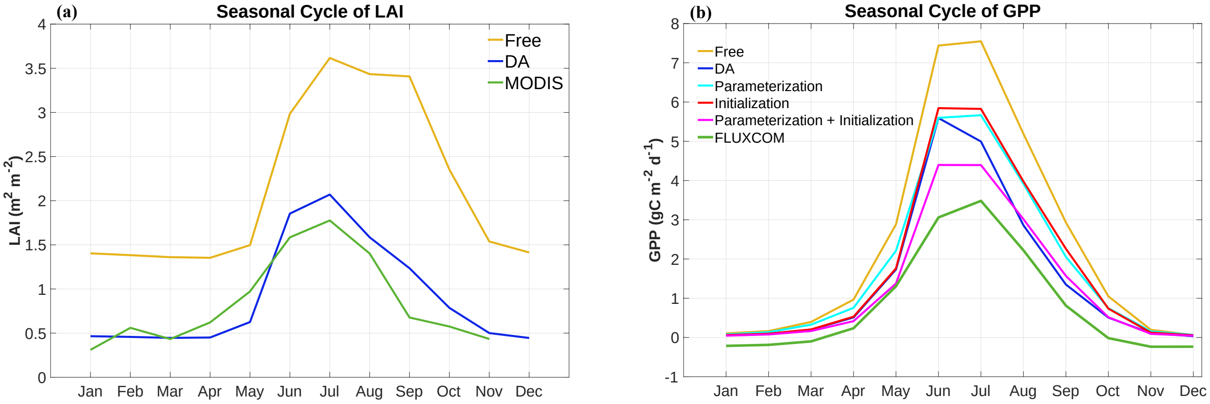


Figure 5. (a) Seasonal cycle of LAI from CLM5.0 free run (light orange), DA run (blue), and MODIS (green) using data in 2015. (b) Seasonal cycle of GPP from different CLM runs in 2015 and FLUXCOM averaged from 2011 to 2013 (green). Parameterization (cyan) implements the new parameterization and starts from the initialization condition identical to the free run. Initialization (red) starts from the initialization condition identical to the DA run. Parameterization + Initialization (pink) implements the new parameterization and starts from the initialization condition identical to the DA run. Note that the free run and DA run shown here are results of the first ensemble member for a fair comparison to the other three CLM runs all of which are driven by the first member of the CAM6 ensemble.

The error in modeled LAI in CLM was corrected through DA as evidenced by the significant reduction in LAI (Figure 5a). However, the bias in carbon uptake was not fully removed. GPP in the DA run (Figure 5b) was still highly biased compared to the FLUXCOM GPP estimate (Jung et al., 2019; Tramontana et al., 2016). This indicates that other model parameterizations associated with GPP could be incorrect. Building on the work of Rogers et al. (2019), we developed and implemented a new parameterization that considered the effect of temperature on the light response curve and implemented it in CLM, with the aim to further reduce bias in the model estimate of GPP in the ABoVE region.

The implementation of the new parameterization (cyan) in CLM reduces model error in GPP compared to the free run (orange) which does not include the parameterization, and it has a similar effect in forecasting GPP as the provision of a better initialization (red) achieved by data assimilation (Figure 5b). This indicates that fixing the bias in the photosynthetic parameters in light response of photosynthesis which is a fast process has a similar effect to improving GPP by fixing the bias in the phenology which is a relatively slow process. When we improved both the initial conditions (through DA) and photosynthesis related parameterization, the forecasted GPP (pink) was the closest to the FLUXCOM benchmark dataset with the greatest model error reduction. It would be interesting to evaluate the relative effect of DA and new parameterization on deciduous versus evergreen forest. However, the data for deciduous forests are too limited (only two grid cells are dominated by deciduous trees and affected by DA) to provide a credible comparison.

4 Discussion

Failure to accurately model LAI and aboveground carbon pools leads to significant errors in projections of regional GPP. Removing biases of 58% and 72% in LAI and aboveground biomass through DA (Table 3) resulted in a 40.6% reduction in GPP (gC/m^2/year) (Table S2). Overall, DA of LAI and biomass significantly improved CLM simulations of the carbon and hydrologic cycles, as well as in representing the functional relationships between LAI and other variables (aboveground biomass, total biomass, GPP, and evapotranspiration; Figure 4).

DA likely resulted in realistic allocation of carbon to above ground biomass as it improved vegetation height estimates when compared to independent airborne and spaceborne lidar (Figure 3).

