Quantifying and reducing uncertainty in global carbon cycle predictions: lessons and perspectives from 15 years of data assimilation studies with the ORCHIDEE Terrestrial Biosphere Model

N. Macbean, C. Bacour, N. Raoult, V. Bastrikov, E. Koffi, S. Kuppel, F. Maignan, Catherine Ottle, M. Peaucelle, D. Santaren, et al.

Global Biogeochemical Cycles

REVIEW ARTICLE
10.1029/2021GB007177

Special Section:
Understanding carbon-climate feedbacks

Quantifying and Reducing Uncertainty in Global Carbon Cycle Predictions: Lessons and Perspectives From 15 Years of Data Assimilation Studies With the ORCHIDEE Terrestrial Biosphere Model

N. MacBean[1], C. Bacour[2,3], N. Raoult[3], V. Bastrikov[4], E. N. Koffi[5], S. Kuppel[6], F. Maignan[3], C. Ottlé[3], M. Peaucelle[7,8], D. Santaren[3], and P. Peylin[3]

[1]Department of Geography, Indiana University, Bloomington, IN, USA, [2]NOVELTIS, Labège, Toulouse, France, [3]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvett, France, [4]Science Partners, Paris, France, [5]European Commission Joint Research Centre, Ispra, Varese, Italy, [6]Géosciences Environnement Toulouse, CNRS - IRD - UPS - CNES, Gif-sur-Yvett, France, [7]Computational & Applied Vegetation Ecology Laboratory (CAVELAB), Ghent Universtiy, Ghent, Belgium, [8]INRAE, Université de Bordeaux, UMR ISPA, Villenave-d'Ornon, Bordeaux, France

Key Points:

- Considerable progress has been made in constraining ORCHIDEE terrestrial biosphere model regional to global scale CO_2 fluxes within a data assimilation system
- Results highlight the importance of optimizing initial C stocks - in addition to C cycle related parameters - using global-scale datasets
- Challenges remain in utilizing the wide variety of available data, particularly when characterizing the observation error covariance matrix

Correspondence to:
N. MacBean,
nlmacbean@gmail.com

Citation:
MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E. N., Kuppel, S., et al. (2022). Quantifying and reducing uncertainty in global carbon cycle predictions: Lessons and perspectives from 15 Years of data assimilation studies with the ORCHIDEE terrestrial biosphere model. *Global Biogeochemical Cycles*, *36*, e2021GB007177. https://doi.org/10.1029/2021GB007177

Received 29 AUG 2021
Accepted 20 MAY 2022

Abstract Predicting terrestrial carbon, C, budgets and carbon-climate feedbacks strongly relies on our ability to accurately model interactions between vegetation, C and water cycles, and the atmosphere. However, C fluxes simulated by global, process-based terrestrial biosphere models (TBMs) remain subject to large uncertainties, partly due to unknown or poorly calibrated parameters. This is because TBMs have not routinely been confronted against C cycle related datasets within a statistical data assimilation (DA) system. In this review, we present 15 years' development of a C cycle DA system for optimizing C cycle parameters of the ORCHIDEE TBM. We analyze the impact of assimilating multiple different C cycle related datasets on regional to global-scale gross and net CO_2 fluxes. We find that assimilating atmospheric CO_2 data is crucial for improving (increasing) ORCHIDEE predictions of the terrestrial land C sink. The improvement is predominantly due to the global-scale constraint these data provide for optimizing initial soil C stocks, which are likely in error due to inaccurate assumptions about steady state spin-up and incomplete knowledge of land use change histories. When comparing the data-constrained ORCHIDEE land C sink estimates to the CAMS atmospheric inversion, we show that while the two approaches agree on the global C sink magnitude, they continue to differ in how the global C sink is partitioned between the northern hemisphere and tropics. We also discuss technical challenges faced in our C cycle DA studies, in particular the difficulty in characterizing the error covariance matrix due to unknown observation biases and/or model-data inconsistencies. We offer our perspectives on how to tackle these challenges that we hope can serve as a roadmap for other TBM groups wishing to develop C cycle DA systems.

1. Introduction

The terrestrial biosphere absorbs ~30% of anthropogenic CO_2 emissions, thus acting as a global carbon, C, sink (Friedlingstein et al., 2019). The terrestrial C sink is a result of complex interactions between C, nutrient, and water cycles in response to a variety of environmental change drivers, including climate change, rising CO_2, and land use and land cover change. Multiple different feedbacks between the C cycle and climate act to moderate the balance between the land C sink and the anthropogenic CO_2 and other greenhouse gas emissions remaining in the atmosphere (Canadell et al., 2021). Accurate representation of C cycle related processes in terrestrial biosphere models (TBMs), which form the land component of earth system models (ESMs), is therefore crucial for improving climate change projections. Key C cycle related processes such as photosynthesis, respiration, stomatal conductance, and soil C cycling (Ball et al., 1987; Farquhar et al., 1980; Parton et al., 1987) were originally developed for TBMs in the late 1980s/early 1990s (Pitman, 2003; Sellers et al., 1997). Since then, numerous additional vegetation and biogeochemical processes have been, and continue to be, implemented in these models (Blyth et al., 2021; Bonan & Doney, 2018; Fisher & Koven, 2020). In parallel, many studies have evaluated TBM C cycle simulations against observations with the aim of understanding the causes of individual model errors and inter-model spread (Anav, Friedlingstein et al., 2013; Keenan et al., 2012; Peng et al., 2015; Piao et al., 2013). Even after two decades of development, there remains high inter-model spread in all aspects of C cycle, including

phenology (Richardson et al., 2012; Anav, Murray-Tortarolo et al., 2013), gross CO_2 uptake (or gross primary productivity, (GPP), Anav et al., 2015), vegetation and soil C stocks and residence time (Friend et al., 2014; Montané et al., 2017; Todd-Brown et al., 2013), the response of biogeochemical cycles to elevated CO_2 or changes in climate (De Kauwe et al., 2014; Huntzinger et al., 2017; Paschalis et al., 2020; Walker et al., 2014; Zaehle et al., 2014), and future projections of global net CO_2 fluxes (cf. Friedlingstein et al., 2014; and, 2006; Arora et al., 2020). These TBM evaluations are typically performed as part of model intercomparison projects (MIPs, e.g., Sitch et al., 2015) and thus have the same climate forcing, the same prescribed vegetation and soil texture maps, and the same simulation protocol. Model errors are therefore due to inaccurate model structure (i.e., which processes are included and the representation of those processes), and/or errors in the fixed values, or parameters, of the model. There is limited understanding about which processes, or which type of model errors, are contributing the most to inter-model spread in global C flux dynamics.

Similarly, these "bottom-up" process-based TBMs are not always in agreement with "top-down" atmospheric inversions on where global C sink is concentrated, the regions contributing to global C cycle interannual variability (IAV), and what the future trend of the land C sink will be (Kondo et al., 2020). Even though the two different approaches roughly agree on the overall magnitude of the current global C sink (Anav, Friedlingstein et al., 2013; Friedlingstein et al., 2019; Kondo et al., 2020), they tend to disagree on regional C budget estimates, and uncertainties in both approaches are high (Anav, Friedlingstein et al., 2013; Huntzinger et al., 2017; Peylin et al., 2013; Schimel et al., 2015). TBMs typically estimate an approximately equal land-atmosphere CO_2 flux in both the tropics and northern hemisphere (NH) (Sitch et al., 2015). In contrast, atmospheric inversions tend to place a much stronger C sink in the NH compared to TBMs, whereas in the tropics they estimate an approximately neutral net CO_2 flux (or small C source) (Anav, Friedlingstein et al., 2013; Liu et al., 2021; Peylin et al., 2013), with a strong natural C sink (possibly due to CO_2 fertilization, Schimel et al., 2015) balanced by C emissions due to deforestation and forest degradation (Gatti et al., 2021; Qin et al., 2021). This remains true when using more spatially comprehensive satellite-based (OCO-2) CO_2 retrievals instead of ground-based tower CO_2 mole fraction data (Peiro et al., 2022). Even without considering the differences with atmospheric inversions, there is a huge uncertainty in long-term TBM/ESM projections of the net land CO_2 flux (~15 PgC spread across models; Arora et al., 2020), with some models even predicting the land may become a source of C to the atmosphere. While TBM simulations may agree with observations that the mean C sink and its long-term trend are likely dominated by high biomass forested regions (Ahlström et al., 2015; Schimel et al., 2015; Sitch et al., 2015), some modeling studies have shown that global C cycle IAV is dominated by semi-arid regions (Ahlström et al., 2015; Poulter et al., 2014; Zhang et al., 2018). However, these results are not consistently supported in the literature. Other studies suggest that tropical rainforests are the dominant contributor due to variability in water stores and fluxes (Worden et al., 2021). In their review, Piao et al. (2020) clearly point to tropical land ecosystems as driving global C cycle IAV, with approximately equal contributions from semi-arid and other tropical regions. Recent studies have shown that TBMs drastically underestimate semi-arid region net CO_2 flux IAV (MacBean et al., 2021; Teckentrup et al., 2021); thus, TBMs clearly require improvements for modeling semi-arid ecosystems, and it remains to be seen exactly what role that drylands play in driving global C cycle IAV.

While there are many studies that have documented model errors or inter-model spread, there are comparatively fewer studies that have used data to quantify and reduce uncertainty in those TBMs within a statistical data assimilation (DA) system. Bayesian DA methods allow model parameters or state variables to be optimized based on new information contained in the data via a statistical framework that considers uncertainties in both the model and the data (Rayner et al., 2019). However, most simulations that are used in MIPs and future climate change projections, such as the Coupled Model Intercomparison Project (CMIP) used for the climate change projections in IPCC reports, typically represent only one simulation following a given protocol - uncertainty due to different model structures or parameter errors is not assessed. For example, of the ~50 modeling groups participated in the latest CMIP (CMIP6), only a few have developed global scale C cycle DA systems (Albergel et al., 2017; Fox et al., 2018; Kaminski et al., 2013; Peylin et al., 2016; Raoult et al., 2016; Schürmann et al., 2016). This means that most of the TBMs used in IPCC climate change projections have not been widely confronted against data within a statistical optimization framework. We do not have any estimate of individual model uncertainty on global estimates of key C cycle variables, nor do we know if the inter-model spread described above could be reduced if the models were confronted against data within such a DA system.

Nevertheless, a growing number of groups are developing global scale C cycle DA systems, and the considerable progress made has been documented in numerous informative studies. Multiple different regional to global-scale data streams have been assimilated with the goal of either updating the state variables (e.g., Bonan et al., 2020; Fox et al., 2018; Ling et al., 2019; Quaife et al., 2008) or optimizing the model parameters (Bloom et al., 2016; Peylin et al., 2016; Rayner et al., 2005; Scholze et al., 2016; Schürmann et al., 2016). These C cycle DA systems have been used to constrain C allocation and residence time for different C pools (Bloom et al., 2016), to examine the impact of drought on vegetation (Albergel et al., 2019), to understand the role of forest disturbance due to management (Pinnington et al., 2017), and to better estimate regional to global C budgets and temporal variability in surface CO_2 fluxes and atmospheric CO_2 concentrations (Haverd et al., 2013; Konings, Bloom et al., 2019; Peylin et al., 2016; Rayner et al., 2005; Trudinger et al., 2016). More comprehensive reviews of C cycle related DA studies are provided by Rayner (2010), Kaminski et al. (2013), Scholze et al. (2017) and Exbrayat et al. (2019). Despite this progress, numerous DA technical and methodological challenges related remain, including: issues of biased observations and model-data inconsistencies, differences in assimilation results depending on the DA experimental configuration, and uncertainties in initial soil C stocks as a result of the steady-state assumption in model spin-up and lack of information on human activity over historical periods, to name just a few. Therefore, there is an urgent need to increase research into TBM DA system development and to encourage all modeling groups to contribute to these efforts. With growing numbers of datasets, longer observation record lengths, and rapidly increasing model complexity, there has never been a better time, nor more urgent need, for developing TBM C cycle DA systems. Understanding and constraining uncertainties in TBM predictions of key characteristics of the global C cycle is imperative if we are to accurately estimate long-term carbon-climate feedbacks. Without accurate projections of carbon-climate feedbacks, we cannot reliably estimate the level of CO_2 emissions that will keep us within a given threshold rise in global temperature (Jones & Friedlingstein, 2020).

In the hope of guiding other modeling groups who wish to develop TBM C cycle DA systems, in this review we document the progress made, highlight successful studies, and discuss the above mentioned challenges we faced when developing a global scale C cycle DA system to constrain C cycle-related parametric uncertainty in the ORCHIDEE (Organizing Carbon and Hydrology In Dynamic EcosystEms) TBM, which forms the land component of the French Institut Pierre Simon Laplace (IPSL) ESM (Dufresne et al., 2013). We document more than 15 years of C cycle DA studies ranging from the site scale to the globe. In Section 2, we briefly describe the ORCHIDEE DA system. In Sections 3 and 4, we present the range of C cycle DA studies conducted to date. Section 3 focuses on the global-scale C cycle assimilation studies that mostly constrained ORCHIDEE version AR5 (vAR5) using long-running C cycle datasets, while Section 4 highlights our ongoing, mostly site-scale assimilation studies with more recent versions of ORCHIDEE (v2.0+), in addition to various "branches" of the model that are under development, using novel datastreams and methods. We discuss the impact of our assimilation experiments on the key characteristics of the global C cycle discussed above and their expected impact on our understanding of carbon-climate feedbacks. We also provide an overview of our "DA methods" studies that have investigated technical issues related to multiple data stream assimilation and choice of optimization algorithm (Section 5). All these studies have provided us with a wide-ranging perspective on the challenges we face in constraining TBM parameters. We present our perspectives on these challenges, as well as our proposed solutions, in the final Section 6 of this review.

2. The ORCHIDEE Terrestrial Biosphere Model and the ORCHIDEE Data Assimilation System

2.1. ORCHIDEE Terrestrial Biosphere Model

ORCHIDEE is a global, process-based TBM that simulates the surface energy balance, hydrological and biogeochemical cycling, and the influence of anthropogenic activity on terrestrial ecosystems at a half hourly time step. In uncoupled mode, the model is run with meteorological forcing data or climate reanalyses (Table 1). All plant species are grouped into broad plant functional types (PFTs) depending on their physiological, structural, and phenological characteristics, as well as the main climate biomes in which they exist. ORCHIDEE vAR5 was used in CMIP5, which contributed to the IPCC fifth Assessment Report, and is described in (Krinner et al., 2005). The ORCHIDEE "trunk" (current main version of the model) has been updated since vAR5. ORCHIDEE v2.0, which is the version contributing to CMIP6, includes the following developments: (a) an 11-layer mechanistic description of soil hydrology and associated modifications as described in Ducharne et al. (in prep); (b) an

Global Biogeochemical Cycles

10.1029/2021GB007177

Table 1
Details of All Data Used in the ORCHIDEE C Cycle DA Studies, Including Assimilated Data and Meteorological Forcing

Data product	Temporal frequency	Observation period	Spatial resolution / footprint	No. of sites / grid cells	Satellite / network / product	Observation operator	ORCHIDEE version	Reference	ORCHIDEE DA Study
Assimilated data: Multi-site, multi-PFT global-scale studies									
Flux tower NEE and LE	Daily	1996–2006 (dependent on site)	~1–2km	78 (2–24/PFT)	FLUXNET	–	vAR5	Papale et al. (2006); https://fluxnet.org/data/la-thuile-dataset/	Kuppel et al. (2012; 2014) + various site-based studies
NDVI	Daily	2000–2008	0.72° × °0.72° (aggregated from 5km MOD09CMG)	90 (15/PFT)	MODIS	Normalized NDVI cf normalized modeled FAPAR	vAR5	https://lpdaac.usgs.gov/products/mod09cmgv006/	MacBean et al. (2015)
SIF	Monthly	2007–2011	0.5° × 0.5°	180 (15/PFT)	GOME-2	Linear relationship btw GPP and SIF	vAR5	Köhler et al. (2015)	MacBean et al. (2018)
SIF	Monthly	2015–2016	0.5° × 0.5°	210 (15/PFT)	OCO-2	Mechanistic SIF model	v2.0	Frankenberg et al. (2014); Sun et al. (2018); ftp://fluo.gps.caltech.edu/data/OCO2/sif_lite_B7300/	Bacour, Maignan, MacBean et al. (2019)
Atmospheric CO_2 mole fraction (conc)	Monthly	2002–2004	–	53	NOAA ESRL	LMDz atmospheric transport model + other CO2 emissions	vAR5	GLOBALVIEW-CO2 (2013); https://gml.noaa.gov/ccgg/flask.html	Peylin et al. (2016); Bacour et al. (in prep)
Assimilated Data: Site-scale studies									
Effective LAI	8 Days	2003–2014	0.5° km	10	–	1D 2-stream radiative transfer model	ORCHIDEE-CAN	EU H2020 MULTIPLY: http://www.multiply-h2020.eu	Unpublished
Total aboveground biomass and biomass increment	Annual	1995–2005; 1996–2007	0.6° ha; 4.7° ha	2	–	–	vAR5	Granier et al. (2008) and Porté et al. (2002)	Thum et al. (2017)
FAPAR	8 Days	2000–2010 (dependent on site)	40° m and 1° km	2	SPOT and MERIS	Beer-lambert law linking to LAI	vAR5	Bacour et al. (2006); Baret et al. (2007); Weiss et al. (2007)	Bacour et al. (2015)
Volumetric soil moisture	Daily	2008–2014 (dependent on site)	*In situ*	18	ISMN	–	v2.0	Dorigo et al. (2011; 2013); https://ismn.geo.tuwien.ac.at/en/	Raoult et al. (2021)

Global Biogeochemical Cycles

10.1029/2021GB007177

Table 1
Continued

Data product	Temporal frequency	Observation period	Spatial resolution / footprint	No. of sites / grid cells	Satellite / network / product	Observation operator	ORCHIDEE version	Reference	ORCHIDEE DA Study
Meteorological Forcing Data									
In situ meterological forcing	30 Minutes	1996–2014 (Dependent on site)	~1–2° km	Dependent on study	FLUXNET	–	Dependent on study	https://fluxnet.org/data/	Kuppel et al. (2012; 2014); Bacour et al. (2015); Raoult et al. (2021)
Climate reanalyses	3-hourly[b]; 6-hourly[a,c]	Dependent on study	0.72°[b]; 0.5°[a,c]	Global	ERA-Interim[b]; CRU-NCEP v5[c] and v7[a]	–	Dependent on study	Dee et al. (2011)[b]; Harris et al. (2014)[c]; Viovy (2018)[a]	MacBean et al. (2015)[b]; Peylin et al. (2016)[b]; MacBean et al. (2018)[c]; Bacour et al. (2019)[a]

[a] The 6-hourly 0.5° resolution CRU-NCEP v7 climate forcing data (Viovy, 2018) that was used in the Bacour et al. (2019) study reviewed in this paper.

