Optimizing Phenology Parameters Drastically Improves Terrestrial Biosphere Model Underestimates of Dryland Net CO2 Flux Inter-Annual Variability

Kashif Mahmud[1], Joel Biederman[2], Russ Scott[3], Marcy Litvak[4], Thomas Kolb[5], Tilden Meyers[6], Praveena Krishnan[6], Vladislav Bastrikov[7], and Natasha MacBean[1]

[1]Indiana University
[2]USDA-ARS Southwest Watershed Research Center
[3]United States Department of Agriculture, Agricultural Research Service, Tucson, AZ 85719, USA
[4]University of New Mexico
[5]Northern Arizona University
[6]5NOAA/ARL Atmospheric Turbulence and Diffusion Division
[7]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL

November 30, 2022

Abstract

Dryland ecosystems occupy ˜40% of the land surface and are thought to dominate the inter-annual variability (IAV) and long-term trend of the global carbon (C) cycle. Therefore, it is imperative that global terrestrial biosphere models (TBMs), which form the land component of IPCC earth system models, are able to accurately simulate dryland vegetation and biogeochemical processes. However, compared to more mesic ecosystems, TBMs have not been widely tested or optimized against in situ dryland ecosystem CO2 fluxes. Here, we address this gap using a Bayesian data assimilation system and 89 site-years of daily net CO2 flux (net ecosystem exchange - NEE) data from 12 southwest US Ameriflux sites spanning forest, shrub and grass dryland ecosystems to evaluate and optimize the C cycle related parameters of the ORCHIDEE TBM. We find that the default (prior) model simulations drastically underestimate both the mean annual NEE and the NEE IAV. By testing different assimilation scenarios, we showed that optimizing phenology parameters dramatically improves the model ability across all sites to capture both the magnitude and sign of the NEE IAV. At high elevation forested sites, which are a mean C sink, optimizing parameters related to C allocation, respiration and turnover reduces the underestimate in simulated mean annual NEE. Our study demonstrates that all TBMs need to be calibrated specifically for dryland ecosystems before they are used to determine dryland contributions to global C cycle variability and long-term carbon-climate feedbacks.

Optimizing Phenology Parameters Drastically Improves Terrestrial Biosphere Model Underestimates of Dryland Net CO_2 Flux Inter-Annual Variability

K. Mahmud[1], J. A. Biederman[2], R. L. Scott[2], M. E. Litvak[3], T. Kolb[4], T. P. Meyers[5], P. Krishnan[5,6], V. Bastrikov[7,8], and N. MacBean[1]

[1]Department of Geography, Indiana University, Bloomington, IN 47405, USA

[2]Southwest Watershed Research Center, United States Department of Agriculture, Agricultural Research Service, Tucson, AZ 85719, USA

[3]Department of Biology, University of New Mexico, Albuquerque, NM, 87131, USA

[4]School of Forestry, Northern Arizona University, Flagstaff, AZ, 86011, USA

[5]NOAA/ARL Atmospheric Turbulence and Diffusion Division, Oak Ridge, TN, 37830, USA

[6]Oak Ridge Associated Universities, Oak Ridge, TN, 37830, USA

[7]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, F-91191, France

[8]Now at: Science Partners, Paris, 75010, France

Corresponding author: Kashif Mahmud (kmahmud@iu.edu)

Key Points:

- ORCHIDEE terrestrial biosphere model drastically underestimates dryland mean annual net CO_2 fluxes and their inter-annual variability (IAV)
- Optimizing phenology, carbon allocation, and respiration parameters are crucial for capturing net CO_2 flux mean and IAV
- Models need to be optimized against dryland CO_2 flux data to achieve accurate predictions of dryland's role in global C cycle variability

Abstract

Dryland ecosystems occupy ~40% of the land surface and are thought to dominate the inter-annual variability (IAV) and long-term trend of the global carbon (C) cycle. Therefore, it is imperative that global terrestrial biosphere models (TBMs), which form the land component of IPCC earth system models, are able to accurately simulate dryland vegetation and biogeochemical processes. However, compared to more mesic ecosystems, TBMs have not been widely tested or optimized against in situ dryland ecosystem CO_2 fluxes. Here, we address this gap using a Bayesian data assimilation system and 89 site-years of daily net CO_2 flux (net ecosystem exchange - NEE) data from 12 southwest US Ameriflux sites spanning forest, shrub and grass dryland ecosystems to evaluate and optimize the C cycle related parameters of the ORCHIDEE TBM. We find that the default (prior) model simulations drastically underestimate both the mean annual NEE and the NEE IAV. By testing different assimilation scenarios, we showed that optimizing phenology parameters dramatically improves the model ability across all sites to capture both the magnitude and sign of the NEE IAV. At high elevation forested sites, which are a mean C sink, optimizing parameters related to C allocation, respiration and turnover reduces the underestimate in simulated mean annual NEE. Our study demonstrates that all TBMs need to be calibrated specifically for dryland ecosystems before they are used to determine dryland contributions to global C cycle variability and long-term carbon-climate feedbacks.