Data assimilation is an effective tool to adjust model states to initial conditions that match observations. Initialization of LSM carbon pools is a challenging but essential step in projecting the future state of the land carbon sink. The practice of model spin-up to equilibrium is time and resource consuming and the assumption of equilibrium is not realistic for most ecological systems (Luo et al., 2015). Running models for millennia, analytical solvers, model vectorization and state data assimilation can be used to more effectively initialize LSMs (Hoffman et al., 2005, Jeong et al., 2008; Luo et al., 2011; Ajami et al., 2014; Liao et al., 2023).

Removing biases caused by processes that influence LAI is a necessary first step to improving estimates of GPP and forecasting carbon storage in LSMs. A detailed modelling study in the artic boreal zone identifies CLM specific issues with LAI phenology, mistimed peak GPP and high GPP for some plant functional types (Birch et al., 2022). Our DA approach adjusted both the timing and magnitude of LAI (Figure 5a) and resulted in significant improvement in GPP relative to FLUXCOM estimates (Figure 5b). Failure to predict LAI is a persistent problem across many ESMs and there has been limited improvement in model projections of LAI from CMIP5 to CMIP6 (Mahowald et al., 2016; Song et al., 2021). A comparison of high latitude LAI in seven Earth System models with the LAI3gv.1 product (Zhu et al., 2013) either overestimated or underestimated LAI (Winkler et al., 2019). Nearly all (24 of 27) CMIP6 models overestimate satellite estimates of global mean LAI (aggregated from three satellite products) and 9 of these models show bias of more than 50% (Song et al., 2021). A study of biophysical processes mediated by leaves in four land surface models attributed biases in interannual variability of LAI to parameterization of the carbon allocation and phenology schemes in these models (Forzieri et al., 2018). Alteration of the phenology model in the CARDAMOM model changes the sensitivity of carbon storage to climate (Norton et al., 2023). We have previously shown that DA can be used globally to remove bias in CLM from poorly parameterized controls of carbon allocation, phenology; adjusting LAI by 23% to align with satellite observations results in an 18% reduction in global GPP and a 6% reduction in global latent heat estimated by CLM (Fox et al., 2022). However, state DA does not improve prognostic modelling of LAI and biomass so model development to better represent the controls of LAI remains a priority.

Correcting bias in LAI revealed that the processes controlling photosynthesis in CLM5.0 also appear to be inaccurate. After LAI bias was removed, GPP was significantly (78.2%, Table S3) higher than the FLUXCOM data product (Figure 5b). Comparing field measurements to model assumptions of photosynthetic parameters have revealed significant over-estimates of apparent Vcmax, Jmax and $\boldsymbol{\Phi}_{PSII}$ in CLM4.5 (Rogers et al., 2017; Rogers et al., 2019). In CLM5.0, Vcmax is now estimated by the Leaf Utilization of Nitrogen for Assimilation (LUNA) model (Ali et al., 2015). Photosynthetic nitrogen is allocated between Vcmax and Jmax depending on factors that influence the daily nitrogen use efficiency of each process (Ali et al., 2015, Lawrence et al., 2019). The performance of LUNA in this region was found to be questionable particularly with respect to temperature sensitivity and seasonal dynamics (Birch et al., 2021). In this study we used empirical estimates of low temperature inhibition of photosynthetic capacity (Rogers et al., 2019) in our reparameterization of photosynthesis in CLM. This approach reduced CLM's estimates of GPP to more closely match the FLUXCOM product (Figure 5b). Birch et al (2021) did not explore temperature inhibition of $\mathbf{\Phi}_{PSII}$ as we have done here, but by adjusting Vcmax downwards at low temperature, their approach has a similar effect of reducing GPP. Further work is needed to resolve how seasonal changes in nitrogen allocation and low temperatures influence photosynthetic capacity in the arctic. We suggest these investigations first ensure minimal bias in LAI.

Assimilating biomass had a modest effect on GPP because total biomass is a weaker constraint on leaf carbon than LAI and biomass is assimilated less frequently than LAI. The correlation between biomass and leaf carbon is not as strong as that between LAI and leaf carbon, and so, at the assimilation step leaf carbon is altered less by the change of biomass than by the change of LAI Moreover, biomass is only assimilated once per year compared to 45 times a year for LAI. Because the leaf carbon pool is rapidly changing in CLM, the impact of biomass on leaf carbon at the assimilation step goes away quickly. Assimilating biomass did influence biomass pools that change more slowly and influences wood (stem and root) carbon pools and decomposition (litter and soil) carbon pools. The change of tree height (Figure 3) induced by the change of biomass will alter momentum roughness length and displacement height. These are two key parameters in calculating wind, temperature, and humidity profiles of the surface boundary layer, which control the sensible and latent heat fluxes representing land-atmosphere interactions (Zeng et al., 1998). Also, tree height impacts the under-canopy atmospheric stability (Sakaguchi and Zeng, 2009). The decrease of tree height (Figure 3) caused by the change of biomass will decrease the under-canopy stability and increase the turbulent transfer coefficient, causing the heat and water vapor transfer from the ground to the canopy air to increase.