[b] The 3-hourly 0.72° resolution ERA-Interim climate forcing data (Dee et al., 2011) that was used in the MacBean et al. (2015) and Peylin et al. (2016) studies reviewed in this paper.

[c] The 6-hourly 0.5° resolution CRU-NCEP v5 climate forcing data (Harris et al., 2014) that was used in the MacBean et al. (2018) study reviewed in this paper.

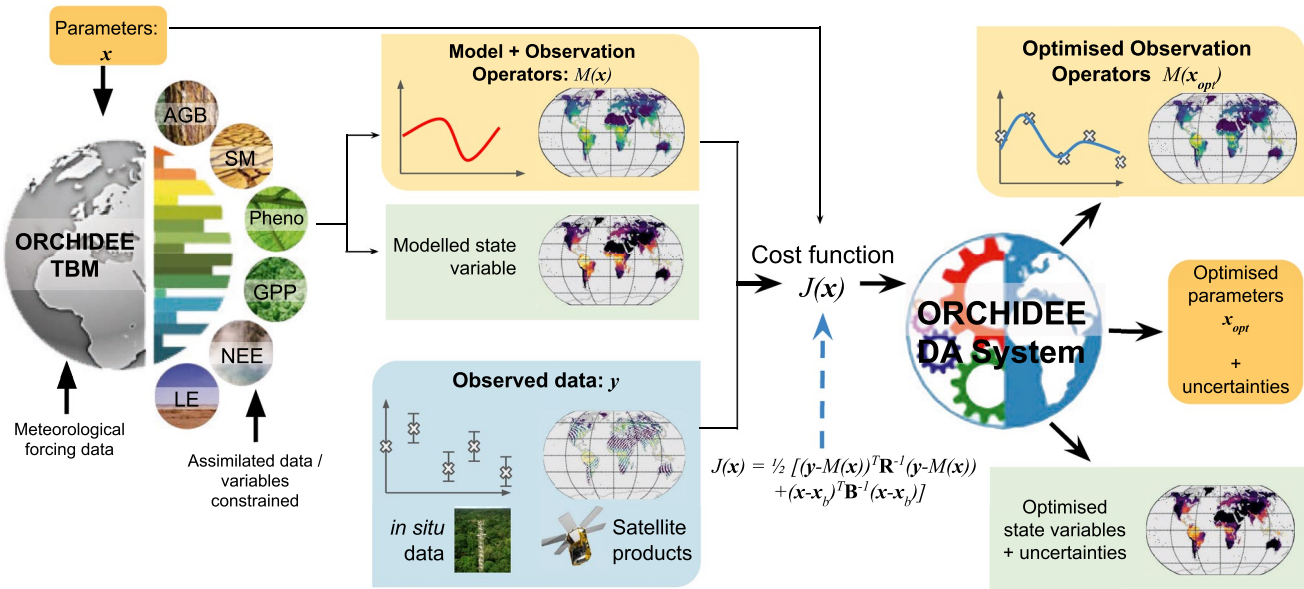

Figure 1. Schematic showing the different components of the ORCHIDEE Data Assimilation System.

analytical solution for the set of equations for photosynthesis, stomatal conductivity and internal CO_2 concentration in the leaf (described in Vuichard et al., 2019); (c) an update of the soil thermal properties and extension of the soil depth for heat diffusion (Wang et al., 2016); (d) a three-layer snow scheme (Wang et al., 2013); (e) a spatially explicit observation-derived estimate for background albedo and optimized vegetation and snow albedo coefficients; and (f) new PFT maps based on the European Space Agency Climate Change Initiative Land Cover product (Poulter et al., 2015). In addition to bare soil, there were 12 vegetated PFTs in vAR5, and from v2.0 onwards there are 15 vegetated PFTs. Different PFTs share the same processes but with different parameters, which are fixed values that are used in the model calculations, with the exception of the phenology schemes that are PFT dependent. The ORCHIDEE-CN-CAN version used to assimilate effective LAI and Free Air CO_2 Enrichment (FACE) experiment data in Section 4.3 has been updated to include: (a) addition of a coupled carbon-nitrogen scheme (Vuichard et al., 2019); and (b) a physical description of vertical canopy structure with a 2-stream radiative transfer (RT) scheme (in addition to other model developments related to forest management) (Naudts et al., 2015). In all assimilations reviewed here we do not account for any anthropogenic activity in the ORCHIDEE simulations, including land use change, agriculture and forest and grassland management.

2.2. ORCHIDAS System and Different Inversion Methods Used

The ORCHIDEE DA system is variational DA system in which all observations within the assimilation time window are included in the optimization (https://orchidas.lsce.ipsl.fr; Figure 1). The optimization of model parameters requires the "minimization" of a cost function, $\boldsymbol{J(x)}$ (Figure 1 - where $\boldsymbol{x}$ is the parameter vector), which describes the difference between the modeled and observed variables considering uncertainties in both. We use a Bayesian DA framework in which prior information on the model parameters (default model values and their uncertainties) is also included in $\boldsymbol{J(x)}$ (Figure 1). Statistical DA frameworks use an inversion algorithm to minimize the cost function, therefore finding the best fit between the model and data considering all uncertainties. Most of the ORCHIDEE C cycle DA studies have used the L-BFGS-B gradient descent inversion algorithm (Byrd et al., 1995), which requires the tangent linear (TL) or adjoint of the model to calculate the gradient of $\boldsymbol{J(x)}$ at each iteration of the inversion (further details can be found in Kuppel et al., 2014; Peylin et al., 2016). The algorithm assumes a quasi-linear model and Gaussian error distributions. Unlike some other gradient descent algorithms, L-BFGS-B allows for bounds on the parameters to be accounted for in the inversion. However, some studies detailed here have used the genetic algorithm (GA; Haupt & Haupt, 2004) to minimize the cost function (further details of the specific implementation with ORCHIDEE are provided in Santaren et al., 2014). The GA is a so-called "global search" method, in which the gradient of the cost function is not needed to find the minimum of $\boldsymbol{J(x)}$; instead, the algorithm efficiently searches the entire parameter space: At each iteration checks

whether the new parameter has reduced the model-data fit and then either accepts or rejects that parameter vector based on the improvement (or lack thereof) in the model-data fit. The accepted parameters form the posterior parameter probability distributions. The GA, like all global search methods, is more flexible in that it does not require assumptions of a quasi-linear model or Gaussian error distributions; however, it is computationally more expensive to run. We discuss our studies investigating the differences between these two classes of algorithm more in Section 5.3.

Prior parameter values are set as the default values in the ORCHIDEE model. The bounds are prescribed based on literature reviews and expertise of the ORCHIDEE Project Team. The prior uncertainty is typically assumed to be 40% of the parameter value range following Kuppel et al. (2012). Unless otherwise specified, the observation error covariance matrix (**R**, Figure 1) in the ORCHIDEE DA system is based on the RMSE between the observations and the model to take into account the fact that **R** should account for both model and observation error. In all our DA studies, the observation (including model) and parameter errors are assumed to be uncorrelated and therefore both **R** and B (prior parameter error covariance matrix) are diagonal matrices. We discuss the implications of this simple approach to characterizing **R** and **B** in various sections throughout this paper. We calculate the posterior errors on the assimilated variables, $\mathbf{R}_{\mathbf{post}}$, using the posterior parameter covariance matrix and the assumption of local linearity (Tarantola, 2013).

2.3. Datasets Used in Assimilation Studies

Many C cycle related datasets at different temporal and spatial resolutions have been used to constrain the C cycle parameters of ORCHIDEE. Table 1 summarizes the data used in both the global and site scale key studies discussed in this review, in addition to the observation operator that is used to link the data to the ORCHIDEE model processes. Further details on the data characteristics, pre-processing prior to assimilation, and detailed descriptions of the observation operators can be found in the respective ORCHIDEE DA studies.

2.4. Assimilation Experimental Set-Up

Specific details of all studies described here can be found in each of the cited papers. In brief, any global-scale C cycle DA study that used site-based data (e.g., in Section 3.1), or selected grid cells from satellite datasets (Sections 3.2 and 3.3), included all sites in a so-called "multi-site" (MS) assimilation for each PFT. Some studies also compared the results of the MS assimilation to a case in which each site was assimilated separately ("single site", SS, assimilations). The number of flux tower sites (see Table 1) included for each PFT was dependent on the available data. Satellite grid cells were randomly chosen from a selection of locations that had a majority fractional cover of the PFT being optimized (deciduous forest and grass PFTs only). To simulate atmospheric CO_2 concentration (Section 3.4), we coupled the ORCHIDEE model to the LMDz atmospheric transport model (Hourdin et al., 2006). Note that the coupling with LMDZ involved the use of pre-calculated transport response functions (i.e., sensitivity of atmospheric concentrations at each station to surface fluxes) as detailed in Peylin et al. (2016). We also added to the transport model the other main carbon cycle flux components: that is, fossil fuel emissions, ocean flux estimates based on pCO_2 ocean data, and biomass burning emissions (with partial forest regrowth also factored in) (see Peylin et al., 2016 for details). Inverting the coupled ORCHIDEE-LMDz model allowed us to use the atmospheric CO_2 concentration data as a much broader scale constraint on net CO_2 surface fluxes (and related parameters).

Due to computational constraints (i.e., the complexity of inverting large matrices when all observations are used simultaneously), the first version of our multiple data stream assimilation system that assimilated satellite, flux tower, and atmospheric CO_2 concentration data used a so-called step-wise approach to assimilate all three data-streams (Peylin et al., 2016). In this step-wise version of the system, the three datasets were assimilated in the following order: in step 1 satellite NDVI data were assimilated to constrain leaf phenology parameters [i.e., the same set-up as described in MacBean et al., (2015)]; in step 2 the flux tower NEE and LE data were used to constrain photosynthesis, respiration, and phenology parameters, in addition to some parameters related to water fluxes (as described in Kuppel et al., 2014); and finally, in step 3, the atmospheric CO_2 data were used to further constrain the same set of parameters as in step 2 but with the addition of regional "K_{soilC}" parameters (30 spatially coherent regions), which act as scalars on the slow and passive soil C pool at the beginning of the assimilation window (Peylin et al., 2016). The scaling of the slow and passive soil C pools (with decade to century scale

turnover times) by K_{soilC} is designed to account for biases in soil C content due to the fact we cannot accurately simulate historical changes in land cover and land use and their impact on C cycling during the spin-up procedure. We optimize a regional scale K_{soilC}, as opposed to a site-based K_{soilC}, for two reasons: first, because it is impossible to scale either SS, or even MS, estimates of K_{soilC} up to broader scales and second, because the land use history and resulting edaphic conditions at each site are likely highly site specific and thus not generalizable to broader scales. The 30 chosen regions were designed for potentially coherent land use and land management history. The posterior parameter and error covariance matrices were propagated at each step and the posterior parameter values from the previous step were used as prior values in the following step. In Section 3.4 we compare the results of the step-wise assimilation of all three datastreams to a "simultaneous" assimilation in which all three datastreams are assimilated in one assimilation experiment.

The parameters selected for optimization were chosen based on site-based sensitivity analyses using the Morris (1991) method as well as expert judgment and prior model testing. Full lists of parameters can be found in each of the studies, but typically included parameters related to leaf phenology (e.g., temperature and moisture thresholds controlling the start of leaf growth and senescence and leaf age), photosynthesis (e.g., maximum carboxylation rate and temperature optima for photosynthesis), soil water availability (e.g., empirical controls on rooting depth profile and water stress functions), autotrophic respiration (e.g., fraction of biomass for growth respiration and temperature dependence of maintenance respiration), and heterotrophic respiration and soil decomposition (e.g., soil Q10, moisture limitation on soil decomposition, litter height), and energy balance (e.g., the reference surface roughness and albedo). In addition, various parameters of the observation operators that link the ORCHIDEE model processes to the respective datasets were included in the assimilations.

Validation of the simulations with optimized parameters was performed in each of the studies either using additional time periods for the sites assimilated, or additional sites that were not included in the assimilation, or both. Validation against independent datasets for variables not included in the assimilation was also performed for many studies. These validation results are not reviewed here but details can be found in each of the studies cited.

3. Highlights of Global-Scale C Cycle DA Studies With ORCHIDEE

3.1. Eddy Covariance Flux Tower NEE and LE Data to Constrain "Fast" C Cycle Parameters

The earliest ORCHIDEE C cycle parameter optimization studies using a Bayesian DA framework were performed using eddy covariance flux tower NEE data at individual sites (Bacour et al., 2015; Santaren et al., 2007; Verbeeck et al., 2011; Williams et al., 2009). Kuppel et al. (2012) were one of the first to move beyond simple assimilation of C and water flux data at individual sites by simultaneously NEE and LE data from multiple sites ("MS" assimilation). Kuppel et al. (2012) compared the assimilation results from single-site ("SS") assimilations using flux data from different temperate broadleaf flux tower sites to an assimilation in which data from all sites were included in the same MS assimilation. They found that the MS assimilation could find a unique parameter set that resulted in as much of a reduction in RMSE at each of the sites as the equivalent SS assimilation (and better than taking the average of the SS optimized parameter values). This was a crucial step in scaling parameter optimization up to global scales, because ultimately TBMs require one set of parameters per PFT for a global-scale simulation. Kuppel et al. (2014) went on to apply a MS NEE and LE parameter optimization framework to all PFTs for which flux tower data were available from the La Thuile fluxnet database (Table 1). This was the first time that such a MS assimilation had been used to constrain C cycle related parameters across a global scale network of sites. The assimilations resulted in a much-improved NEE (and LE) seasonal cycle for most PFTs, with the exception of tropical broadleaf evergreen trees (TrBE). The TrBE optimizations were likely hampered by the same model structural errors related to root zone water up identified in Verbeeck et al. (2011), whereby deep-root-mediated water access during the dry season, and the resulting sustained productivity, are not captured by that version of the ORCHIDEE model. Numerous other studies continue to use site-based flux tower data to constrain various C cycle related processes in ORCHIDEE, before applying the model to answer a range of global change related questions (Druel et al., 2017; Li et al., 2017; Liu et al., 2018; Peng et al., 2013).

3.2. Satellite Normalized Difference Vegetation Index (NDVI) Data to Constrain Leaf Phenology

MacBean et al. (2015) used satellite NDVI data (Table 1) to constrain the ORCHIDEE phenology parameters for the six deciduous PFTs. NDVI is a measure of vegetation "greenness" and as such can be used to monitor

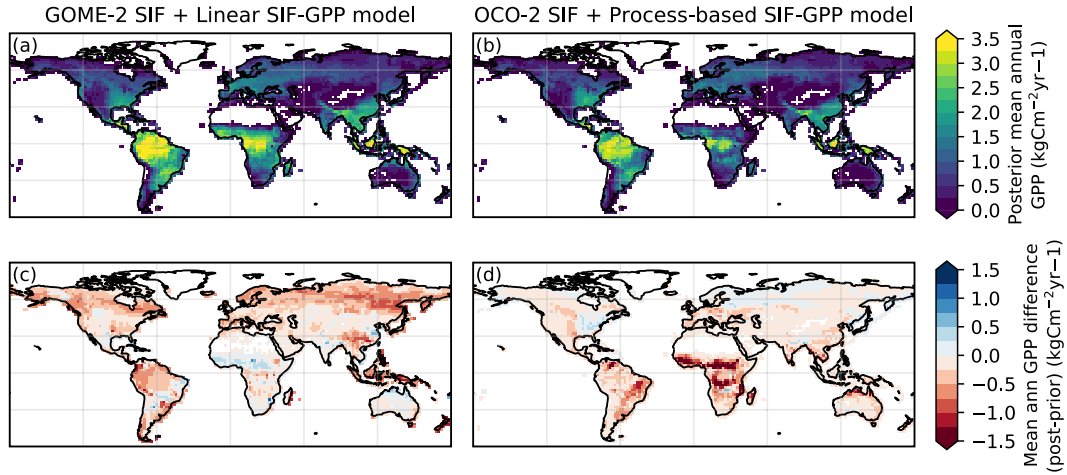

Figure 2. Global maps of mean annual GPP (averaged over 2000–2009) at 2 × 2 resolution: (a) and (b) show the posterior simulation after optimization with GOME-2 SIF over 2007–2011 using the linear SIF-GPP relationship (MacBean et al., 2018), and OCO-2 SIF over 2015–2016 using a physical or process-based model linking SIF and GPP (Bacour, Maignan, MacBean et al., 2019), respectively; (c) and (d) show maps of the difference in mean annual GPP (posterir minus prior simulation) for MacBean et al. (2018) and Bacour, Maignan, MacBean et al. (2019), respectively.

seasonal changes in vegetation dynamics. MacBean et al. (2015) showed that across all four temperate and boreal deciduous PFTs the prior model simulated too long a growing season, that is, leaf senescence started much later than was observed in the data. The assimilations resulted in an earlier start to the senescence (mostly through elevating the senescence temperature threshold) and therefore a much improved fit between the model and the data. However, MacBean et al. (2015) also showed that the prior model did not accurately simulate the LAI dynamics for the semi-arid and tropical deciduous PFTs and neither the SS or MS assimilations could correct for these model errors. These results highlighted a secondary goal of parameter DA, that once the model has been calibrated it is then easier to identify areas of possible structural errors.