While both GPP and ecosystem respiration were reduced by data assimilation, ecosystem respiration was less improved due to the worse soil carbon pool and resulted in worse NEE (Figure 4). Assimilating biomass caused a decrease in aboveground respiration (improving ecosystem respiration Figure S3b in Supporting Information S1) but was less successful in constraining belowground carbon stocks (Figure S4 in Supporting Information S1). Greater improvements in ER and more realistic soil carbon may be achieved by assimilating soil carbon observations into CLM to constrain soil carbon directly, though data are limited, and high uncertainty remains a concern (Jackson et al., 2017). The slight increase in soil carbon introduced by DA, the subsequent increased soil respiration (Figure S4 in Supporting Information S1) was likely caused by disequilibrium in soil carbon. Limited by computational resources, we were unable to spin up each of the 40 members in the ensemble to equilibrium individually. We assumed the equilibrium states of each ensemble member were similar and spun up one ensemble member for over 1000 years to quasi-equilibrium, as the initial condition to spin up each of the 40 members. This caused soil carbon to be in disequilibrium across approximately one third of the study domain. Coupling the EAKF with faster approaches to model spin-up (Liao et al., 2023) could allow defensible initialization while also allowing model states to be in disequilibrium (Luo et al., 2015).

5 Conclusion

Predicting ecosystem responses to environmental change relies on understanding many related processes simultaneously and, because many processes are imperfectly understood or difficult of parameterize, simplifications are necessary. In general, highly simplified model processes of leaf phenology, leaf carbon allocation, and turnover interact to predict LAI using some simple assumptions. The LSMs compared within the CMIP6 protocol show broad agreement in land carbon storage under historical conditions but projected annual carbon land-atmosphere flux in the ensemble ranged from approximately 0 to 15 PgC year^{-1} after 140 years (Spafford and MacDougal, 2021). GPP was the most common benchmark presented by the CMIP6 modeling teams, with LAI evaluated less frequently (Spafford and MacDougal, 2021). If different LSMs have altered photosynthetic controls to counterbalance persistent issues with LAI (Song et al., 2021), this would explain the difference between historical and future performance. It is challenging to correctly evaluate the implementation of photosynthetic processes in models that incorrectly estimate LAI. Our work suggests that DA can facilitate model development by overcoming model bias in highly uncertain processes. It also underscores the need for progress in understanding phenology, leaf carbon allocation and turnover. Given the mechanistic connection between carbon stocks and processes controlling carbon, water, and energy fluxes, improving model predictions of carbon stocks remains a priority in biogeochemical research.

Acknowledgements

XLH, AMF, HD, CD, WJG, WKS and DJPM gratefully acknowledge funding support from NASA Terrestrial Ecology Grant 80NSSC19M0103 and high-performance computing support from Cheyenne (https://doi.org/10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. AR was supported by the Next-Generation Ecosystem Experiments (NGEE Arctic) project that is supported by the Office of Biological and Environmental Research in the Department of Energy, Office of Science, and through the United States Department of Energy contracts No. DE-SC0012704 to Brookhaven National Laboratory and No. DE-AC02-05CH11231 to Lawrence Berkeley National Laboratory.

Open Research

Files and scripts (Huo 2024) used to conduct model experiments, generate figures and perform statistical analysis are archived at https://doi.org/10.5281/zenodo.10480817. The README in the repository provides detailed descriptions of each folder and outlines the connection between files within each folder. All data (Huo 2024) are archived on CyVerse https://data.cyverse.org/dav-anon/iplant/home/huoxl90/ABoVE_DA_Data. The slightly modified ILAMB code (Huo 2024) used to evaluate model performance are available at https://doi.org/10.5281/zenodo.10480704, and model simulations (Huo 2024) which were compared to the benchmark data and the comprehensive evaluation result are accessible from: https://data.cyverse.org/dav-anon/iplant/home/huoxl90/ABoVE_DA_Data/ILAMB_ModelData_Results/. The CLM model version used in the study is a developing version of CLM5.1 and was slightly modified (Huo 2024) to comply with the purpose of data assimilation (https://doi.org/10.5281/zenodo.10480768).

References

Ajami, H., Evans, J. P., McCabe, M. F., & Stisen, S. (2014). Technical Note: Reducing the spin-up time of integrated surface water–groundwater models. *Hydrol. Earth Syst. Sci.*, *18*(12), 5169-5179. https://doi.org/10.5194/hess-18-5169-2014

Albergel, C., Calvet, J. C., de Rosnay, P., Balsamo, G., Wagner, W., Hasenauer, S., Naeimi, V., Martin, E., Bazile, E., Bouyssel, F., & Mahfouf, J. F. (2010). Cross-evaluation of modelled and remotely sensed surface soil moisture with in situ data in southwestern France. *Hydrol. Earth Syst. Sci.*, *14*(11), 2177-2191. https://doi.org/10.5194/hess-14-2177-2010

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbu, A. L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Le Moigne, P., Decharme, B., Mahfouf, J. F., & Calvet, J. C. (2017). Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area. *Geosci. Model Dev.*, *10*(10), 3889-3912. https://doi.org/10.5194/gmd-10-3889-2017

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

Anderson, J. L. (2007). An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus A: Dynamic Meteorology and Oceanography*, *59*(2), 210-224. https://doi.org/10.1111/j.1600-0870.2006.00216.x

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

Bacour, C., Peylin, P., MacBean, N., Rayner, P. J., Delage, F., Chevallier, F., Weiss, M., Demarty, J., Santaren, D., Baret, F., Berveiller, D., Dufrêne, E., & Prunet, P. (2015). Joint assimilation of eddy covariance flux measurements and FAPAR products over temperate forests within a process-oriented biosphere model. *Journal of Geophysical Research: Biogeosciences*, *120*(9), 1839-1857. https://doi.org/10.1002/2015JG002966

Bernacchi, C. J., Singsaas, E. L., Pimentel, C., Portis Jr, A. R., & Long, S. P. (2001). Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell & Environment*, *24*(2), 253-259. https://doi.org/10.1111/j.1365-3040.2001.00668.x

Birch, L., Schwalm, C. R., Natali, S., Lombardozzi, D., Keppel-Aleks, G., Watts, J., Lin, X., Zona, D., Oechel, W., Sachs, T., Black, T. A., & Rogers, B. M. (2021). Addressing biases in Arctic–boreal carbon cycling in the Community Land Model Version 5. *Geosci. Model Dev.*, *14*(6), 3361-3382. https://doi.org/10.5194/gmd-14-3361-2021

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., Melton, J. R., Pongratz, J., Turton, R. H., Yoshimura, K., & Yuan, H. (2021). Advances in Land Surface Modelling. *Current Climate Change Reports*, *7*(2), 45-71. https://doi.org/10.1007/s40641-021-00171-5

Bolhar-Nordenkampf, H. R., Hofer, M., & Lechner, E. G. (1991). Analysis of light-induced reduction of the photochemical capacity in field-grown plants. Evidence for photoinhibition? *Photosynth Res*, *27*(1), 31-39. https://doi.org/10.1007/bf00029974

Bonan, G. (2019). *Climate Change and Terrestrial Ecosystem Modeling*. Cambridge University Press. https://doi.org/10.1017/9781107339217

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M., & Swenson, S. C. (2011). Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *Journal of Geophysical Research: Biogeosciences*, *116*(G2). https://doi.org/10.1029/2010JG001593

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. *Journal of Geophysical Research: Biogeosciences*, *117*(G2). https://doi.org/10.1029/2011JG001913

Boussetta, S., Balsamo, G., Dutra, E., Beljaars, A., & Albergel, C. (2015). Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction. *Remote Sensing of Environment*, *163*, 111-126. https://doi.org/10.1016/j.rse.2015.03.009

Braghiere, R. K., Fisher, J. B., Miner, K. R., Miller, C. E., Worden, J. R., Schimel, D. S., & Frankenberg, C. (2023). Tipping point in North American Arctic-Boreal carbon sink persists in new generation Earth system models despite reduced uncertainty. *Environmental Research Letters*, *18*(2), 025008. https://doi.org/10.1088/1748-9326/acb226

Brown, S. (2002). Measuring carbon in forests: current status and future challenges. *Environmental Pollution*, *116*(3), 363-372. https://doi.org/10.1016/S0269-7491(01)00212-3

Caspersen, J. P., Pacala, S. W., Jenkins, J. C., Hurtt, G. C., Moorcroft, P. R., & Birdsey, R. A. (2000). Contributions of Land-Use History to Carbon Accumulation in U.S. Forests. *Science*, *290*(5494), 1148-1151. https://doi.org/doi:10.1126/science.290.5494.1148