3.3. Satellite Solar-Induced Chlorophyll Fluorescence (SIF) Data to Constrain GPP

The advent of satellite measures of solar-induced chlorophyll fluorescence (SIF) in the early 2010s allowed an independent constraint on modeled GPP. Compared to space-borne estimates of FAPAR or vegetation indices (e.g., NDVI) that characterize vegetation dynamics, satellite SIF data are more closely related to photosynthesis activity and hence to GPP. Also, contrary to *in situ* GPP estimates, which are derived from flux tower NEE measurements using a flux partitioning model [with associated uncertainties (Baldocchi et al., 2015)], space-borne SIF estimates provide a *global* scale constraint on GPP. The early SIF studies showed that the satellite SIF data were linearly correlated with GPP at broad spatial and temporal scales, with the exception that the slope of the linear relationship differed between different PFTs (Guanter et al., 2012). MacBean et al. (2018) used these assumptions of linearity to apply a simple linear observation operator for each vegetated PFT to link satellite SIF data (GOME-2; Table 1) and modeled GPP to constrain ORCHIDEE vAR5 photosynthesis and phenology parameters. MacBean et al. (2018) demonstrated that SIF data can provide a strong constraint (considerable decrease in GPP uncertainty) on modeled GPP. Large reductions in the magnitude of GPP were observed for all PFTs except those in semi-arid regions (Figure 2c), resulting in a reduction in the global mean annual GPP budget of ∼30 PgC. yr^{-1} for the 2000–2009 period; thus, removing ORCHIDEE's strong positive bias in simulated C uptake.

The objective of using a simple linear observation operator between GPP and SIF in MacBean et al. (2018) was to test quickly just how beneficial the newly available satellite SIF data could be in constrining model GPP parameters. Following the work of MacBean et al. (2018), which used ORCHIDEE vAR5, Bacour, Maignan, MacBean et al. (2019) developed a mechanistic, or process-based, SIF observation operator in ORCHIDEE v2.0 to better leverage the information content of SIF signals for constraining photosynthesis and phenology processes. The SIF observation operator in Bacour, Maignan, MacBean et al. (2019) relies on two coupled scale-dependent modules: (a) a canopy scale RT modeling approach that relies on an emulator of the RT module embedded in the SCOPE model (van der Tol et al., 2009) to calculate the fluorescence flux emerging at the top of the canopy,

and (b) a new leaf scale model to calculate the regulation of fluorescence related to meteorological forcing and plant stress. The empirical linear relationship approach of MacBean et al. (2018) is limited to coarse spatial and temporal scales where the assumption of a linear relationship between SIF and GPP holds. At finer spatiotemporal scales fluorescence can be decoupled from photosynthetic uptake, particularly in the presence of heat, drought or biotic stress (Magney et al., 2020); therefore, the mechanistic SIF model developed in Bacour, Maignan, MacBean et al. (2019) can be applied at finer spatial scales, with the possibility of constraining model processes using *in situ* SIF data (Parazoo et al., 2020). With this process-based SIF observation operator, Bacour, Maignan, MacBean et al. (2019) used OCO-2 SIF data to constrain ORCHIDEE v2.0 photosynthesis and phenology parameters in addition to two parameters of the SIF model. The resulting optimization also had a considerable impact on the simulated GPP with a decrease of the global mean annual GPP budget by 29 $PgC.yr^{-1}$ over the 2000–2009 period (approximately the same reduction in GPP as MacBean et al., 2018). At regional scale, the correction of GPP was the most pronounced in the southern hemisphere (Figure 2d), and over the Tropics in particular - in contrast to MacBean et al. (2018 - Figure 2c).

Both SIF optimization studies proved that we now have a useful independent constraint on GPP (separate to other proxy measurements of GPP) that can, in theory, be used to help partition the gross from the net CO_2 fluxes (although this is yet to be tested). We have yet to compare the improvement in model-data fit with respect to GPP using the linear versus mechanistic SIF observation operators with the same ORCHIDEE version. We note that the strong decrease in GPP in both studies is model dependent since other global-scale TBM SIF DA studies have found an increase in global mean annual GPP (Norton et al., 2019). In future work we will also test whether the simultaneous optimization of both NDVI and SIF provides independent constraints on LAI and GPP, respectively, and how that might differ for different PFTs.

3.4. Atmospheric CO_2 Data as a Global Scale Constraint on C Fluxes and Stocks

Peylin et al. (2016) combined flux tower NEE and LE data (following Kuppel et al., 2014 - Section 3.1) and satellite measures of vegetation greenness (following MacBean et al., 2015 - Section 3.2) together with a network of atmospheric CO_2 mole fraction data in a global scale step-wise assimilation (Section 2.4). To our knowledge, this study was the first to assimilate three different global-scale data streams to constrain C cycle related parameters of a process-based TBM. The inclusion of atmospheric CO_2 data in the assimilation: (a) corrected the too strong a positive trend in simulated atmospheric CO_2, mainly via small decreases in the regional K_{soilC} values; (b) resulted in a better prediction of the north-south gradient in atmospheric CO_2 predictions; and (c) allowed for further improvement on the amplitude and phase of seasonal cycle in simulated atmospheric CO_2 via adjustment of parameters related mainly (but not only) to heterotrophic respiration. Thus, as expected, atmospheric CO_2 data provided a broad scale constraint on C flux related parameters, and in particular those related to soil C cycling.

Despite the success of the Peylin et al. (2016) study in correcting predictions of trends in atmospheric CO_2 concentration, arguably the more optimal configuration for assimilating multiple datastreams is to include them all in one assimilation - a so-called "simultaneous" assimilation. A step-wise approach to assimilating multiple data streams is equivalent to the simultaneous, but only if you can accurately characterize and propagate the full background and error covariance matrices between each step. To date, this has been difficult to achieve (and is not fully done in Peylin et al., 2016) because we do not know the full structure of the observation error covariance, **R**, matrix that could include model and observation biases, spatiotemporal autocorrelations, and error correlations between different data streams (see Sections 5 and 6.3 for more discussion). The more complex simultaneous assimilation, mixing site and global scale simulations, has been recently implemented by Bacour et al. (2022). Bacour et al. (2022) also found a strong influence of a reduction in the K_{soilC} parameter that decreased the prior predicted trend in atmospheric CO_2 data - thus better matching the atmospheric CO_2 time series.

3.5. Impact for Global and Regional Budgets C Budgets and Comparison With Atmospheric Inversions

What impact did these global-scale parameter optimization studies have on regional to global scale C budget predictions? Neither the global assimilations of flux tower NEE and LE (Kuppel et al., 2014) nor satellite NDVI (MacBean et al., 2015) resulted in much of a change in the mean annual GPP when compared to the prior (default) version of ORCHIDEE vAR5 (Figure 3). For the NDVI assimilation this was perhaps not that surprising given that most of the changes in the temperate and boreal deciduous leaf phenology were related to leaf senescence

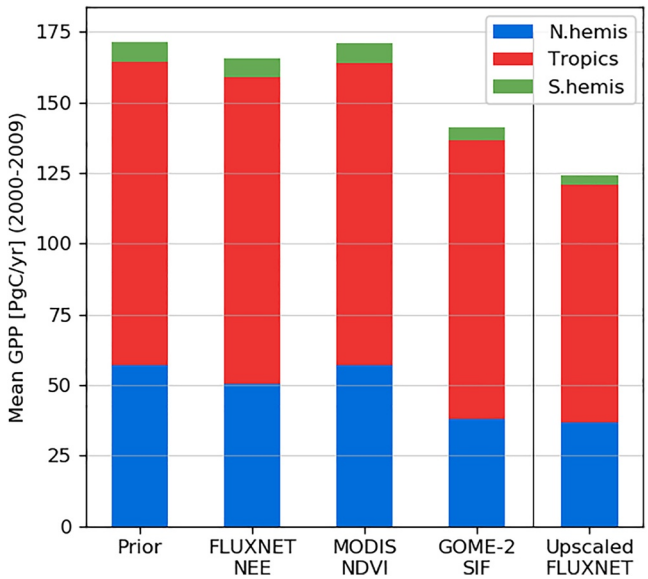

Figure 3. Global annual GPP ($PgCyr^{-1}$) simulated by ORCHIDEE vAR5 averaged over the 2000–2009 period and separated into northern hemisphere (30°N–90°N - blue), tropics (30°S–30°N - red), and southern hemisphere (30°S–90°S - green) regions. The prior model (left bar) is compared with assimilations using FLUXNET NEE to constrain photosynthesis, phenology, respiration, and latent energy flux parameters (Kuppel et al., 2014); MODIS NDVI to constrain phenology parameters (MacBean et al., 2015); GOME-2 SIF data to constrain photosynthesis and phenology parameters, plus parameters of the linear relationship between GPP and SIF (MacBean et al., 2018) and to the right of the solid vertical line an independent data-driven estimate from upscaled FLUXNET flux tower data using the model tree ensemble approach (Jung et al., 2011). ORCHIDEE simulations were run with ERA-Interim climate reanalysis data at LMDz resolution (2.5 × 3.75°). The Bacour, Maignan, MacBean et al. (2019) study is not included in this plot because a later version of ORCHIDEE (v2.0) was used in these assimilations, which hinders the comparison to the other assimilation studies presented in this figure that were performed with vAR5.

parameters. It is likely that during this time GPP is already limited by other climate or soil nutrient factors; hence, an earlier end to the growing season might not impact broader-scale GPP budgets as much as changes to spring onset would have. Therefore, it is worth noting that performing an optimization of leaf phenology using satellite NDVI might have a greater impact in other models that do have biases in spring leaf onset (e.g., Richardson et al., 2012). We did not optimize any phenology parameters for evergreen PFTs because there was no evergreen phenology scheme in ORCHIDEE at the time of these studies. That has since changed (Chen et al., 2020; Peaucelle, Ciais, et al., 2019); therefore, future phenology parameter optimization studies across all PFTs may result in larger changes in simulated C fluxes across both taiga and tropical rainforest regions. Kuppel et al. (2014) noted that part of the improvement in their in situ NEE simulations resulted from compensating effects between GPP and R_{eco}, highlighting the potential limits of solely using NEE to constrain gross C budgets at local to global scales.

Contrary to the flux tower and NDVI assimilations, the global-scale assimilation of satellite SIF data into ORCHIDEE vAR5 (MacBean et al., 2018) did result in a considerable decrease in global mean annual GPP. The greater decrease in the northern hemisphere GPP resulted in a shift in the global GPP distribution, with the post-optimization simulations showing the tropics are a more dominant region of C uptake (Figure 3), a result consistent with Parazoo et al. (2014) who used GOSAT SIF data to constrain the GPP simulated by TBMs participating in the TRENDY model intercomparison. The regional to global scale GPP estimates after the SIF optimization better matched estimates from the upscaled flux tower estimate from (Jung et al., 2011), although we note that this independent data-derived estimate is ultimately the result of a machine learning method with its own uncertainties, and not data in and of itself.

In Figure 4 we compare the prior ORCHIDEE global mean annual NEE budget (2000–2009) to the two posterior simulations (stepwise and simultaneous) that were constrained with atmospheric CO_2 data (Peylin et al., 2016 and Bacour et al. (2022); Section 3.4). In both studies, the improved fit to atmospheric CO_2 trend via a reduction in the K_{soilC} parameter resulted in a much stronger mean global net terrestrial CO_2 sink than the prior (Figure 4 second and third bars) – mostly via an increase in the net CO_2 sink in the tropics for the step-wise assimilation, and both the NH and the tropics for the simultaneous. Consistent with the impact of the NDVI, NEE and SIF assimilations, the assimilations that included atmospheric CO_2 also resulted in a reduction in global mean annual GPP over the 2000–2009 period (possibly in response to a too strong CO_2 fertilization effect in the model - Figure A1a); however, the reduction in R_{eco} was larger than the reduction in GPP (Figure A1b) due to the decrease in the initial total soil carbon via the reduction in K_{soilC}, and thus the overall result was an increase in the mean NEE. Over this short time period, the resultant increase in total above- and belowground C stocks was negligible (almost no change in aboveground biomass in either the NH or the tropics, with a 0.5% increase in total litter and soil C in the NH, and 0.35% increase in the tropics).

We compared the prior and posterior natural terrestrial CO_2 flux simulations (stepwise and simultaneous) to the Global Carbon Project's 2018 Global Carbon Budget's (GCB; Le Quéré et al., 2018) residual terrestrial CO_2 sink estimate (GCB residual S_{LAND} – gray markers) and the mean and standard deviation from the 16 TRENDY model intercomparison v7 DGVM S_{LAND} values (yellow markers) (Figure 4 first three bars). The default ORCHIDEE vAR5 drastically underestimates the prior global terrestrial flux (-0.23 $PgC.yr^{-1}$) compared to the GCB and DGVM mean S_{LAND} values (-2.9 ± 0.9 and -2.7 ± 0.7 $PgC.yr^{-1}$, respectively) (Figure 4 first bar). Constraining the model with MODIS NDVI, FLUXNET NEE, and atmospheric CO_2 data resulted in a natural terrestrial net CO_2 flux for both the stepwise and simultaneous assimilation that was much more comparable with GCB and

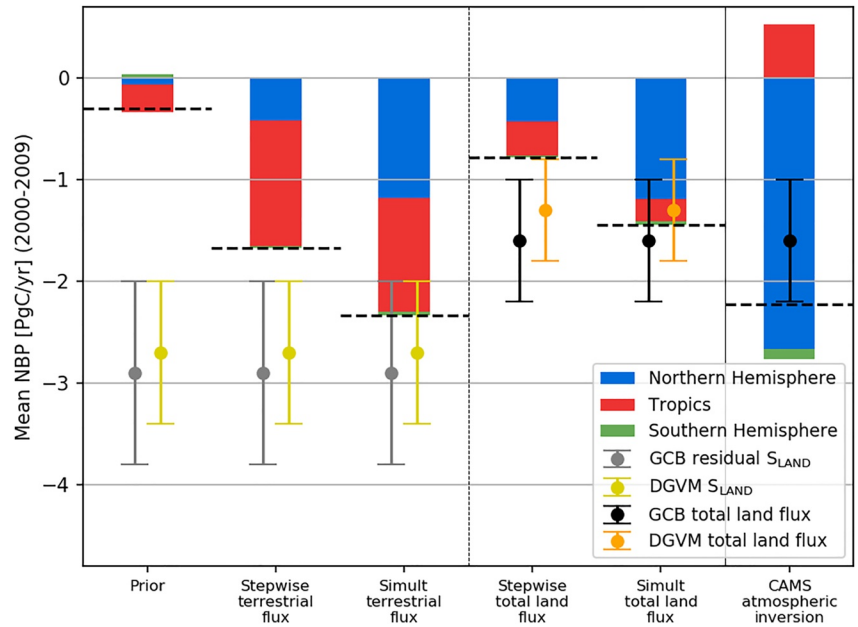

Figure 4. Global annual mean net biome productivity (PgC.yr^{-1}) (averaged over 2000–2009) simulated by ORCHIDEE vAR5 (to the left of the vertical solid line) compared to the CAMS atmospheric inversion (Chevallier et al., 2005) (right of the vertical solid line), which used ORCHIDEE vAR5 surface net CO_2 fluxes as prior information in the inversion (Le Quéré et al., 2018). The prior model (left bar) is compared with the step-wise (Peylin et al., 2016) and simultaneous (Bacour et al. (2022)) assimilation of MODIS NDVI, FLUXNET NEE and LE, and atmospheric CO_2 concentration data (Table 1 and see Sections 3.1 to 3.4). Left of the vertical dashed line are the natural terrestrial fluxes (given that biomass burning emissions and partial forest regrowth were accounted for in the atmospheric CO_2 assimilations – Section 2.4). These simulations are compared to an independent estimate of the terrestrial CO_2 sink (S_{LAND}) over the same 2000–2009 period from the Global Carbon Project's (GCP) 2018 Global Carbon Budget (GCB) (Le Quéré et al., 2018) (S_{LAND} is calculated as the residual of all other budget terms and represented by the gray marker) and to the mean S_{LAND} estimate from the 16 TRENDY v7 DGVMs (yellow marker). S_{LAND} is due to impacts of elevated CO_2, N deposition, and climate change on plant growth and soil C storage and does not take into account net CO_2 fluxes due to LUC. Right of the vertical dashed line shows the same step-wise and simultaneous simulations but with the biomass burning emissions we imposed in the assimilation subtracted, resulting in the total land flux that can be more readily compared to the CAMS inversion. These simulations and the CAMS inversion are compared to the GCB total land flux (i.e., the sum of the terrestrial CO_2 sink plus net LUC emissions - $S_{LAND} + E_{LUC}$ in Le Quéré et al. (2018) - black markers) and for the ORCHIDEE simulations the mean TRENDY v7 DGVM total land flux (orange markers). Error bars represent ±1 s.d. about the mean. The overall global net CO_2 flux for each of the ORCHIDEE simulations and CAMS inversion is shown by the horizontal black dashed lines. The ORCHIDEE simulations and the CAMS inversion are separated into northern hemisphere (30°–90°N - blue), tropics (30°S–30°N - red), and southern hemisphere (30°–90°S - green) regions. ORCHIDEE simulations were run with ERA-Interim climate reanalysis data (Table 1) aggregated to LMDz resolution (2.5 × 3.75). Negative values indicate a sink of C into vegetation and soil; positive values a source of C to the atmosphere.

TRENDY v7 DGVM mean S_{LAND} (Figure 4 second and third bars), although the simultaneous (−2.34 PgC.yr^{-1}) better matched the GCB and DGVM mean values compared to the stepwise (−1.67 PgC.yr^{-1}).