Collatz, G. J., Ball, J. T., Grivet, C., & Berry, J. A. (1991). Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agricultural and Forest Meteorology*, *54*(2), 107-136. https://doi.org/10.1016/0168-1923(91)90002-8

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., & Randerson, J. T. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2731-2754. https://doi.org/10.1029/2018MS001354

Ds345.0. (2020). *CAM6 Data Assimilation Research Testbed (DART) Reanalysis* Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/JG1E-8525

Delbart, N., Ciais, P., Chave, J., Viovy, N., Malhi, Y., & Le Toan, T. (2010). Mortality as a key driver of the spatial distribution of aboveground biomass in Amazonian forest: results from a dynamic vegetation model. *Biogeosciences*, *7*(10), 3027-3039. https://doi.org/10.5194/bg-7-3027-2010

Demarty, J., Chevallier, F., Friend, A. D., Viovy, N., Piao, S., & Ciais, P. (2007). Assimilation of global MODIS leaf area index retrievals within a terrestrial biosphere model. *Geophysical Research Letters*, *34*(15). https://doi.org/10.1029/2007GL030014

Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. A., Chazdon, R. L., Weishampel, J. F., & Prince, S. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, *79*(2), 305-319. https://doi.org/10.1016/S0034-4257(01)00281-4

El Gharamti, M., Raeder, K., Anderson, J., & Wang, X. (2019). Comparing Adaptive Prior and Posterior Inflation for Ensemble Filters Using an Atmospheric General Circulation Model. *Monthly Weather Review*, *147*(7), 2535-2553. https://doi.org/10.1175/MWR-D-18-0389.1

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C 3 species. *Planta*, *149*(1), 78-90. https://doi.org/10.1007/bf00386231

Forzieri, G., Duveiller, G., Georgievski, G., Li, W., Robertson, E., Kautz, M., Lawrence, P., Garcia San Martin, L., Anthoni, P., Ciais, P., Pongratz, J., Sitch, S., Wiltshire, A., Arneth, A., & Cescatti, A. (2018). Evaluating the Interplay Between Biophysical Processes and Leaf Area Changes in Land Surface Models. *Journal of Advances in Modeling Earth Systems*, *10*(5), 1102-1126. https://doi.org/10.1002/2018MS001284

Fox, A. M., Hoar, T. J., Anderson, J. L., Arellano, A. F., Smith, W. K., Litvak, M. E., MacBean, N., Schimel, D. S., & Moore, D. J. P. (2018). Evaluation of a Data Assimilation System for Land Surface Models Using CLM4.5. *Journal of Advances in Modeling Earth Systems*, *10*(10), 2471-2494. https://doi.org/10.1029/2018MS001362

Fox, A. M., Huo, X., Hoar, T. J., Dashti, H., Smith, W. K., MacBean, N., Anderson, J. L., Roby, M., & Moore, D. J. P. (2022). Assimilation of Global Satellite Leaf Area Estimates Reduces Modeled Global Carbon Uptake and Energy Loss by Terrestrial Ecosystems. *Journal of Geophysical Research: Biogeosciences*, *127*(8), e2022JG006830. https://doi.org/10.1029/2022JG006830

Franklin, O., Johansson, J., Dewar, R. C., Dieckmann, U., McMurtrie, R. E., Brännström, Å., & Dybzinski, R. (2012). Modeling carbon allocation in trees: a search for principles. *Tree Physiology*, *32*(6), 648-666. https://doi.org/10.1093/treephys/tpr138

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., . . . Woodward, F. I. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO₂. *Proceedings of the National Academy of Sciences*, *111*(9), 3280-3285. https://doi.org/doi:10.1073/pnas.1222477110

Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, *125*(554), 723-757. https://doi.org/10.1002/qj.49712555417

Gower, S. T., Krankina, O., Olson, R. J., Apps, M., Linder, S., & Wang, C. (2001). Net Primary Production and Carbon Allocation Patterns of Boreal Forest Ecosystems. *Ecological Applications*, *11*(5), 1395-1411. https://doi.org/10.2307/3060928

Groom, Q. J., & Baker, N. R. (1992). Analysis of Light-Induced Depressions of Photosynthesis in Leaves of a Wheat Crop during the Winter. *Plant Physiol*, *100*(3), 1217-1223. https://doi.org/10.1104/pp.100.3.1217

Hoffman, F. M., Vertenstein, M., Kitabata, H., & White, J. B. (2005). Vectorizing the Community Land Model. *The International Journal of High Performance Computing Applications*, *19*(3), 247-260. https://doi.org/10.1177/1094342005056113