Given both the stepwise and simultaneous simulations shown here included a constraint from atmospheric CO_2 data in the assimilations, we also compared the global mean annual total land flux (2000–2009) from the stepwise and simultaneous (Figure 4 fourth and fifth bars) to the CAMS atmospheric inversion (Chevallier et al., 2005; Le Quéré et al., 2018) (Figure 4 sixth bar). CAMS uses the same prior ORCHIDEE vAR5 simulations as the prior in their inversion; therefore, these ORCHIDEE simulations can be more directly compared with the CAMS inversion. To calculate the ORCHIDEE total land flux we subtracted the net biomass burning emissions (that included partial forest regrowth – i.e., net land use change emissions equivalent to E_{LUC} in Le Quéré et al., 2018) that we had accounted for in the assimilations with atmospheric CO_2 data (see Section 2.4). These estimates are therefore comparable to the GCB and TRENDY v7 DGVM total land flux (which equals $S_{LAND} - E_{LUC}$) GCB and TRENDY v7 DGVM mean and standard deviation estimates of the total land flux are shown in black and orange markers in Figure 4. The stepwise, simultaneous and CAMS show markedly different total land flux values (−0.78, −1.45, and −2.23 PgC.yr^{-1}, respectively). CAMS predicts a much stronger mean net total land

flux (horizontal black dashed lines in Figure 4) over the 2000–2009 period than either the stepwise, but is more comparable to the simultaneous. The simultaneous mean total land flux compares well with both the GCB and TRENDY v7 DGVM values (-1.6 ± 0.6 and -1.3 ± 0.5 PgC.yr^{-1}, respectively), while the stepwise and CAMS estimates show a much weaker and stronger mean total land flux, respectively.

The partitioning between regions is very different when comparing the two optimized ORCHIDEE vAR5 estimates and CAMS. CAMS puts a much stronger sink in the NH and a source for the tropics (similar to other atmospheric inversions described in Peylin et al., 2013 and supported by data from Gatti et al., 2021 and Qin et al., 2021, although see Yin et al., 2020). The simultaneous does place a greater C sink in the NH, thus better matching the CAMS inversion than the stepwise, which shows a roughly equal split in the total land flux between the NH and tropics. However, both ORCHIDEE estimates show a small C sink in the tropics, unlike CAMS. Bacour et al. (2022) demonstrated that posterior ORCHIDEE terrestrial CO_2 sink values vary considerably depending on the assimilation configuration (i.e., number of years and different datasets included in the optimization). One of the key differences that could explain the difference between the stepwise and simultaneous regional to global scale C budget estimates is that only three years of atmospheric CO_2 data were used in the stepwise (compared to 10 years in the simultaneous), which means that the stepwise is likely biased by assimilating such a short period of data. Thus, in the results presented here, constraining a process-based model with the same prior fluxes and the same atmospheric CO_2 data used in the CAMS atmospheric inversion has not resolved the differences in regional partitioning or global total land flux estimates between the "bottom-up" process-based ORCHIDEE versus "top-down" CAMS atmospheric inversion. The optimized ORCHIDEE versions have been confronted against atmospheric CO_2 and therefore can be considered to have accounted for all processes contributing to changes in atmospheric CO_2 concentrations. However, the fact that errors related to missing processes in the model (including land management processes that are not accounted for in the biomass burning emissions used in the atmospheric CO_2 data assimilations) are likely aliased onto posterior parameter estimates during the optimization may mean that ORCHIDEE regional or global budgets are not correctly estimated. Other sources of errors in both the process-based models and inversions may be contributing to the different regional budget estimates, but more research is needed to uncover the causes of discrepancies between these "bottom-up" versus "top-down" approaches. Bastos et al. (2020) and Kondo et al. (2020) suggest possible solutions for reconciling estimates from these two approaches. Nonetheless, the simultaneous prediction of a stronger land C sink in the NH than in the tropics in the simultaneous assimilation (Bacour et al., 2022) is more in line with other atmospheric inversion estimates (Peiro et al., 2022) and contrasts with the roughly equal NH-Tropical land partitioning predicted by many (non-optimized) TBMs (or at least TBM ensemble means; Sitch et al., 2015).

In contrast to the shift in regional to global scale magnitude of the mean annual C flux budgets resulting from assimilating a number of different datastreams, none of the global scale assimilation studies to date have had a dramatic impact on the IAV or medium-term trend in either the gross or net CO_2 fluxes. Figure 5 shows NEE and GPP examples for the NH and tropics, but the same is true for all regions (and for R_{eco}, not shown). We note that this is in contrast to the dramatic impact that assimilating atmospheric CO_2 data has on the simulated trend in atmospheric CO_2 data (Section 3.4 and see Figure 6 in Peylin et al., 2016). This result suggests that the optimizations are not currently altering the model's ability to capture the impacts of inter-annual changes in climate extremes on regional to global scale C budgets. Next steps in addressing this issue will include modifications to the cost function to specifically account for longer timescales (e.g., Desai, 2010) and anomalous extremes will also be considered.

4. Ongoing C Cycle DA Studies With ORCHIDEE Using Novel Datasets and Methods

4.1. Optimization of Effective LAI Using the Two-Stream Radiative Transfer Model Implemented in ORCHIDEE-CAN

LAI is a difficult variable to assimilate in TBMs due to the inconsistent values found across different products applying a range of different RT algorithms; hence,why NDVI was used to constrain leaf phenology parameters in MacBean et al. (2015) and Peylin et al. (2016). With the development of a multi-layered canopy in ORCHIDEE (ORCHIDEE-CAN; Naudts et al., 2015), the 1-D two-stream radiative transfer (RT) scheme of (Pinty et al., 2006) was implemented into the model. Since in a 1-D scheme it is nearly impossible to account for the orientation of the randomly distributed vegetation scatters, this scheme uses effective LAI (eLAI; a statistical description of the vertical distribution of leaf mass that accounts for their orientation and clumping) instead of

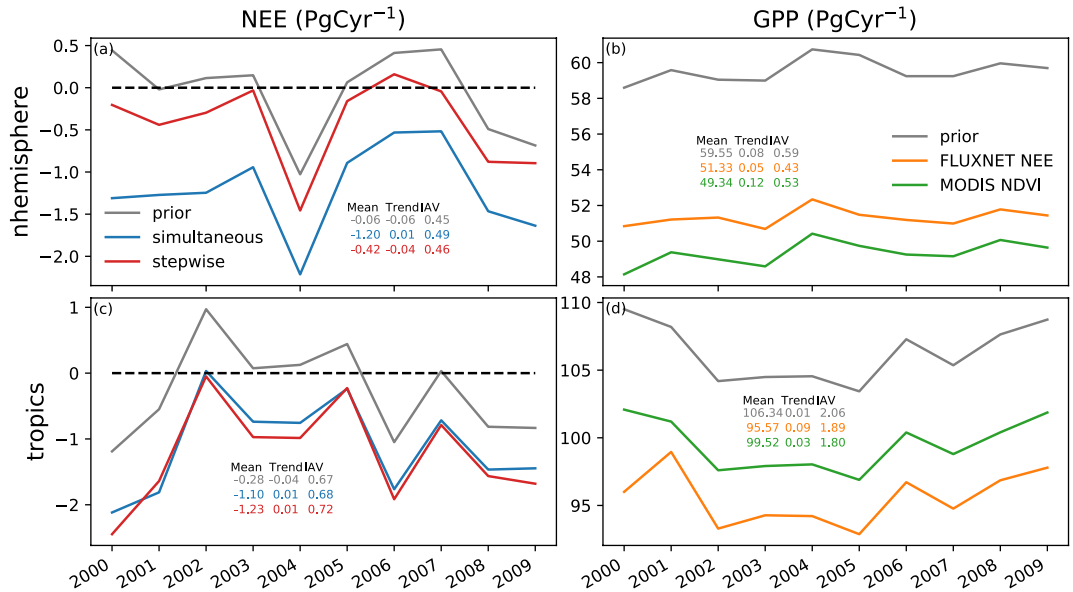

Figure 5. Time series of annual NEE (left column) and GPP (right column) over the 2000–2009 period for the northern hemisphere (top row) and tropics (bottom row) comparing the prior ORCHIDEE simulations (gray curve) with, for NEE: the stepwise (red curve) and simultaneous (blue curve) assimilation of flux tower NEE, satellite NDVI and atmospheric CO_2 data; and for GPP: flux tower data only (orange curve) and satellite NDVI only (green curve). Horizontal grey dashed line shows net zero CO_2 flux.

true LAI, to reproduce the radiative fluxes of a 3-D model. In as yet unpublished work, satellite retrieved eLAI was assimilated from the EU H2020 MULTIPLY Project platform (http://www.multiply-h2020.eu) to constrain the phenology of ORCHIDEE-CAN. As the MULTIPLY eLAI data uses the same assumptions and retrieval algorithms as ORCHIDEE-CAN, we did not have to rely on normalizing the LAI retrievals, as is sometimes done when assimilating true LAI (Chen et al., 2018; Huang et al., 2015). Normalization usually means we lose some of the information content of the measurements and are only able to improve the timing of vegetation dynamics, not the magnitude. By directly assimilating eLAI ("eLAI opt") at a number of in situ broadleaf deciduous and needleleaf evergreen forest flux tower sites, the model fit to observed LAI was improved for all sites (blue crosses all below the 1.0 horizontal dashed line in Figure 6). In contrast, calibrating the model using normalized eLAI ("norm opt") only improved the fit at the deciduous sites (last five sites in Figure 6). This work therefore highlighted the importance of (a) developing the model so it more accurately represents what is observed, and (b) the value of directly simulating measured variables.

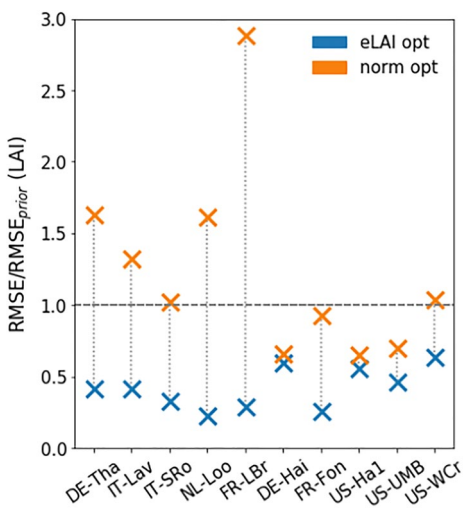

Figure 6. Changes in LAI model-data fit over a number of FLUXNET sites for "eLAI opt", in which effective LAI was directly assimilated, and "norm opt", in which normalized effective LAI data were assimilated. The dashed horizontal line at 1.0 represents the initial model fit at each site, points found below this line show an improvement in fit (measured by RMSE), points above show a degradation in fit. The first five sites are evergreen needleleaf forests and the latter five are deciduous broadleaf sites.

4.2. Can Free Air CO_2 Enrichment (FACE) Data be Used to Optimize Modeled Carbon-Nitrogen Interactions Under Scenarios of Elevated CO_2?

Carbon–nitrogen interactions are essential in understanding global terrestrial ecosystem productivity and the response of plants to rising atmospheric CO_2. Following the inclusion of the nitrogen, N, cycle in ORCHIDEE (Vuichard et al., 2019), members of the ORCHIDEE DA group have started experiments calibrating the new model against FACE elevated CO_2 experimental manipulation data and investigating the difference in the modeled N limitation on the fertilizing effect of CO_2 between the default and optimized model. These DA experiments are useful since future model projections are likely to exceed historical and present-day conditions; therefore, optimizing against

present day observations does not necessarily give us confidence in model projections under a changing climate and rising atmospheric CO_2 (Wieder et al., 2019). The modeled CO_2 fertilization effect on GPP at the two FACE sites included in the assimilation (the Oak Ridge broadleaf deciduous forest - Norby et al., 2010, and the Duke needleleaf evergreen dominated forest site (with deciduous understorey) - McCarthy et al., 2010) was reduced after optimization of data from the elevated CO_2 plot. These results are promising for providing data-constrained model estimates of how C stocks will evolve under elevated CO_2 (e.g., Jiang et al., 2020), which is encouraging given the net impacts of CO_2 fertilization on different components of biogeochemical cycles, carbon-water interactions, and vegetation dynamics is currently a key source of uncertainty in TBMs (Canadell et al., 2021; De Kauwe et al., 2014; Walker et al., 2014, 2021; Zaehle et al., 2014).

4.3. Parameter Optimization as a Tool for Improving CO_2 Flux Simulations in Regions Sensitive to Climate Variability and Carbon-Climate Feedbacks

The possible loss of C stored in arctic soils from permafrost thaw is one of the more uncertain carbon-climate feedbacks (Canadell et al., 2021). TBM groups are continuing to implement the relevant processes needed to improve predictions of the magnitude of this feedback over the coming centuries to millennia (Bowring et al., 2019; Lawrence et al., 2019; Melton et al., 2019). Parameter calibration will be crucial to those efforts (e.g., Schneider von Deimling et al., 2015). In the ORCHIDEE group, several studies have already made headway in calibrating hydrological, soil thermal, and biogeochemical parameters at a number of northern tundra, permafrost, and peatland sites (Dantec-Nédélec et al., 2017; Salmon et al., 2022). Parameter optimizations using soil temperature and moisture data at two arctic forest and tundra sites allowed improvements in vertical soil and heat transfers simulated by a version of ORCHIDEE with the latest hydrology and snow schemes (Dantec-Nédélec et al., 2017). This work presents a crucial step in simulating future permafrost thaw and carbon-water-vegetation interactions in those regions. Salmon et al. (2022) performed the first optimization of CH_4 related parameters with eddy covariance CH_4 fluxes at 14 peatland sites with the ORCHIDEE-PCH4 model. They found generally higher CH_4 emissions across Eurasian and North American peatland regions as a result of using values from the MS assimilation. However, SS assimilations revealed variability in methanogenesis at different sites depending on substrate limitation and seasonal fluctuations in soil temperature. The latter result highlights the need for more research into coupled hydrological-soil thermal-biogeochemical interactions when considering CO_2 and CH_4 flux estimates in high latitude ecosystems, and further model developments (and parameter calibrations) in line with improved process understanding. Ongoing work in this area within the ORCHIDEE group is focused on applying soil drying manipulation experimental data within ORCHIDAS to better estimate the impact of lowered water tables on C stocks and fluxes (Kwon et al., 2022).

The results of semi-arid site phenology optimization from the MacBean et al. (2015) study (Section 3.2), in addition to other ORCHIDEE model evaluation studies showing poor performance of simulated vegetation dynamics in semi-arid regions (MacBean et al., 2021; Traore et al., 2014), motivated the focus on optimization of semi-arid C and water fluxes using a more recent version of ORCHIDEE (v2).2; (Mahmud et al., 2021). They demonstrated using flux tower NEE data from 12 southwestern US semi-arid tree-, shrub-, and grassland-dominated sites that the strong prior model underestimate in mean annual NEE and NEE IAV could be accounted for by optimizing C cycle parameters (contrary to the lack of change in regional C cycle IAV from earlier optimizations - Figure 5) - mostly via constraint on the phenology parameters. In future work we need to expand semi-arid C related parameter optimization to other semi-arid sites worldwide. By repeating global scale simulations with TBMs that have been optimized at a wider range of semi-arid ecosystem sites, we will be able to determine if semi-arid ecosystems are as important in controlling global C cycle IAV as previously suggested (e.g., Ahlström et al., 2015; Poulter et al., 2014).

4.4. Impact of Optimizing Surface Soil Moisture Drydowns on Carbon-Vegetation-Water Interactions

The strong coupling between the C and water cycles has motivated the use of soil moisture (SM) datasets in our recent assimilation studies. By using International SM network data to calibrate SM drydowns simulated by ORCHIDEE following significant rainfall events for a range of different PFTs (Table 1), Raoult et al. (2021) were able to improve the soil drying rates in the model without degrading the fit to other fluxes (e.g., GPP). By better simulating drydowns, it is possible that we improve the simulation of root water uptake and therefore better capture the response of the C fluxes to both drought and to rainfall pulses, which is a particular feature of

semi-arid ecosystems that are thought to dominate global C cycle IAV (Huxman et al., 2004) (see Section 4.3). This study provides an example of the need for more future DA studies coupling both C and hydrology related datasets in the assimilation system (e.g., Moore et al., 2008).

4.5. Accounting for the Ecological Properties of Ecosystems to Better Constrain the Parameter Optimization

Past attempts in optimizing the ORCHIDEE model with multiple sources of data have mainly focused on model performances, with less of a focus on the ecological and physiological consistency of optimized parameters. By optimizing the model parameters against flux tower-derived GPP estimates at a range of different forest and C3 grassland sites, Peaucelle, Bacour et al. (2019) were able to explore the relationships between trait-related parameters as well as their variability with climate conditions for each PFT. They showed that optimized parameter values are consistent with leaf-scale traits and well-known trade-offs observed at the leaf level, echoing the leaf economic spectrum theory (Wright et al., 2004). This exercise demonstrated that it is ecologically sound to use known trade-offs between parameters to better define the **B** matrix (as in Ziehn et al., 2011; Bloom & Williams, 2015; and Pinnington et al., 2016) and thus potentially reduce the risk of model parameter equifinality (where multiple different parameter values result in a similar minimum of the cost function). Accounting for the ecological properties of ecosystems is also relevant given the increasing development of trait-based approaches (Franklin et al., 2020; Scheiter et al., 2013) to represent the acclimation of plant physiological traits and thus improve projections of ecosystem responses and feedbacks to climate change.

5. Technical Challenges Addressed in ORCHIDEE DA Studies

5.1. Challenges of Assimilating Multiple Data Streams

In addition to the main focus on improving global C cycle predictions via parameter optimization, the ORCHIDEE DA team has performed several studies exploring the impact of using a combination of different datastreams in an assimilation, particularly with regard to the impact assimilating only one datastream has on other variables (Bacour et al., 2015; Bacour et al., 2022; Thum et al., 2017). Such studies are important because a common assumption is that assimilating additional independent datastreams should improve the constraint on different parts of the model and therefore result in much improved C budget estimates. However, these mostly site-level studies have highlighted that assimilating one datastream can often degrade the fit to other variables, usually because of a mismatch between what the model represents (including the observation operator) and what the data are actually measuring, incomplete characterization of the R matrix (including both model structural and observation errors), or the choice of parameters is not appropriate. Bacour et al. (2015) demonstrated at two temperate forest sites that assimilating satellite observations related to vegetation dynamics (in this case, FAPAR - Table 1) can degrade the fit to flux tower C flux observations. They found that assimilating both FAPAR and flux tower NEE resolved this issue, resulting in an improvement in RMSE to both datastreams. Bacour et al. (2022) find the same issue, and solution, at global scale when assimilating normalized NDVI and comparing to multiple flux tower site NEE records. Bacour et al. (2015) suggested that the joint assimilation of the two datastreams results in a posterior parameter set that is a compromise from the case where either datastream is assimilated alone, analogous to Kuppel et al. (2012, 2014) who showed that MS assimilations of flux data find a "middle-ground" set of optimized parameter values with a similar, or even better, model fit to the data compared to averaging individual site values. Including more data in the assimilation potentially "smoothes" the cost function and therefore lowers the chance that the algorithm gets trapped in local minima. These studies thus suggest that it is preferable to include more data streams in an assimilation.

Thum et al. (2017) explored the potential benefit from assimilating aboveground biomass data at two different temperate forest sites, in addition to C fluxes. The goal of including biomass data was to constrain "slower" process parameters related to C allocation and biomass pool turnover, as well as those related to the "faster" photosynthesis, respiration, and phenology parameters. Assimilating total biomass data improved the fit to the data, but only by estimating too rapid a biomass turnover. This is because the AR5 version of the model assumed a constant turnover rate due to mortality and did not consider stand scale dynamics of disturbance and management; therefore, optimization against biomass data results in an unrealistically high turnover rate. These results

highlighted that assimilation of certain data is only possible if the model accurately represents all processes that contribute to those measurements, otherwise that model structural error will be "aliased" onto the parameters.

5.2. Challenges Associated With Biased Observations and Model-Data Inconsistencies

Mathematically speaking, many of the issues described in Section 5.1 are related to an incorrect characterization of the **R** matrix - that is, model structural error and/or observation biases are not properly accounted for in the error covariances, resulting in the degradation of model-data fit to non assimilated variables or and "aliasing" of that bias onto the parameters, which results in inaccurate posterior parameter values (Wutzler & Carvalhais, 2014). The same result can occur when observation biases are present. Global scale models are computationally expensive to run; thus, the inversion algorithms chosen must be as efficient as possible. However, the assumptions involved with the algorithms typically used to constrain global scale models (e.g., quasi-linear models and Gaussian error distributions with zero mean) are often violated if biases are not accounted for in the R matrix. The result may be inaccurate posterior parameter values, even when it appears that the parameter uncertainty has been reduced, as demonstrated in synthetic assimilation experiments in MacBean et al. (2016) in which the "true" parameter values of a simple C cycle model were known. It is therefore not enough to simply examine the reduction in parameter uncertainty, or how the reduction in parameter uncertainty propagates to reduced uncertainty on state variables. This issue was further highlighted in Bacour, Maignan, Peylin et al. (2019), who demonstrated that biased real-world SIF data results in a strong bias component to the prior mean squared deviation between the model and data that can be entirely removed via parameter optimization. They therefore argued that data products (in this case, SIF) need to be examined for potential biases, and that those biases should either be removed prior to assimilation or properly accounted for in the cost function. This study also highlighted just how careful modelers should be when choosing a data set to assimilate from a variety of available products.

Several ORCHIDEE DA studies have shown a degradation to the C fluxes when assimilating measures of vegetation dynamics (NDVI or FAPAR) (Sections 4.1 and 5.1). Several hypotheses have been proposed for why satellite measures of vegetation activity are causing a degradation in the fit to C fluxes, including scale mismatches between the model and data resulting in a bias that we are not accounting for in the R matrix (Section 6.3), or possible large discrepancies in how vegetation canopies are represented in the model versus reality. In ORCHIDEE vAR5, vegetation canopies are approximated using the so-called "big leaf" approach, in which the canopy for each PFT is represented by a single uniform green column with PFT-dependent maximum LAI, and with a simple beer-lambert law describing attenuation of light through the canopy. No vertical canopy structure nor spatial heterogeneity within the grid cell is represented. However, even the use of effective LAI within the two-stream RT model of ORCHIDEE-CAN (Section 4.1) degraded the fit to flux tower GPP. Therefore, use of effective LAI and a more complex RT model did not resolve this issue. Furthermore, unpublished work investigating the impact on C fluxes when satellite NDVI is assimilated at grassland sites results in the same degradation; therefore, inaccurate representation of forest canopies and RT schemes cannot be the only issue. This is an unsolved issue and we continue to investigate. One possible explanation is that LAI (and FAPAR) and GPP may be too tightly coupled in the model, when in reality they may not be as well correlated for some PFTs and/or at certain times of the year. This has been observed in several different biomes when comparing satellite-derived SIF and NDVI/FAPAR with in situ GPP, with SIF tracking photosynthetic activity much better than NDVI/FAPAR (Joiner et al., 2014; Smith et al., 2018; Walther et al., 2016). This may point to a model structural issue, or, it may simply point to the need to optimize phenology parameters in a first step using satellite NDVI/FAPAR, followed by parameters related to photosynthesis and the use of carbohydrate reserves for plant growth using GPP/SIF in a second step (as suggested by Kolassa et al. (2020) following their use of satellite FAPAR to calibrate vegetation parameters).

5.3. Technical C Cycle DA Studies

A third focus in past ORCHIDEE DA studies has been to explore different DA system configurations for parameter optimization. Kuppel et al. (2013) moved beyond previous ORCHIDEE DA studies that simply define the **R** matrix as the prior RMSE between model and data [or by simply inflating the observation uncertainties by the number of observations as in (Kuppel et al., 2012; Thum et al., 2017)] by exploring a method for quantifying the model structural error and its impact on C cycle parameter optimization with NEE based on simple model-minus-observation mismatch statistics (Desroziers et al., 2005). They found that this error has a standard

deviation of 1.5–1.7 gC m^{-2} day^{-1} in the NEE space, without a significant temporal correlation beyond the lag of a few days but with a large spatial persistence that can be approximated with an exponential decay of e-folding length of 500 km. The same Desroziers et al. (2005) consistency checks were used by Bacour et al. (2022) to determine if the prior parameter errors (set to 40% of the range of variation for each parameter) were over- or underestimated. Their experiments suggest that the 40% is likely an overestimate (i.e., the prior parameter errors are too high), although it is highly dependent upon the assimilation set-up (data used and associated observation errors).

Two papers by Santaren et al. (2014) and Bastrikov et al. (2018) examined the advantages and disadvantages of using a gradient descent (L-BFGS-B) versus a "global search" (GA) inversion method for parameter estimation with flux tower data (see Section 2.2). Global search algorithms have the advantage of fewer assumptions compared to gradient descent methods, which may be violated in the case of using complex TBMs and/or biased observations (see Section 5.2). In a SS assimilation, both studies found that the GA was better able to find the minimum of the cost function because, unlike L-BFGS-B, the GA algorithm did not get caught in so-called "local minima". Bastrikov et al. (2018) expanded the investigation to multi-site (MS) optimizations across a wider range of PFTs and found that in the MS case the difference between the two methods in terms of reducing the cost function was smaller. They suggest that this smaller difference between the two methods is because adding more data into the assimilation acts to "smooth" the cost function - thus, making it easier for a gradient-based method not to get trapped in local minima. Nonetheless, Bastrikov et al. (2018) demonstrate that the GA is still the superior method for adequately exploring the parameter space and for finding the "true" parameters (as evidenced by performing synthetic experiments). The caveat is that it is currently too computationally expensive to run the GA using a greater number of data points than is typically used in a MS flux tower assimilation. If we add in a greater number of sites, global gridded datasets, or a greater number of different datasets, the time it would take to perform just one assimilation with the GA would be prohibitively expensive. More recent ORCHIDEE DA studies using L-BFGS-B have relied on the use of multiple initial parameter guesses in order to improve the convergence efficiency to the global minimum of the cost function (Bacour, Maignan, MacBean et al., 2019; MacBean et al., 2015; Peaucelle, Bacour, et al., 2019). As the computational efficiency of running the ORCHIDEE model and the optimization code continues to increase, it may be feasible to run all our optimizations using the GA (or another global search algorithm such as Metropolis Hastings Markov Chain Monte Carlo) in the near future.

6. Perspectives and Roadmap for Future C Cycle DA Studies

6.1. The Challenge of Initializing Regional to Global-Scale Aboveground Biomass and Soil C Stocks

Peylin et al. (2016) and Bacour et al. (2022) demonstrated the need to have accurate initial estimates of soil C stocks in order to correctly predict trends in atmospheric CO_2. If ESMs cannot accurately simulate current trends in atmospheric CO_2 then we cannot have confidence in our predictions of whether the land will remain a sink of C, nor in the allowable C emissions for remaining under a given threshold rise in temperature. The widely used approach for model spin-up results in biased estimates of soil C (Exbrayat et al., 2014; Schwalm et al., 2019). It is likely that our current approach of spinning up the C stocks to equilibrium (i.e., the steady state assumption) prior to the industrial revolution (~1,750) is flawed because humans have been actively managing the land for thousands of years before this point (Ellis et al., 2021) and we do not have enough information on the land management and land use history for every point on the land surface (Pongratz et al., 2018). Indeed, model-data fit within a DA system has been shown to improve when the steady state assumption is relaxed (Carvalhais et al., 2008, 2010). Therefore, we believe that one key objective of near-term C cycle DA efforts should be on accurate initialization of model biomass and C cycle stocks (Luo et al., 2016) either after spin up at the beginning of the industrial revolution (although data on this time period are scarce, but see Matthes et al., 2016) or after the transient run and before the historical simulation at the beginning of the assimilation window as in Peylin et al. (2016) and Carvalhais et al. (2008, 2010). This can be achieved by including scalars on initial C pools as parameters in the optimization (e.g., K_{soilC} in Peylin et al., 2016, and possibly an equivalent K_{biomass} scalar for the aboveground biomass stocks). Another option would be to include the spinup period in the parameter optimization; however, the computational expense of including the spinup within a regional to global scale assimilation prohibits this as a realistic option. An alternative approach might be to use state DA with gridded estimates of changes in C stocks - even satellite estimates of aboveground biomass can propagate through to adjustments in

belowground C stocks (e.g., Fox et al., 2018). This would also have the benefit of helping to keep the models up-to-date with how C stocks are responding to recent land use and cover changes. A third option would be to focus data-collection efforts on accurately estimating global gridded soil C stocks, which can be used to initialize the model (e.g., the International Soil Carbon Network, ISCN; Harden et al., 2018; Nave et al., 2016). Parameter optimization could then be performed after accurate soil C stocks values have been prescribed and state DA approaches could then be used to adjust aboveground biomass with continuing land use and cover changes. It is worth noting that while we can optimize initial C stocks at site level (e.g., using a K_{soilC} parameter), for global scale simulations it is likely necessary to constrain the soil C stocks at regional to global scales using global-scale data (e.g., global gridded data or global atmospheric CO_2 concentrations) rather than networks of *in situ* data. We have yet to make full use of global gridded datasets in ORCHIDEE C cycle DA studies. This is a clear goal for the future.

6.2. Novel Datasets Needed for Constraining Predictions of Carbon-Climate Feedbacks

Despite the fact that Peylin et al. (2016) were one of the first to use three separate C cycle related datastreams to constrain global TBM C flux estimates, these three sources of data represent just a fraction of the data that are now available for C cycle model optimization. We must utilize the data from the increasing number of both *in situ* eddy covariance CO_2 flux sites and tall tower atmospheric CO_2 mole fraction measurement stations, as well as longer running records at many existing sites. In Sections 3 and 4, we discussed some of the newer datasets that we have started to use to constrain different variables in the ORCHIDEE model, including satellite-derived SIF products, effective LAI, and elevated CO_2 manipulation data. Other ongoing studies within the ORCHIDEE DA group aim to use tree ring width data to optimize C allocation, and carbonyl sulphide (COS) measurements to constrain GPP. Initial work has focused on developing models - or observation operators - linking the relevant processes already implemented in ORCHIDEE to the new data. Jeong et al. (2020) have proposed a novel approach for linking ORCHIDEE-CAN to tree ring width estimates from the International Tree Ring Data Bank that accounts for sampling biases in tree ring chronologies, while Barichivich et al. (2021) used C and oxygen isotopes derived from tree rings as an integrative benchmark of simulated tree growth and physiological responses to drought. Follow up work will use these data formerly within the DA system to constrain relevant C uptake and allocation and stomatal conductance related parameters. A process-based model for vegetation COS uptake has been developed for ORCHIDEE (Maignan et al., 2021). To be able to assimilate the COS flux measured at the ecosystem scale, we have also modeled the soil COS contribution (Abadie et al., in prepartion). The next step will be to assimilate COS observed fluxes to constrain the simulated GPP. A complementary approach would be to exploit COS atmospheric concentration data in a similar way to atmospheric CO_2 data - that is, by coupling ORCHIDEE to the LMDz transport model. Remaud et al. (2021) have already used atmospheric CO_2 and COS concentration data in an atmospheric inversion of LMDz. Analogous to using both NEE and GPP, these two sources of constraint on both net and gross surface CO_2 fluxes resulted in a 2 $PgC.yr^{-1}$ increase in the NH net C sink.

As we noted above, better understanding and predictions of carbon-climate feedbacks will rely on accurate estimation of longer-term C cycling and changes in above- and belowground C stocks; therefore, we must find new methods and data to address this issue. The global soil respiration database (Jian et al., 2021) and ISCN should be more widely used in TBM C Cycle DA. Soil radiocarbon measurements (Lawrence et al., 2020) are also a promising source of information for constraining rates of soil C cycling (Shi et al., 2020). With the increasing development of models that represent the vertical and spatial heterogeneity of forest canopies (e.g., Fisher et al., 2018; Naudts et al., 2015), new remote sensing datasets related to forest canopy biomass and structure (e.g., the full waveform lidar data of vegetation height from the GEDI mission; Dubayah et al., 2020) or aboveground biomass from the ESA BIOMASS mission (Quegan et al., 2019), methods for detecting crown fractional coverage (Brandt et al., 2020), and novel machine learning approaches linking ground-based forest inventory data with remote sensing of vegetation height (e.g., Xu et al., 2021) will be instrumental in aiding initialization and parameterization of the new vegetation demographic models. Scholze et al. (2017) and Exbrayat et al. (2019) provide useful reviews of how satellite data can be utilized in C cycle DA systems. These satellite datasets will be particularly useful for high carbon regions such as the Arctic boreal zone or tropical rainforests that are particularly sensitive to carbon-climate feedbacks yet have a relative paucity of *in situ* field stations. Similarly, with recent developments of mechanistic plant hydraulic schemes in TBMs (Kennedy et al., 2019; Li et al., 2021; Naudts et al., 2015), simultaneous assimilation of C and water stocks and fluxes will

be needed to constrain these complex, interacting processes. This should help to improve predictions of deep root water uptake and therefore ecosystem vulnerability or resilience to water stress (Fan et al., 2017; Kleidon & Heimann, 1997). Microwave satellite derived estimates of vegetation optical depth data (e.g., Konings et al., 2016; Liu et al., 2011; Moesinger et al., 2020) may also be useful in this regard, particularly for separating out leaf growth versus plant water responses to water stress (Feldman et al., 2021; Konings, Rao et al., 2019). Finally, we recommend that modeling groups rapidly increase their use of experimental manipulation data in parameter optimization studies, as we have documented in a pilot study described in Section 4.2 (and see Medlyn et al., 2015). Ultimately, accurate predictions of carbon-climate feedbacks will rely on models being able to accurately simulate how ecosystem processes respond to elevated CO_2 and changing climate; therefore, it is imperative that we make full use of the experimental data that can give us insights into these responses. As with all past C cycle DA studies, it will take time to test the best approaches for how to best use novel data-streams within a DA system.

6.3. Challenge of Accurate Characterization of the Prior Observation and Parameter Error Covariance Matrices, R and B

As discussed in Sections 5.1 to 5.3, accurate characterization of the **R** and B matrices is difficult to achieve; however, this is a key goal of the ORCHIDEE DA group, as it is for many other groups in the land modeling DA community. Most land model DA studies to date have only prescribed the diagonal elements of R - meaning that we assume that observation (including model structural errors) errors are independent of each other. This is not the case in reality: correlated errors can exist in individual datasets due to spatial and temporal autocorrelation, as well as between two or more different datasets of related variables. The magnitude of the model structural component of **R** is also typically poorly constrained in TBM C cycle DA. As discussed in Section 5.3, Kuppel et al. (2013) presented a novel approach adapted from the atmospheric DA community for estimating the off-diagonal elements of the **R** matrix, but more "DA science" research needs to be conducted in this area. As we argued in Section 5.2, data providers and users need to identify observation biases and/or inconsistencies between the model and the data and incorporate that information into the **R** matrix.

In a similar vein, we need to increase our use of the types of consistency tests employed in Bacour et al. (2022) for defining the variance terms of the B matrix (Section 5.3). In another study with the JULES TBM, Raoult et al. (2016) performed similar tests for determining the prior parameter errors. They included a multiplicative factor on the prior parameter ("background") error term in J(x), which, in the absence of more information on the prior errors, was tuned to ensure a normal distribution with sufficiently high variance to not be too constrained to the prior value, nor too broad that the parameters hit the bounds of their ranges. This ensured the inversion problem was well posed and well conditioned. Using known ecological relationships between parameters (i.e., error correlations between parameters - see Section 4.5) is also critical for a better characterization of the off-diagonal B matrix terms. We also need more information (e.g., field measurements) on parameter ranges, which necessitates that model parameters are measurable. Without this extra information and better characterization of B, the optimization likely will be ill posed, resulting in model equifinality. Synthetic experiments with pseudo-observations and known parameter values should help to identify the types issues with characterizing R and B presented here, and thus should form one of the first steps in a rigorous DA framework.