Houghton, R. A. (2005). Aboveground Forest Biomass and the Global Carbon Balance. *Global Change Biology*, *11*(6), 945-958. https://doi.org/10.1111/j.1365-2486.2005.00955.x

Huang, X., Cheng, F., Wang, J. L., Yi, B. J., & Bao, Y. L. (2023). Comparative Study on Remote Sensing Methods for Forest Height Mapping in Complex Mountainous Environments. *Remote Sensing*, *15*(9) 2275-2023. https://doi.org/10.3390/rs15092275

Huo, X. (2024). XueliHuo/ABoVE_DA_LAIandBiomass: ABoVE_DA_LAIandBiomass. [Software] *Zenodo*. https://doi.org/10.5281/zenodo.10480817

Huo, X. (2024). ABoVE_DA_Data. [Dataset] *Cyverse*. https://data.cyverse.org/dav-anon/iplant/home/huoxl90/ABoVE_DA_Data

Huo, X. (2024). XueliHuo/ILAMB: ILAMB_ABoVE_DA. [Software] *Zenodo*. https://doi.org/10.5281/zenodo.10480704

Huo, X. (2024). ABoVE_DA_Data/ILAMB_ModelData_Results. [Dataset] *Cyverse*. https://data.cyverse.org/dav-anon/iplant/home/huoxl90/ABoVE_DA_Data/ILAMB_ModelData_Results

Huo, X. (2024). XueliHuo/CTSM: CLM_ABoVE_DA. [Software] *Zenodo*. https://zenodo.org/records/10480768

Jackson, R. B., Lajtha, K., Crow, S. E., Hugelius, G., Kramer, M. G., & Piñeiro, G. (2017). The Ecology of Soil Carbon: Pools, Vulnerabilities, and Biotic and Abiotic Controls. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 419-445. https://doi.org/10.1146/annurev-ecolsys-112414-054234

Jeong, J.-H., Ho, C.-H., Chen, D., & Park, T.-W. (2008). Land Surface Initialization Using an Offline CLM3 Simulation with the GSWP-2 Forcing Dataset and Its Impact on CAM3 Simulations of the Boreal Summer Climate. *Journal of Hydrometeorology*, *9*(6), 1231-1248. https://doi.org/10.1175/2008JHM941.1

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., & Reichstein, M. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, *6*(1), 74. https://doi.org/10.1038/s41597-019-0076-8

Keith, H., Mackey, B. G., & Lindenmayer, D. B. (2009). Re-evaluation of forest biomass carbon stocks and lessons from the world's most carbon-dense forests. *Proceedings of the National Academy of Sciences*, *106*(28), 11635-11640. https://doi.org/doi:10.1073/pnas.0901970106

Kromdijk, J., Głowacka, K., Leonelli, L., Gabilly, S. T., Iwai, M., Niyogi, K. K., & Long, S. P. (2016). Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science*, *354*(6314), 857-861. https://doi.org/doi:10.1126/science.aai8878

Kumar, S. V., M. Mocko, D., Wang, S., Peters-Lidard, C. D., & Borak, J. (2019). Assimilation of Remotely Sensed Leaf Area Index into the Noah-MP Land Surface Model: Impacts on Water and Carbon Fluxes and States over the Continental United States. *Journal of Hydrometeorology*, *20*(7), 1359-1377. https://doi.org/10.1175/JHM-D-18-0237.1

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., . . . Zeng, X. (2019). The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245-4287. https://doi.org/10.1029/2018MS001583

Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., & Gower, S. T. (2002). Lidar Remote Sensing of Above-Ground Biomass in Three Biomes. *Global Ecology and Biogeography*, *11*(5), 393-399. http://www.jstor.org/stable/3182648

Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., Hunter, M. O., & de Oliveira Jr., R. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, *32*(22). https://doi.org/10.1029/2005GL023971

Liao, C., Lu, X., Huang, Y., Tao, F., Lawrence, D. M., Koven, C. D., Oleson, K. W., Wieder, W. R., Kluzek, E., Huang, X., & Luo, Y. (2023). Matrix Approach to Accelerate Spin-Up of CLM5. *Journal of Advances in Modeling Earth Systems*, *15*(8), e2023MS003625. https://doi.org/10.1029/2023MS003625

Ling, X. L., Fu, C. B., Guo, W. D., & Yang, Z.-L. (2019). Assimilation of Remotely Sensed LAI Into CLM4CN Using DART. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2768-2786. https://doi.org/10.1029/2019MS001634

Litton, C. M., Raich, J. W., & Ryan, M. G. (2007). Carbon allocation in forest ecosystems. *Global Change Biology*, *13*(10), 2089-2109. https://doi.org/10.1111/j.1365-2486.2007.01420.x