6.4. Further C Cycle DA Methods and Inter-Comparison Studies Needed

Throughout this review we have highlighted that the posterior parameter values and resultant estimates of C cycle budgets can vary considerably depending on the specific DA system configuration used (e.g., the data type included, stepwise vs. simultaneous assimilation of multiple datastreams, the record length and frequency of the observations, the number of sites, single vs. multi-site optimization, which PFTs are optimized, the number parameters included and to which variables they are sensitive, prior uncertainties on parameters and assimilated quantities, error correlations, and which terms/metrics are included in cost function, among other factors). As a result of this issue and of ongoing ORCHIDEE developments, relatively few parameters derived from our C cycle DA studies have been incorporated in the ORCHIDEE trunk version. In the earliest days of terrestrial ecosys-

tem model C cycle DA studies different aspects of the DA system configuration were explored in more detail (Braswell et al., 2005; Keenan et al., 2013; Moore et al., 2008; Ricciuto et al., 2008, 2011; Santaren et al., 2007; Zobitz et al., 2014). We have continued this research direction (see Section 5) but we argue that much more technical DA research is needed, including:

1. More studies to determine how the DA experiment configuration is impacting our predictions of all aspects of regional to global scale C budgets, IAV and future trends.
2. Extensive assessments by all modeling groups of the impact of different sources of uncertainty in their models, in order to better target model development and optimization efforts.
3. To fully characterize the contributions of different sources of uncertainty we should adopt ensemble approaches that can not only explore the full parameter space, but also include different driving datasets and versions of the model that account for different representations of physical processes.
4. Development of more consistent and robust C cycle DA frameworks, which would include initial steps of global sensitivity analyses (Morris, 1991), synthetic DA experiments (Sections 5.2, 5.3, 6.3), and a wider range of consistency checks and diagnostic metrics (e.g., Desroziers et al., 2005) to assess DA system performance. This type of framework should be used prior to assimilating "real" data so we can have more confidence that the DA system set-up we are using can reliably and accurately estimate and reduce uncertainty on C cycle parameters and processes.
5. These DA frameworks should be easily adaptable and robust to changes in the underlying model version. Keeping the DA system up-to-date with TBM developments is not an easy task, particularly when you have to derive the TL or adjoint version of the code. In the ORCHIDEE DA team a tool has been developed to do the required preprocessing of any version of ORCHIDEE so the TL version of the model can be easily derived using the automatic differentiation tool TAF (Transformation of Algorithms in Fortran; Giering et al., 2005). However, we need to test other approaches such as the 4D ensemble variational (4DEnVar) DA method (e.g., Pinnington et al., 2020) that may be helpful in this regard. In the case of 4DEnVar a TL or adjoint version is not needed and the computational time required for the assimilation is much lower. Such approaches can therefore accommodate increasing complexity in TBMs, which usually require longer simulation times.
6. Increasing model complexity does not always increase predictive power (Famiglietti et al., 2021). Therefore, model sensitivity and DA studies should be expanded to examine the relative benefit of increasing model complexity (and realism) versus the constraint on model structural and parametric uncertainty that different datasets can achieve (Famiglietti et al., 2021; Bacour et al., 2022). Indeed, improved predictive power with more complex model developments may only be realized once observations have been used within a formal DA framework to constrain model parametric uncertainty (Famiglietti et al., 2021). Model development would also benefit from applying DA approaches in a hypothesis testing approach, specifically for testing different assumptions about how various processes can be represented in models (Mahmud et al., 2018; Walker et al., 2018).
7. More calibrated MIPs (i.e., global DA system intercomparison studies) are needed to determine if the spread between models seen in so many model benchmarking studies (see Introduction) is reduced once parameter uncertainty has been accounted for. Such studies would present a clear signal to the broader CMIP modeling community that accounting for parameter uncertainty within a rigorous, formalized DA system is a much needed and worthwhile endeavor.
8. Finally, we need to explore: (a) which parameters (and therefore, which processes) are mostly responsible for the changes we see in the optimized global-scale simulations; (b) whether the data-constrained model can capture observed changes in the land C sink at finer regional scales within a given biome (e.g., Hubau et al., 2020); and, (c) what the impact of the optimizations will be in future projections of carbon-climate feedbacks.

7. Conclusions

Our global C cycle DA studies have shown that optimizing the C cycle parameters of a model can dramatically change in which region the model estimates the strongest C sink; therefore, it is imperative that modeling studies aiming to address this question use a formal DA system to quantify and reduce parametric (and ideally, structural) uncertainty. The studies that the ORCHIDEE DA group have performed to date have often

been pioneering studies in the C cycle DA field. However, much more technical "DA science" work needs to be done in testing different TBM C cycle DA configurations (see Section 6.4) before we can reliably use data-constrained TBMs to answer questions such as which regions dominate the global land C sink and IAV, which C cycle processes and which C pools are responsible for the net uptake of CO_2, and the resilience of ecosystem functioning in response to future climate change under possible confounding effects of elevated CO_2 (Gampe et al., 2021; Yuan et al., 2019). In particular, more attention is needed to address the C cycle DA challenges related to spin-up and accurate initial C stocks, and to identify observation biases and inconsistencies between what is represented in the data and what is observed in reality so that we can better characterize the prior observation error covariance matrix. DA science studies need to continue in coordination with the process-based model developments still needed to account for the complex interactions between the terrestrial biosphere, climate change, rising CO_2, and land use change and management. DA allows us to make optimal use of information from both models and data; thus, it is a crucial tool for improving predictions of carbon-climate feedbacks. We argue that the TBM community should move rapidly toward development of DA systems, and the preferential use of data-constrained simulations in the annual global carbon budget and IPCC/CMIP future climate change projections.

Appendix A: Partitioning of Net CO_2 Fluxes Into GPP and R_{eco} for Assimilations Including Atmospheric CO_2 Concentration Data

Figure A1 shows the separation of regional to global net ecosystem exchange (NEE) shown in Figure 4 into the two component gross C fluxes: gross CO2 uptake (GPP) and ecosystem respiration (R_eco).

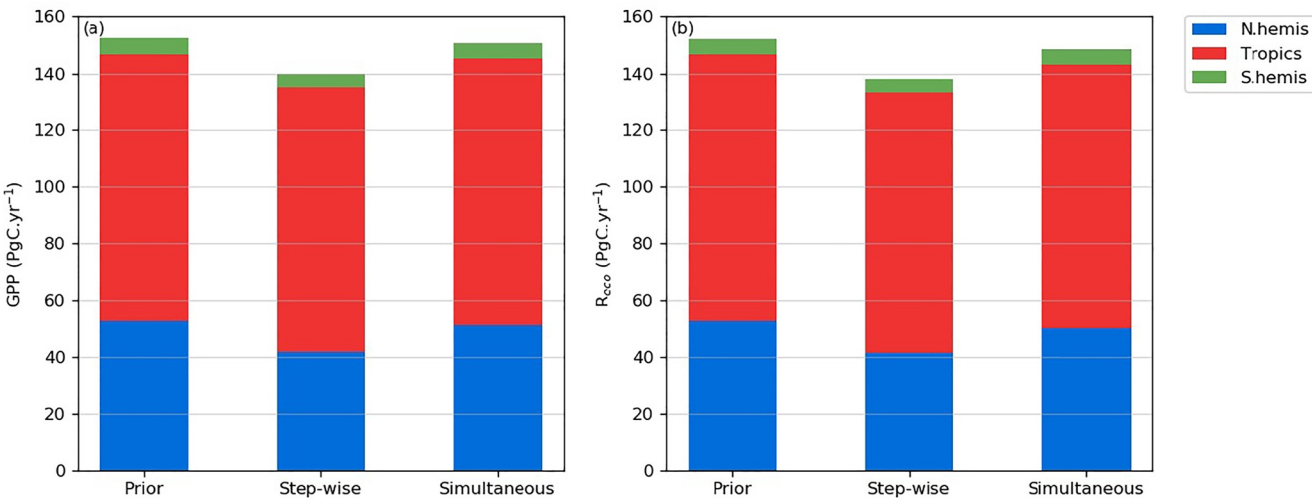

Figure A1. Global annual (a) GPP and (b) R_{eco} (PgCyr-1) simulated by ORCHIDEE vAR5 averaged over the 2000–2009 period and separated into northern hemisphere (30°N–90°N - blue), tropics (30°S–30°N - red), and southern hemisphere (30°S–90°S - green) regions.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The ORCHIDEE vAR5 (and later versions) model code and documentation are publicly available via the ORCHIDEE wiki page (http://forge.ipsl.jussieu.fr/orchidee/browser/) under the CeCILL license (http://www.cecill.info/index.en.html, CeCILL, 2020). The associated ORCHIEE documentation can be found at https://forge.ipsl.jussieu.fr/orchidee/wiki/Documentation. The ORCHIDEE model code is written in Fortran90 and is

maintained and developed under an SVN version control system at the Institute Pierre Simon Laplace (IPSL) in France. The ORCHIDAS data assimilation scheme (in Python) is available through a dedicated web site (https://orchidas.lsce.ipsl.fr). Simulation post-processing and plotting scripts for the figures presented in this paper were performed in Python and are provided on NM's GitHub repository: https://github.com/nmacbean/ORCHIDEE-Param-DA-Review-GBC/ https://doi.org/10.5281/zenodo.6621632, MacBean et al., 2022.

Acknowledgments
We would like to thank the ORCHIDEE Project Team for developing and maintaining the ORCHIDEE code and for providing the ORCHIDEE versions used in this study. We also thank all the data providers that have made this work possible, including the FLUXNET Site PIs and data management team, developers of the NASA MODIS NDVI products, and the GOME-2 and OCO-2 SIF products, all of which are cited in papers referenced in this study. Most of the studies and results presented here were supported by the EU's seventh Framework FP7-SPACE-2009-1 Program for Research and Development "CARBONES" Project under grant agreement number 242316, the EU Horizon 2020 H2020-EO-2015 Program "MULTIPLY" Project under grant agreement number 687320, the EU Horizon 2020 H2020-SC5-2016-2017 Program "VERIFY" Project under grant agreement number 776810, and the French CNES-TOSCA (Centre Nationale d'Etudes Spatiale – Terre, Océan, Surfaces continentales, Atmosphère) Programme-supported FluOR and ECOFLUO Projects. NR was partially supported by the European Space Agency (ESA) as part of an ESA Climate Change Initiative (CCI) Fellowship (ESA ESRIN/Contract No. 4000133601) (project DESTRESS).

References

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., et al. (2015). Carbon cycle. The dominant role of semi-arid ecosystems in the trend and variability of the land CO_2 sink. *Science*, *348*(6237), 895–899.

Albergel, C., Dutra, E., Bonan, B., Zheng, Y., Munier, S., Balsamo, G., et al. (2019). Monitoring and forecasting the impact of the 2018 summer heatwave on vegetation. *Remote Sensing*, *11*(5), 520.

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.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., et al. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. *Reviews of Geophysics*, *53*(3), 785–818. https://doi.org/10.1002/2015rg000483

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., et al. (2013). Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth system models. *Journal of Climate*, *26*(18), 6801–6843. https://doi.org/10.1175/jcli-d-12-00417.1

Anav, A., Murray-Tortarolo, G., Friedlingstein, P., Sitch, S., Piao, S., & Zhu, Z. (2013). Evaluation of land surface models in reproducing satellite derived leaf area index over the high-latitude northern hemisphere. Part II: Earth system models. *Remote Sensing*, *5*(8), 3637–3661. https://doi.org/10.3390/rs5083637

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., et al. (2020). Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences*, *17*(16), 4173–4222. https://doi.org/10.5194/bg-17-4173-2020

Bacour, C., Baret, F., Béal, D., Weiss, M., & Pavageau, K. (2006). Neural network estimation of LAI, fAPAR, fCover and LAI×Cab, from top of canopy MERIS reflectance data: Principles and validation. *Remote Sensing of Environment*, *105*(4), 313–325. https://doi.org/10.1016/j.rse.2006.07.014

Bacour, C., MacBean, N.,Chevallier, F., Léonard, S., Koffi, E. N., & Peylin, P. (2022). *Assimilation of multiple different datasets results in large differences in regional to global-scale NEE and GPP budgets simulated by a terrestrial biosphere model*. Submitted.

Bacour, C., Maignan, F., MacBean, N., Porcar-Castell, A., Flexas, J., Frankenberg, C., et al. (2019). Improving estimates of gross primary productivity by assimilating solar-induced fluorescence satellite retrievals in a terrestrial biosphere model using a process-based SIF model. *Journal of Geophysical Research: Biogeosciences*, *124*(11), 3281–3306. https://doi.org/10.1029/2019jg005040

Bacour, C., Maignan, F., Peylin, P., Macbean, N., Bastrikov, V., Joiner, J., et al. (2019). Differences between OCO-2 and GOME-2 SIF products from a model-data fusion perspective. *Journal of Geophysical Research: Biogeosciences*, *124*(10), 3143–3157.

Bacour, C., Peylin, P., MacBean, N., Rayner, P. J., Delage, F., Chevallier, F., et al. (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

Baldocchi, D., Sturtevant, C., & Contributors, F. (2015). Does day and night sampling reduce spurious correlation between canopy photosynthesis and ecosystem respiration? *Agricultural and Forest Meteorology*, *207*, 117–126.

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In *Progress in photosynthesis research* (pp. 221–224). Springer.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., et al. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. *Remote Sensing of Environment*, *110*(3), 275–286.

Barichivich, J., Peylin, P., Launois, T., Daux, V., Risi, C., Jeong, J., & Luyssaert, S. (2021). A triple tree-ring constraint for tree growth and physiology in a global land surface model. *Biogeosciences*, *18*(12), 3781–3803.

Bastos, A., O'Sullivan, M., Ciais, P., Makowski, D., Sitch, S., Friedlingstein, P., et al. (2020). Sources of uncertainty in regional and global terrestrial CO_2 exchange estimates. *Global Biogeochemical Cycles*, *34*(2). https://doi.org/10.1029/2019gb006393

Bastrikov, V., MacBean, N., Bacour, C., Santaren, D., Kuppel, S., & Peylin, P. (2018). Land surface model parameter optimisation using in situ flux data: Comparison of gradient-based versus random search algorithms (a case study using ORCHIDEE v1.9.5.2). *Geoscientific Model Development*, *11*(12), 4739–4754.

Bloom, A. A., Exbrayat, J.-F., van der Velde, I. R., Feng, L., & Williams, M. (2016). The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(5), 1285–1290.

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

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., et al. (2021). Advances in land surface modelling. *Current Climate Change Reports*, *7*(2), 45–71.

Bonan, B., Albergel, C., Zheng, Y., Barbu, A. L., Fairbairn, D., Munier, S., & Calvet, J. C. (2020). An ensemble square root filter for the joint assimilation of surface soil moisture and leaf area index within the land data assimilation system LDAS-monde: Application over the euro-mediterranean region. *Hydrology and Earth System Sciences*, *24*(1), 325–347.

Bonan, G. B., & Doney, S. C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. *Science*, *359*(6375), eaam8328.

Bowring, S. P., Lauerwald, R., Guenet, B., Zhu, D., Guimberteau, M., Tootchi, A., et al. (2019). ORCHIDEE MICT-LEAK (r5459), a global model for the production, transport, and transformation of dissolved organic carbon from Arctic permafrost regions–Part 1: Rationale, model description, and simulation protocol. *Geoscientific Model Development*, *12*(8), 3503–3521.

Brandt, M., Tucker, C. J., Kariryaa, A., Rasmussen, K., Abel, C., Small, J., et al. (2020). An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature*, *587*(7832), 78–82.

Braswell, B. H., Sacks, W. J., Linder, E., & Schimel, D. S. (2005). Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. *Global Change Biology*, *11*(2), 335–355. https://doi.org/10.1111/j.1365-2486.2005.00897.x

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, *16*(5), 1190–1208. https://doi.org/10.1137/0916069

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., da Cunha, L. C., Cox, P. M., Alexey, V., et al. (2021). Global carbon and other biogeochemical cycles and feedbacks. In P., Zhai, A., Pirani, S. L., Connors, C., Péan, S., Berger, & N., Caud (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change [Masson-Delmotte, V]*. Cambridge University Press. In Press.

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., et al. (2010). Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints. *Global Change Biology*, *16*(10), 2813–2829.

Carvalhais, N., Reichstein, M., Seixas, J., Collatz, G. J., Pereira, J. S., Berbigier, P., et al. (2008). Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval. *Global Biogeochemical Cycles*, *22*(2). https://doi.org/10.1029/2007gb003033

CeCILL. (2020). Retrieved from https://cecill.info/index.en.html

Chen, X., Maignan, F., Viovy, N., Bastos, A., Goll, D., Wu, J., et al. (2020). Novel representation of leaf phenology improves simulation of Amazonian evergreen forest photosynthesis in a land surface model. *Journal of Advances in Modeling Earth Systems*, *12*(1), e2018MS001565.

Chen, Y., Zhang, Z., & Tao, F. (2018). Improving regional winter wheat yield estimation through assimilation of phenology and leaf area index from remote sensing data. *European Journal of Agronomy*, *101*, 163–173. https://doi.org/10.1016/j.eja.2018.09.006

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., et al. (2005). Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data. *Journal of Geophysical Research*, *110*(D24). https://doi.org/10.1029/2005jd006390

Dantec-Nédélec, S., Ottlé, C., Wang, T., Guglielmo, F., Maignan, F., Delbart, N., et al. (2017). Testing the capability of ORCHIDEE land surface model to simulate Arctic ecosystems: Sensitivity analysis and site-level model calibration. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1212–1230.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597.

De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Wang, Y.-P., et al. (2014). Where does the carbon go? A model–data intercomparison of vegetation carbon allocation and turnover processes at two temperate forest free-air CO_2 enrichment sites. *New Phytologist*, *203*(3), 883–899. https://doi.org/10.1111/nph.12847

Desai, A. R. (2010). Climatic and phenological controls on coherent regional interannual variability of carbon dioxide flux in a heterogeneous landscape. *Journal of Geophysical Research*, *115*. https://doi.org/10.1029/2010jg001423

Desroziers, G., Berre, L., Chapnik, B., & Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3385–3396. https://doi.org/10.1256/qj.05.108

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., et al. (2011). The international soil moisture network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, *15*(5), 1675–1698. https://doi.org/10.5194/hess-15-1675-2011

Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., et al. (2013). Global automated quality control of in situ soil moisture data from the international soil moisture network. *Vadose Zone Journal*, *12*(3). https://doi.org/10.2136/vzj2012.0097

Druel, A., Peylin, P., Krinner, G., Ciais, P., Viovy, N., Peregon, A., et al. (2017). Towards a more detailed representation of high-latitude vegetation in the global land surface model ORCHIDEE (ORC-HL-VEGv1.0). *Geoscientific Model Development*, *10*(12), 4693–4722. https://doi.org/10.5194/gmd-10-4693-2017

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., et al. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, *1*, 100002. https://doi.org/10.1016/j.srs.2020.100002

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., et al. (2013). Climate change projections using the IPSL-CM5 Earth system model: From CMIP3 to CMIP5. *Climate Dynamics*, *40*(9–10), 2123–2165.

Ellis, E. C., Gauthier, N., Klein Goldewijk, K., Bliege Bird, R., Boivin, N., Díaz, S., et al. (2021). People have shaped most of terrestrial nature for at least 12,000 years. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(17). https://doi.org/10.1073/pnas.2023483118

Exbrayat, J. F., Bloom, A. A., Carvalhais, N., Fischer, R., Huth, A., MacBean, N., & Williams, M. (2019). Understanding the land carbon cycle with space data: Current status and prospects. *Surveys in Geophysics*, *40*(4), 735–755.

Exbrayat, J.-F., Pitman, A. J., & Abramowitz, G. (2014). Response of microbial decomposition to spin-up explains CMIP5 soil carbon range until 2100. *Geoscientific Model Development*, *7*(6), 2683–2692.

Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., et al. (2021). Optimal model complexity for terrestrial carbon cycle prediction. *Biogeosciences*, *18*(8), 2727–2754.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(40), 10572–10577.

Farquhar, G. D., von Caemmerer, S. V., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta*, *149*(1), 78–90.

Feldman, A. F., Short Gianotti, D. J., Konings, A. G., Gentine, P., & Entekhabi, D. (2021). Patterns of plant rehydration and growth following pulses of soil moisture availability. *Biogeosciences*, *18*(3), 831–847.

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453.

Fisher, R. A., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., et al. (2018). Vegetation demographics in Earth system models: A review of progress and priorities. *Global Change Biology*, *24*(1), 35–54.

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*(10), 2471–2494. https://doi.org/10.1029/2018ms001362

Frankenberg, C., O'Dell, C., Berry, J., Guanter, L., Joiner, J., Köhler, P., et al. (2014). Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2. *Remote Sensing of Environment*, *147*, 1–12. https://doi.org/10.1016/j.rse.2014.02.007

Franklin, O., Harrison, S. P., Dewar, R., Farrior, C. E., Brännström, Å., Dieckmann, U., et al. (2020). Organizing principles for vegetation dynamics. *Nature Plants*, *6*(5), 444–453.

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.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., et al. (2019). Global carbon budget 2019. *Earth System Science Data*, *11*(4), 1783–1838.

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.

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., et al. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO_2. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3280–3285.

Gampe, D., Zscheischler, J., Reichstein, M., O'Sullivan, M., Smith, W. K., Sitch, S., & Buermann, W. (2021). Increasing impact of warm droughts on northern ecosystem productivity over recent decades. *Nature Climate Change*. https://doi.org/10.1038/s41558-021-01112-8

Gatti, L. V., Basso, L. S., Miller, J. B., Gloor, M., Gatti Domingues, L., Cassol, H. L. G., et al. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, *595*(7867), 388–393.

Giering, R., Kaminski, T., & Slawig, T. (2005). Generating efficient derivative code with TAF. *Future Generation Computer Systems*, *21*(8), 1345–1355. https://doi.org/10.1016/j.future.2004.11.003

GLOBALVIEW CO2. (2013). Cooperative Global Atmospheric Data Integration Project, updated annually, Multi-laboratory compilation of synchronized and gap-filled atmospheric carbon dioxide records for the period 1979–2012 (obspack_co2_1_GLOBALVIEW-CO2_2013_v1.0.4_2013-12-23), compiled by NOAA Global Monitoring Division: Boulder, Colorado, USA Data product. https://doi.org/10.3334/OBSPACK/1002

Granier, A., Bréda, N., Longdoz, B., Gross, P., & Ngao, J. (2008). Ten years of fluxes and stand growth in a young beech forest at Hesse, North-eastern France. *Annals of Forest Science*, *65*(7), 1.

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., et al. (2012). Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. *Remote Sensing of Environment*, *121*, 236–251. https://doi.org/10.1016/j.rse.2012.02.006

Harden, J. W., Hugelius, G., Ahlström, A., Blankinship, J. C., Bond-Lamberty, B., Lawrence, C. R., et al. (2018). Networking our science to characterize the state, vulnerabilities, and management opportunities of soil organic matter. *Global Change Biology*, *24*(2), e705–e718.

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations - The CRU TS3.10 Dataset. *International Journal of Climatology*, *34*(3), 623–642.

Haupt, R. L., & Haupt, S. E. (2004). *Practical genetic algorithms*. John Wiley & Sons.

Haverd, V., Raupach, M. R., Briggs, P. R., Canadell, J. G., Isaac, P., Pickett-Heaps, C., et al. (2013). Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles. *Biogeosciences*, *10*(3), 2011–2040.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., et al. (2006). The LMDZ4 general circulation model: Climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics*, *27*(7–8), 787–813. https://doi.org/10.1007/s00382-006-0158-0

Huang, J., Tian, L., Liang, S., Ma, H., Becker-Reshef, I., Huang, Y., et al. (2015). Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. *Agricultural and Forest Meteorology*, *204*, 106–121. https://doi.org/10.1016/j.agrformet.2015.02.001

Hubau, W., Lewis, S. L., Phillips, O. L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., et al. (2020). Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature*, *579*(7797), 80–87.

Huntzinger, D. N., Michalak, A. M., Schwalm, C., Ciais, P., King, A. W., Fang, Y., et al. (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific Reports*, *7*(1), 4765.

Huxman, T. E., Snyder, K. A., Tissue, D., Leffler, A. J., Ogle, K., Pockman, W. T., et al. (2004). Precipitation pulses and carbon fluxes in semiarid and arid ecosystems. *Oecologia*, *141*(2), 254–268.

Jeong, J., Barichivich, J., Peylin, P., Haverd, V., McGrath, M. J., Vuichard, N., et al. (2020). *Using the International Tree-Ring Data Bank (ITRDB) records as century-long benchmarks for land-surface models*. https://doi.org/10.5194/gmd-2020-29

Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., et al. (2021). A restructured and updated global soil respiration database (SRDB-V5). *Earth System Science Data*, *13*(2), 255–267.

Jiang, M., Medlyn, B. E., Drake, J. E., Duursma, R. A., Anderson, I. C., Barton, C. V., et al. (2020). The fate of carbon in a mature forest under carbon dioxide enrichment. *Nature*, *580*(7802), 227–231.

Joiner, J., Yoshida, Y., Vasilkov, A. P., Schaefer, K., Jung, M., Guanter, L., et al. (2014). The seasonal cycle of satellite chlorophyll fluorescence observations and its relationship to vegetation phenology and ecosystem atmosphere carbon exchange. *Remote Sensing of Environment*, *152*, 375–391.

Jones, C. D., & Friedlingstein, P. (2020). Quantifying process-level uncertainty contributions to TCRE and carbon budgets for meeting Paris Agreement climate targets. *Environmental Research Letters: ERL [Web Site]*, *15*(7), 074019.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Altaf Arain, M., et al. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research*, *116*. https://doi.org/10.1029/2010jg001566

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., et al. (2013). The BETHY/JSBACH carbon cycle data assimilation system: Experiences and challenges. *Journal of Geophysical Research: Biogeosciences*, *118*(4), 1414–1426. https://doi.org/10.1002/jgrg.20118

Keenan, T. F., Baker, I., Barr, A., Ciais, P., Davis, K., Dietze, M., et al. (2012). Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO_2 exchange. *Global Change Biology*, *18*(6), 1971–1987. https://doi.org/10.1111/j.1365-2486.2012.02678.x

Keenan, T. F., Davidson, E. A., Munger, J. W., & Richardson, A. D. (2013). Rate my data: Quantifying the value of ecological data for the development of models of the terrestrial carbon cycle. *Ecological Applications: Ecological Society of America*, *23*(1), 273–286.

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., da Costa, A. C. L., & Gentine, P. (2019). Implementing plant hydraulics in the community land model, version 5. *Journal of Advances in Modeling Earth Systems*, *11*(2), 485–513. https://doi.org/10.1029/2018ms001500

Kleidon, A., & Heimann, M. (1997). *A method of determining rooting depth from a terrestrial biosphere model and its impacts on the global water- and carbon cycle*.

Köhler, P., Guanter, L., & Joiner, J. (2015). A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data. *Atmospheric Measurement Techniques*, *8*(6), 2589–2608. https://doi.org/10.5194/amt-8-2589-2015

Kolassa, J., Reichle, R. H., Koster, R. D., Liu, Q., Mahanama, S., & Zeng, F. W. (2020). An observation-driven approach to improve vegetation phenology in a global land surface model. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2020MS002083.

Kondo, M., Patra, P. K., Sitch, S., Friedlingstein, P., Poulter, B., Chevallier, F., et al. (2020). State of the science in reconciling top-down and bottom-up approaches for terrestrial CO budget. *Global Change Biology*, *26*(3), 1068–1084.
Konings, A. G., Bloom, A. A., Liu, J., Parazoo, N. C., Schimel, D. S., & Bowman, K. W. (2019). Global satellite-driven estimates of heterotrophic respiration. *Biogeosciences*, *16*(11), 2269–2284.
Konings, A. G., Piles, M., Rötzer, K., McColl, K. A., Chan, S. K., & Entekhabi, D. (2016). Vegetation optical depth and scattering albedo retrieval using time series of dual-polarized L-band radiometer observations. *Remote Sensing of Environment*, *172*, 178–189. https://doi.org/10.1016/j.rse.2015.11.009
Konings, A. G., Rao, K., & Steele-Dunne, S. C. (2019). Macro to micro: Microwave remote sensing of plant water content for physiology and ecology. *New Phytologist*, *223*(3), 1166–1172.
Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., et al. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(1). https://doi.org/10.1029/2003gb002199
Kuppel, S., Chevallier, F., & Peylin, P. (2013). Quantifying the model structural error in carbon cycle data assimilation systems. *Geoscientific Model Development*, *6*(1), 45–55. https://doi.org/10.5194/gmd-6-45-2013
Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., & Richardson, A. D. (2012). Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*, *9*(10), 3757–3776. https://doi.org/10.5194/bg-9-3757-2012
Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., & Cescatti, A. (2014). Model–data fusion across ecosystems: From multisite optimizations to global simulations. *Geoscientific Model Development*, *7*(6), 2581–2597. https://doi.org/10.5194/gmd-7-2581-2014
Kwon, M., Ballantyne, A., Ciais, P., Qiu, C., Salmon, E., Raoult, N., et al. (2022). Lowering water table reduces carbon sink strength and carbon stock in northern circumarctic peatlands. Submitted to global change biology.
Lawrence, C. R., Beem-Miller, J., Hoyt, A. M., Monroe, G., Sierra, C. A., Stoner, S., et al. (2020). An open-source database for the synthesis of soil radiocarbon data: International Soil Radiocarbon Database (ISRaD) version 1.0. *Earth System Science Data*, *12*(1), 61–76.
Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (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.
Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., et al. (2018). Global carbon budget 2018. *Earth System Science Data*, *10*(4), 2141–2194.
Li, L., Yang, Z., Matheny, A. M., Zheng, H., Swenson, S. C., Lawrence, D. M., et al. (2021). Representation of plant hydraulics in the noah-MP land surface model: Model development and multiscale evaluation. *Journal of Advances in Modeling Earth Systems*, *13*(4). https://doi.org/10.1029/2020ms002214
Li, Y., Yang, H., Wang, T., MacBean, N., Bacour, C., Ciais, P., et al. (2017). Reducing the uncertainty of parameters controlling seasonal carbon and water fluxes in Chinese forests and its implication for simulated climate sensitivities. *Global Biogeochemical Cycles*, *31*(8), 1344–1366. https://doi.org/10.1002/2017gb005714
Ling, X. L., Fu, C. B., Yang, Z. L., & Guo, W. D. (2019). Comparison of different sequential assimilation algorithms for satellite-derived leaf area index using the Data Assimilation Research Testbed (version Lanai). *Geoscientific Model Development*, *12*(7), 3119–3133.
Liu, D., Li, Y., Wang, T., Peylin, P., MacBean, N., Ciais, P., et al. (2018). Contrasting responses of grassland water and carbon exchanges to climate change between Tibetan Plateau and Inner Mongolia. *Agricultural and Forest Meteorology*, *249*, 163–175. https://doi.org/10.1016/j.agrformet.2017.11.034
Liu, J., Baskaran, L., Bowman, K., Schimel, D., Bloom, A. A., Parazoo, N. C., et al. (2021). Carbon monitoring system flux net biosphere exchange 2020 (CMS-flux NBE 2020). *Earth System Science Data*, *13*(2), 299–330.
Liu, Y. Y., de Jeu, R. A. M., McCabe, M. F., Evans, J. P., & van Dijk, A. I. J. M. (2011). Global long-term passive microwave satellite-based retrievals of vegetation optical depth. *Geophysical Research Letters*, *38*(18). https://doi.org/10.1029/2011gl048684
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*(1), 40–56.
MacBean, N., Bastrikov, V., Bacour, C., & Raoult, N. (2022). *nmacbean/ORCHIDEE-Param-DA-Review-GBC/v1*. Zenodo, https://doi.org/10.5281/zenodo.6621632
MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., et al. (2018). Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. *Scientific Reports*, *8*(1), 1973.
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
MacBean, N., Peylin, P., Chevallier, F., Scholze, M., & Schürmann, G. (2016). Consistent assimilation of multiple data streams in a carbon cycle data assimilation system. *Geoscientific Model Development*, *9*(10), 3569–3588. https://doi.org/10.5194/gmd-9-3569-2016
MacBean, N., Scott, R. L., Biederman, J. A., Peylin, P., Kolb, T., Litvak, M. E., et al. (2021). Dynamic global vegetation models underestimate net CO_2 flux mean and inter-annual variability in dryland ecosystems. *Environmental Research Letters: ERL [Web Site]*, *16*(9), 094023.
Magney, T. S., Barnes, M. L., & Yang, X. (2020). On the covariation of chlorophyll fluorescence and photosynthesis across scales. *Geophysical Research Letters*, *47*(23). https://doi.org/10.1029/2020gl091098
Mahmud, K., Medlyn, B. E., Duursma, R. A., Campany, C., & De Kauwe, M. G. (2018). Inferring the effects of sink strength on plant carbon balance processes from experimental measurements. *Biogeosciences*, *15*(13), 4003–4018.
Mahmud, K., Scott, R. L., Biederman, J. A., Litvak, M. E., Kolb, T., Meyers, T. P., et al. (2021). Optimizing carbon cycle parameters drastically improves terrestrial biosphere model underestimates of dryland mean net CO_2 flux and its inter-annual variability. *Journal of Geophysical Research: Biogeosciences*, *126*(10), e2021JG006400.
Maignan, F., Abadie, C., Remaud, M., Kooijmans, L. M. J., Kohonen, K.-M., Commane, R., et al. (2021). Carbonyl sulfide: Comparing a mechanistic representation of the vegetation uptake in a land surface model and the leaf relative uptake approach. *Biogeosciences*, *18*(9), 2917–2955.
Matthes, J. H., Goring, S., Williams, J. W., & Dietze, M. C. (2016). Benchmarking historical CMIP5 plant functional types across the Upper Midwest and Northeastern United States. *Journal of Geophysical Research: Biogeosciences*, *121*(2), 523–535.
McCarthy, H. R., Oren, R., Johnsen, K. H., Gallet-Budynek, A., Pritchard, S. G., Cook, C. W., et al. (2010). Re-assessment of plant carbon dynamics at the Duke free-air CO_2 enrichment site: Interactions of atmospheric [CO_2] with nitrogen and water availability over stand development. *New Phytologist*, *185*(2), 514–528. https://doi.org/10.1111/j.1469-8137.2009.03078.x
Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., et al. (2015). Using ecosystem experiments to improve vegetation models. *Nature Climate Change*, *5*(6), 528–534.
Melton, J. R., Verseghy, D. L., Sospedra-Alfonso, R., & Gruber, S. (2019). Improving permafrost physics in the coupled Canadian land surface scheme (v.3.6.2) and Canadian terrestrial ecosystem model (v.2.1) (CLASS-CTEM). *Geoscientific Model Development*, *12*, 4443–4467. https://doi.org/10.5194/gmd-12-4443-2019

Moesinger, L., Dorigo, W., de Jeu, R., van der Schalie, R., Scanlon, T., Teubner, I., & Forkel, M. (2020). The global long-term microwave vegetation optical depth climate archive (VODCA). *Earth System Science Data*, *12*(1), 177–196. https://doi.org/10.5194/essd-12-177-2020

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*(9), 3499–3517.

Moore, D. J. P., 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 CO_2 and H_2O fluxes. *Agricultural and Forest Meteorology*, *148*(10), 1467–1477. https://doi.org/10.1016/j.agrformet.2008.04.013

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174. https://doi.org/10.1080/00401706.1991.10484804

Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., et al. (2015). A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes. *Geoscientific Model Development*, *8*(7), 2035–2065.

Nave, L., Johnson, K., van Ingen, C., Agarwal, D., Humphrey, M., & Beekwilder, N. (2016). International soil carbon network (ISCN) database v3-1 [Dataset]. International Soil Carbon Network (ISCN). https://doi.org/10.17040/ISCN/1305039

Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E., & McMurtrie, R. E. (2010). CO_2 enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(45), 19368–19373.