Lombardozzi, D. L., Bonan, G. B., Smith, N. G., Dukes, J. S., & Fisher, R. A. (2015). Temperature acclimation of photosynthesis and respiration: A key uncertainty in the carbon cycle-climate feedback. *Geophysical Research Letters*, *42*(20), 8624-8631. https://doi.org/10.1002/2015GL065934

Long, S. P., Humphries, S., & Falkowski, P. G. (1994). Photoinhibition of Photosynthesis in Nature. *Annual Review of Plant Physiology and Plant Molecular Biology*, *45*(1), 633-662. https://doi.org/10.1146/annurev.pp.45.060194.003221

Long, S. P., Postl, W. F., & Bolhár-Nordenkampf, H. R. (1993). Quantum yields for uptake of carbon dioxide in C_3 vascular plants of contrasting habitats and taxonomic groupings. *Planta*, *189*(2), 226-234. http://www.jstor.org/stable/23382109

Lunch, C., K. NEON (National Ecological Observatory Network). Data Tutorial: Compare tree height measured from the ground to a Lidar-based Canopy Height Model. https://www.neonscience.org/resources/learning-hub/tutorials/tree-heights-veg-structure-chm (accessed 02/15/2023).

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., & Schimel, D. S. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, *21*(5), 1429-1442. https://doi.org/10.1890/09-1275.1

Luo, Y., Keenan, T. F., & Smith, M. (2015). Predictability of the terrestrial carbon cycle. *Global Change Biology*, *21*(5), 1737-1751. https://doi.org/10.1111/gcb.12766

Luyssaert, S., Inglima, I., Jung, M., Richardson, A. D., Reichstein, M., Papale, D., Piao, S. L., Schulze, E. D., Wingate, L., Matteucci, G., Aragao, L., Aubinet, M., Beer, C., Bernhofer, C., Black, K. G., Bonal, D., Bonnefond, J. M., Chambers, J., Ciais, P., . . . Janssens, I. A. (2007). CO2 balance of boreal, temperate, and tropical forests derived from a global database. *Global Change Biology*, *13*(12), 2509-2537. https://doi.org/10.1111/j.1365-2486.2007.01439.x

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

Meier, C. (2023). TOS Protocol and Procedure: VST - Measurement of Vegetation Structure. https://data.neonscience.org/api/v0/documents/NEON.DOC.000987vK

Montané, F., Fox, A. M., Arellano, A. F., MacBean, N., Alexander, M. R., Dye, A., Bishop, D. A., Trouet, V., Babst, F., Hessl, A. E., Pederson, N., Blanken, P. D., Bohrer, G., Gough, C. M., Litvak, M. E., Novick, K. A., Phillips, R. P., Wood, J. D., & Moore, D. J. P. (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. *Geosci. Model Dev.*, *10*(9), 3499-3517. https://doi.org/10.5194/gmd-10-3499-2017

NEON (National Ecological Observatory Network). Vegetation structure (DP1.10098.001), RELEASE-2023. https://doi.org/10.48443/73zn-k414

NEON (National Ecological Observatory Network). Ecosystem structure (DP3.30015.001), RELEASE-2023. https://doi.org/10.48443/y26y-sj42

Norton, A. J., Bloom, A. A., Parazoo, N. C., Levine, P. A., Ma, S., Braghiere, R. K., & Smallman, T. L. (2023). Improved process representation of leaf phenology significantly shifts climate sensitivity of ecosystem carbon balance. *Biogeosciences*, *20*(12), 2455-2484. https://doi.org/10.5194/bg-20-2455-2023

National Ecological Observatory Network (NEON). (2023). *Ecosystem structure (DP3.30015.001)* https://doi.org/10.48443/Y26Y-SJ42

Oleson, K., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., … Yang, Z. -L. (2013). Technical description of version 4.5 of the Community Land Model (CLM) (No. NCAR/TN-503+STR). http://dx.doi.org/10.5065/D6RR1W7M

Ogren, E., & Evans, J. (1992). Photoinhibition of Photosynthesis *in situ* in Six Species of *Eucalyptus*. *Functional Plant Biology*, *19*(3), 223-232. https://doi.org/10.1071/PP9920223

Raczka, B., Hoar, T. J., Duarte, H. F., Fox, A. M., Anderson, J. L., Bowling, D. R., & Lin, J. C. (2021). Improving CLM5.0 Biomass and Carbon Exchange Across the Western United States Using a Data Assimilation System. *Journal of Advances in Modeling Earth Systems*, *13*(7), e2020MS002421. https://doi.org/10.1029/2020MS002421