Norton, A. J., Rayner, P. J., Koffi, E. N., Scholze, M., Silver, J. D., & Wang, Y.-P. (2019). Estimating global gross primary productivity using chlorophyll fluorescence and a data assimilation system with the BETHY-SCOPE model. *Biogeosciences*, *16*(15), 3069–3093.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., et al. (2006). Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: Algorithms and uncertainty estimation. *Biogeosciences*, *3*(4), 571–583.

Parazoo, N. C., Bowman, K., Fisher, J. B., Frankenberg, C., Jones, D. B., Cescatti, A., et al. (2014). Terrestrial gross primary production inferred from satellite fluorescence and vegetation models. *Global Change Biology*, *20*(10), 3103–3121.

Parazoo, N. C., Magney, T., Norton, A., Raczka, B., Bacour, C., Maignan, F., et al. (2020). Wide discrepancies in the magnitude and direction of modeled solar-induced chlorophyll fluorescence in response to light conditions. *Biogeosciences*, *17*(13), 3733–3755.

Parton, W. J., Schimel, D. S., Cole, C. V., & Ojima, D. S. (1987). Analysis of factors controlling soil organic matter levels in Great Plains grasslands. *Soil Science Society of America Journal*, *51*(5), 1173–1179.

Paschalis, A., Fatichi, S., Zscheischler, J., Ciais, P., Bahn, M., Boysen, L., et al. (2020). Rainfall manipulation experiments as simulated by terrestrial biosphere models: Where do we stand? *Global Change Biology*, *26*(6), 3336–3355.

Peaucelle, M., Bacour, C., Ciais, P., Vuichard, N., Kuppel, S., Peñuelas, J., et al. (2019). Covariations between plant functional traits emerge from constraining parameterization of a terrestrial biosphere model. *Global Ecology and Biogeography: A Journal of Macroecology*, *28*(9), 1351–1365.

Peaucelle, M., Ciais, P., Maignan, F., Nicolas, M., Cecchini, S., & Viovy, N. (2019). Representing explicit budburst and senescence processes for evergreen conifers in global models. *Agricultural and Forest Meteorology*, *266*, 97–108.

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., et al. (2022). Four years of global carbon cycle observed from the Orbiting Carbon Observatory 2 (OCO-2) version 9 and in situ data and comparison to OCO-2 version 7. *Atmospheric Chemistry and Physics*, *22*(2), 1097–1130.

Peng, S., Ciais, P., Chevallier, F., Peylin, P., Cadule, P., Sitch, S., et al. (2015). Benchmarking the seasonal cycle of CO_2 fluxes simulated by terrestrial ecosystem models. *Global Biogeochemical Cycles*, *29*(1), 46–64. https://doi.org/10.1002/2014gb004931

Peng, S., Piao, S., Shen, Z., Ciais, P., Sun, Z., Chen, S., et al. (2013). Precipitation amount, seasonality and frequency regulate carbon cycling of a semi-arid grassland ecosystem in inner Mongolia, China: A modeling analysis. *Agricultural and Forest Meteorology*, *178–179*, 46–55. https://doi.org/10.1016/j.agrformet.2013.02.002

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., et al. (2016). A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle. *Geoscientific Model Development*, *9*(9), 3321–3346. https://doi.org/10.5194/gmd-9-3321-2016

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., et al. (2013). Global atmospheric carbon budget: Results from an ensemble of atmospheric CO_2 inversions. *Biogeosciences*, *10*(10), 6699–6720.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., et al. (2013). Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends. *Global Change Biology*, *19*(7), 2117–2132.

Piao, S., Wang, X., Wang, K., Li, X., Bastos, A., Canadell, J. G., et al. (2020). Interannual variation of terrestrial carbon cycle: Issues and perspectives. *Global Change Biology*, *26*(1), 300–318.

Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., & Scoby, D. (2020). The land variational ensemble data assimilation framework: LAVENDAR v1.0.0. *Geoscientific Model Development*, *13*(1), 55–69. https://doi.org/10.5194/gmd-13-55-2020

Pinnington, E. M., Casella, E., Dance, S. L., Lawless, A. S., Morison, J. I., Nichols, N. K., et al. (2017). Understanding the effect of disturbance from selective felling on the carbon dynamics of a managed woodland by combining observations with model predictions. *Journal of Geophysical Research: Biogeosciences*, *122*(4), 886–902.

Pinnington, E. M., Casella, E., Dance, S. L., Lawless, A. S., Morison, J. I. L., Nichols, N. K., et al. (2016). Investigating the role of prior and observation error correlations in improving a model forecast of forest carbon balance using Four-dimensional Variational data assimilation. *Agricultural and Forest Meteorology*, *228*(229), 299–314. https://doi.org/10.1016/j.agrformet.2016.07.006

Pinty, B., Lavergne, T., Dickinson, R. E., Widlowski, J.-L., Gobron, N., & Verstraete, M. M. (2006). Simplifying the interaction of land surfaces with radiation for relating remote sensing products to climate models. *Journal of Geophysical Research*, *111*(D2). https://doi.org/10.1029/2005jd005952

Pitman, A. J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *23*(5), 479–510.

Pongratz, J., Dolman, H., Don, A., Erb, K.-H., Fuchs, R., Herold, M., et al. (2018). Models meet data: Challenges and opportunities in implementing land management in Earth system models. *Global Change Biology*, *24*(4), 1470–1487.

Porte, A., Trichet, P., Bert, D., & Loustau, D. (2002). Allometric relationships for branch and tree woody biomass of Maritime pine (Pinus pinaster Aıt.). *Forest Ecology and Management*, *158*(1–3), 71–83.

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., et al. (2014). Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature*, *509*(7502), 600–603.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., et al. (2015). Plant functional type classification for Earth system models: Results from the European space agency's land cover climate change initiative. *Geoscientific Model Development*, *8*(7), 2315–2328.

Qin, Y., Xiao, X., Wigneron, J.-P., Ciais, P., Brandt, M., Fan, L., et al. (2021). Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon. *Nature Climate Change*, *11*(5), 442–448. https://doi.org/10.1038/s41558-021-01026-5

Quaife, T., Lewis, P., De Kauwe, M., Williams, M., Law, B. E., Disney, M., & Bowyer, P. (2008). Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter. *Remote Sensing of Environment*, *112*(4), 1347–1364.

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Minh, D. H. T., et al. (2019). The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space. *Remote Sensing of Environment*, *227*, 44–60. https://doi.org/10.1016/j.rse.2019.03.032

Raoult, N., Ottlé, C., Peylin, P., Bastrikov, V., & Maugis, P. (2021). Evaluating and optimizing surface soil moisture drydowns in the ORCHIDEE land surface model at in situ locations. *Journal of Hydrometeorology*, *22*(4), 1025–1043.

Raoult, N. M., Jupp, T. E., Cox, P. M., & Luke, C. M. (2016). Land-surface parameter optimisation using data assimilation techniques: The adJULES system V1.0. *Geoscientific Model Development*, *9*(8), 2833–2852. https://doi.org/10.5194/gmd-9-2833-2016

Rayner, P. J. (2010). The current state of carbon-cycle data assimilation. *Current Opinion in Environmental Sustainability*, *2*(4), 289–296.

Rayner, P. J., Michalak, A. M., & Chevallier, F. (2019). Fundamentals of data assimilation applied to biogeochemistry. *Atmospheric Chemistry and Physics*, *19*(22), 13911–13932.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., & Widmann, H. (2005). Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global Biogeochemical Cycles*, *19*(2).

Remaud, M., Chevallier, F., Maignan, F., Belviso, S., Berchet, A., Parouffe, A., et al. (2021). *Plant gross primary production, plant respiration and carbonyl sulfide emissions over the globe inferred by atmospheric inverse modelling*. https://doi.org/10.5194/acp-2021-326

Ricciuto, D. M., Davis, K. J., & Keller, K. (2008). A Bayesian calibration of a simple carbon cycle model: The role of observations in estimating and reducing uncertainty. *Global Biogeochemical Cycles*, *22*(2). https://doi.org/10.1029/2006gb002908

Ricciuto, D. M., King, A. W., Dragoni, D., & Post, W. M. (2011). Parameter and prediction uncertainty in an optimized terrestrial carbon cycle model: Effects of constraining variables and data record length. *Journal of Geophysical Research*, *116*(G1). https://doi.org/10.1029/2010jg001400

Richardson, A. D., Anderson, R. S., Altaf Arain, M., Barr, A. G., Bohrer, G., Chen, G., et al. (2012). Terrestrial biosphere models need better representation of vegetation phenology: Results from the north American carbon Program site synthesis. *Global Change Biology*, *18*(2), 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x

Salmon, E., Jégou, F., Guenet, B., Jourdain, L., Qiu, C., Bastrikov, V., et al. (2022). Assessing methane emissions for northern peatlands in ORCHIDEE-PEAT revision 7020. *Geoscientific Model Development*, *15*(7), 2813–2838.

Santaren, D., Peylin, P., Bacour, C., Ciais, P., & Longdoz, B. (2014). Ecosystem model optimization using in situ flux observations: Benefit of Monte Carlo versus variational schemes and analyses of the year-to-year model performances. *Biogeosciences*, *11*(24), 7137–7158. https://doi.org/10.5194/bg-11-7137-2014

Santaren, D., Peylin, P., Viovy, N., & Ciais, P. (2007). Optimizing a process-based ecosystem model with eddy-covariance flux measurements: A pine forest in southern France. *Global Biogeochemical Cycles*, *21*(2). https://doi.org/10.1029/2006gb002834

Scheiter, S., Langan, L., & Higgins, S. I. (2013). Next-generation dynamic global vegetation models: Learning from community ecology. *New Phytologist*, *198*(3), 957–969.

Schimel, D., Stephens, B. B., & Fisher, J. B. (2015). Effect of increasing CO_2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences*, *112*(2), 436–441. https://doi.org/10.1073/pnas.1407302112

Schneider von Deimling, T., Grosse, G., Strauss, J., Schirrmeister, L., Morgenstern, A., Schaphoff, S., et al. (2015). Observation-based modelling of permafrost carbon fluxes with accounting for deep carbon deposits and thermokarst activity. *Biogeosciences*, *12*(11), 3469–3488.

Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., & Quegan, S. (2017). Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems. *Biogeosciences*, *14*(14), 3401–3429.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J. P., & Scipal, K. (2016). Simultaneous assimilation of SMOS soil moisture and atmospheric CO_2 in-situ observations to constrain the global terrestrial carbon cycle. *Remote Sensing of Environment*, *180*, 334–345.

Schürmann, G. J., Kaminski, T., Köstler, C., Carvalhais, N., Voßbeck, M., Kattge, J., et al. (2016). Constraining a land-surface model with multiple observations by application of the MPI-Carbon Cycle Data Assimilation System V1.0. *Geoscientific Model Development*, *9*(9), 2999–3026. https://doi.org/10.5194/gmd-9-2999-2016

Schwalm, C. R., Schaefer, K., Fisher, J. B., Huntzinger, D., Elshorbany, Y., Fang, Y., et al. (2019). Divergence in land surface modeling: Linking spread to structure. *Environmental Research Communications*, *1*(11), 111004.

Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F. G., Berry, J. A., et al. (1997). Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, *275*(5299), 502–509.

Shi, Z., Allison, S. D., He, Y., Levine, P. A., Hoyt, A. M., Beem-Miller, J., et al. (2020). The age distribution of global soil carbon inferred from radiocarbon measurements. *Nature Geoscience*, *13*(8), 555–559.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., et al. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, *12*(3), 653–679. https://doi.org/10.5194/bg-12-653-2015

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. *Geophysical Research Letters*, *45*(2), 748–757.

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Köhler, P., & Magney, T. (2018). Overview of solar-induced chlorophyll fluorescence (SIF) from the orbiting carbon observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP. *Remote Sensing of Environment*, *209*, 808–823. https://doi.org/10.1016/j.rse.2018.02.016

Tarantola, A. (2013). *Inverse problem theory: Methods for data fitting and model parameter estimation*. Elsevier.

Teckentrup, L., De Kauwe, M. G., Pitman, A. J., Goll, D., Haverd, V., Jain, A. K., et al. (2021). Assessing the representation of the Australian carbon cycle in global vegetation models. https://doi.org/10.5194/bg-2021-66

Thum, T., MacBean, N., Peylin, P., Bacour, C., Santaren, D., Longdoz, B., et al. (2017). The potential benefit of using forest biomass data in addition to carbon and water flux measurements to constrain ecosystem model parameters: Case studies at two temperate forest sites. *Agricultural and Forest Meteorology*, *234–235*, 48–65. https://doi.org/10.1016/j.agrformet.2016.12.004

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., & Allison, S. D. (2013). Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, *10*(3), 1717–1736. https://doi.org/10.5194/bg-10-1717-2013

Traore, A., Ciais, P., Vuichard, N., MacBean, N., Dardel, C., Poulter, B., et al. (2014). 1982–2010 trends of light use efficiency and inherent water use efficiency in african vegetation: Sensitivity to climate and atmospheric CO_2 concentrations. *Remote Sensing*, *6*(9), 8923–8944. https://doi.org/10.3390/rs6098923

Trudinger, C. M., Haverd, V., Briggs, P. R., & Canadell, J. G. (2016). Interannual variability in Australia's terrestrial carbon cycle constrained by multiple observation types. *Biogeosciences*, *13*(23), 6363–6383.

Van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., & Su, Z. (2009). An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. *Biogeosciences*, *6*(12), 3109–3129.

Verbeeck, H., Peylin, P., Bacour, C., Bonal, D., Steppe, K., & Ciais, P. (2011). Seasonal patterns of CO_2 fluxes in Amazon forests: Fusion of eddy covariance data and the ORCHIDEE model. *Journal of Geophysical Research*, *116*(G2). https://doi.org/10.1029/2010jg001544

Viovy, N. (2018). CRUNCEP version 7—atmospheric forcing data for the community land model [Dataset] . Retrieved from https://rda.ucar.edu/datasets/ds314.3/

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., et al. (2019). Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): Multi-scale evaluation of gross primary production. *Geoscientific Model Development*, *12*(11), 4751–4779. https://doi.org/10.5194/gmd-12-4751-2019

Walker, A. P., De Kauwe, M. G., Bastos, A., Belmecheri, S., Georgiou, K., Keeling, R. F., et al. (2021). Integrating the evidence for a terrestrial carbon sink caused by increasing atmospheric CO_2. *New Phytologist*, *229*(5), 2413–2445.

Walker, A. P., Hanson, P. J., De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Asao, S., et al. (2014). Comprehensive ecosystem model-data synthesis using multiple data sets at two temperate forest free-air CO_2 enrichment experiments: Model performance at ambient CO_2 concentration. *Journal of Geophysical Research: Biogeosciences*, *119*(5), 937–964. https://doi.org/10.1002/2013jg002553

Walker, A. P., Ye, M., Lu, D., De Kauwe, M. G., Gu, L., Medlyn, B. E., et al. (2018). The multi-assumption architecture and testbed (MAAT v1. 0): R code for generating ensembles with dynamic model structure and analysis of epistemic uncertainty from multiple sources. *Geoscientific Model Development*, *11*(8), 3159–3185.

Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Köhler, P., et al. (2016). Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. *Global Change Biology*, *22*(9), 2979–2996.

Wang, F., Cheruy, F., & Dufresne, J.-L. (2016). The improvement of soil thermodynamics and its effects on land surface meteorology in the IPSL climate model. *Geoscientific Model Development*, *9*(1), 363–381. https://doi.org/10.5194/gmd-9-363-2016

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., et al. (2013). Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model. *Journal of Geophysical Research: Atmospheres*, *118*(12), 6064–6079. https://doi.org/10.1002/jgrd.50395

Weiss, M., Baret, F., Garrigues, S., & Lacaze, R. (2007). LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: Validation and comparison with MODIS collection 4 products. *Remote sensing of Environment*, *110*(3), 317–331.

Wieder, W. R., Lawrence, D. M., Fisher, R. A., Bonan, G. B., Cheng, S. J., Goodale, C. L., et al. (2019). Beyond static benchmarking: Using experimental manipulations to evaluate land model assumptions. *Global Biogeochemical Cycles*, *33*(10), 1289–1309.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., et al. (2009). Improving land surface models with FLUXNET data. *Biogeosciences*, *6*(7), 1341–1359. https://doi.org/10.5194/bg-6-1341-2009

Worden, J., Saatchi, S., Keller, M., Bloom, A. A., Liu, J., Parazoo, N., et al. (2021). *Satellite observations of the tropical terrestrial carbon balance and interactions with the water cycle during the 21st century*.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., et al. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–827.

Wutzler, T., & Carvalhais, N. (2014). Balancing multiple constraints in model-data integration: Weights and the parameter block approach. *Journal of Geophysical Research: Biogeosciences*, *119*(11), 2112–2129. https://doi.org/10.1002/2014jg002650

Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., Pongratz, J., Bloom, A. A., et al. (2021). Changes in global terrestrial live biomass over the 21st century. *Science Advances*, *7*(27). https://doi.org/10.1126/sciadv.abe9829

Yin, Y., Bloom, A. A., Worden, J., Saatchi, S., Yang, Y., Williams, M., et al. (2020). Fire decline in dry tropical ecosystems enhances decadal land carbon sink. *Nature Communications*, *11*(1), 1900.

Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., et al. (2019). Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Science Advances*, *5*(8), eaax1396.

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., et al. (2014). Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO_2 Enrichment studies. *New Phytologist*, *202*(3), 803–822.

Zhang, X., Wang, Y.-P., Peng, S., Rayner, P. J., Ciais, P., Silver, J. D., et al. (2018). Dominant regions and drivers of the variability of the global land carbon sink across timescales. *Global Change Biology*, *24*(9), 3954–3968.

Ziehn, T., Kattge, J., Knorr, W., & Scholze, M. (2011). Improving the predictability of global CO2 assimilation rates under climate change. *Geophysical Research Letters*, *38*(10).

Zobitz, J. M., Moore, D. J., 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*, 73–88.