Raeder, K., Hoar, T. J., El Gharamti, M., Johnson, B. K., Collins, N., Anderson, J. L., Steward, J., & Coady, M. (2021). A new CAM6 + DART reanalysis with surface forcing from CAM6 to other CESM models. *Scientific Reports*, *11*(1), 16384. https://doi.org/10.1038/s41598-021-92927-0

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., von Caemmerer, S., Dietze, M. C., Kattge, J., Leakey, A. D. B., Mercado, L. M., Niinemets, Ü., Prentice, I. C., Serbin, S. P., Sitch, S., Way, D. A., & Zaehle, S. (2017). A roadmap for improving the representation of photosynthesis in Earth system models. *New Phytologist*, *213*(1), 22-42. https://doi.org/10.1111/nph.14283

Rogers, A., Serbin, S. P., Ely, K. S., & Wullschleger, S. D. (2019). Terrestrial biosphere models may overestimate Arctic CO2 assimilation if they do not account for decreased quantum yield and convexity at low temperature. *New Phytologist*, *223*(1), 167-179. https://doi.org/10.1111/nph.15750

Sakaguchi, K., & Zeng, X. (2009). Effects of soil wetness, plant litter, and under-canopy atmospheric stability on ground evaporation in the Community Land Model (CLM3.5). *Journal of Geophysical Research: Atmospheres*, *114*(D1). https://doi.org/https://doi.org/10.1029/2008JD010834

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, *116*(G4). https://doi.org/10.1029/2011JG001708

Singsaas, E. L., Ort, D. R., & DeLucia, E. H. (2001). Variation in measured values of photosynthetic quantum yield in ecophysiological studies. *Oecologia*, *128*(1), 15-23. https://doi.org/10.1007/s004420000624

Song, X., Wang, D.-Y., Li, F., & Zeng, X.-D. (2021). Evaluating the performance of CMIP6 Earth system models in simulating global vegetation structure and distribution. *Advances in Climate Change Research*, *12*(4), 584-595. https://doi.org/10.1016/j.accre.2021.06.008

Spafford, L., & MacDougall, A. H. (2021). Validation of terrestrial biogeochemistry in CMIP6 Earth system models: a review. *Geosci. Model Dev.*, *14*(9), 5863-5889. https://doi.org/10.5194/gmd-14-5863-2021

Stöckli, R., Lawrence, D. M., Niu, G.-Y., Oleson, K. W., Thornton, P. E., Yang, Z.-L., Bonan, G. B., Denning, A. S., & Running, S. W. (2008). Use of FLUXNET in the Community Land Model development. *Journal of Geophysical Research: Biogeosciences*, *113*(G1). https://doi.org/10.1029/2007JG000562

Takagi, K., Yone, Y., Takahashi, H., Sakai, R., Hojyo, H., Kamiura, T., Nomura, M., Liang, N., Fukazawa, T., Miya, H., Yoshida, T., Sasa, K., Fujinuma, Y., Murayama, T., & Oguma, H. (2015). Forest biomass and volume estimation using airborne LiDAR in a cool-temperate forest of northern Hokkaido, Japan. *Ecological Informatics*, *26*, 54-60. https://doi.org/10.1016/j.ecoinf.2015.01.005

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., & Papale, D. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, *13*(14), 4291-4313. https://doi.org/10.5194/bg-13-4291-2016

Wang, J., Farina, M. K., Baccini, A., & Friedl, M. A. (2021). ABoVE: Annual Aboveground Biomass for Boreal Forests of ABoVE Core Domain, 1984-2014. *ORNL DAAC*, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1808

Wieder, W. R., Lawrence, D. M., Fisher, R. A., Bonan, G. B., Cheng, S. J., Goodale, C. L., Grandy, A. S., Koven, C. D., Lombardozzi, D. L., Oleson, K. W., & Thomas, R. Q. (2019). Beyond Static Benchmarking: Using Experimental Manipulations to Evaluate Land Model Assumptions. *Global Biogeochemical Cycles*, *33*(10), 1289-1309. https://doi.org/10.1029/2018GB006141

Zeng, X., Zhao, M., & Dickinson, R. E. (1998). Intercomparison of Bulk Aerodynamic Algorithms for the Computation of Sea Surface Fluxes Using TOGA COARE and TAO Data. *Journal of Climate*, *11*(10), 2628-2644. https://doi.org/10.1029/2008JD010834

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R., & Myneni, R. B. (2013). Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011. *Remote Sensing*, *5*(2), 927-948. https://www.mdpi.com/2072-4292/5/2/927