1 Introduction

Terrestrial ecosystems currently take up ~30% of anthropogenic CO_2 emissions, thus acting as a substantial global carbon (C) sink (Fu et al., 2017) and providing a critical reduction in the rate of global warming. However, while we know the magnitude of the global C sink to a good degree of certainty, our knowledge of other components of the global C cycle are more uncertain. One such knowledge gap is which ecosystems, and/or which processes, are driving inter-annual variability (IAV) in land net C uptake (Fu et al., 2017). Improving our understanding of the IAV characteristics of the global terrestrial C cycle is key to being able to forecast the future of the land C sink and long-term biosphere-climate feedback (Cox et al., 2013).

Recent studies have pointed to drylands (arid and semi-arid ecosystems) as the dominant driver of global terrestrial C cycle IAV (Ahlström et al., 2015; Cleverly et al., 2016; Haverd et al., 2017; Poulter et al., 2014). High annual variability in net CO_2 exchange in response to plant-available moisture is observed in site-based flux studies in these regions (Biederman et al., 2017; Cleverly et al., 2016; Haverd et al., 2017; Scott et al., 2015). However, the global terrestrial biosphere models (TBMs) used in the recent C cycle IAV regional attribution studies have often only been extensively evaluated against data in more mesic ecosystems (e.g. (Peng et al., 2015; Piao et al., 2013; Raczka et al., 2013; Schaefer et al., 2012)), although studies have evaluated models against eddy covariance flux data from Australian dryland sites (Haverd et al., 2013a; Whitley et al., 2016a). TBM optimization (parameter calibration) has also only typically been carried out using temperate and boreal site data (Haverd et al., 2013a; Kuppel et al., 2014). Therefore, there remains a relative gap in model benchmarking and optimization using dryland C cycle related data.

Model benchmarking and optimization studies that have been performed in dryland regions indicate considerable model-data discrepancies in vegetation dynamics, C and water fluxes (Haverd et al., 2013b; MacBean et al., 2015; Renwick et al., 2019; Trudinger et al., 2016; Whitley et al., 2016b; Traore et al., 2014). A recent study by MacBean et al. (submitted) has demonstrated that global TBMs participating in the TRENDY v7 model intercomparison project dramatically underestimate both the mean annual net ecosystem exchange (NEE) and its IAV at a suite of southwestern (SW) US dryland sites due to weak sensitivity of gross primary productivity (GPP) to changing water availability. This analysis is corroborated by (Renwick et al., 2019) who also showed that a semi-deciduous phenology scheme was necessary to accurately predict the magnitude of GPP in dryland shrublands. SW US hydrology modeling studies have also suggested that parameter calibration is needed to realistically represent semi-arid water fluxes because the default parameters diminish model performance (Natasha MacBean et al., 2020; Hogue et al., 2005; Unland et al., 1996). Given the lack of model parameter calibration studies that have included dryland sites in their optimizations, it remains to be seen whether model-data discrepancies in dryland ecosystem NEE simulations are due to inaccurate model processes or uncertain parameters. Parameter uncertainty may be higher for dryland ecosystems given parameter values were initially measured in the field and/or optimized for more mesic temperate and boreal ecosystems.

To address the gap in dryland site model parameter optimization, and to determine if parameter optimization can account for dryland model-data discrepancies in NEE, we used a Bayesian data assimilation (DA) framework to optimize the photosynthesis, phenology, C allocation and turnover, and respiration parameters of the ORCHIDEE TBM using 89 site-years of daily NEE observations of 12 Ameriflux sites spanning SW US semi-arid grass, shrub and forest ecosystems. Following (Biederman et al., 2017, MacBean et al, submitted), we categorize sites based on their mean annual NEE: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses are mostly tree-dominated C sink sites; shrub and grass-dominated sites US-Wkg, US-SRG, US-Seg, US-SRM, and US-Whs "pivot" between a mean annual C sink and source; and the US-Aud grassland is a mean source of C. We used the well-established DA system designed for ORCHIDEE (ORCHIDAS: https://orchidas.lsce.ipsl.fr) (Kuppel et al., 2014; MacBean et al., 2018; Peylin et al., 2016), in which a cost function that represents the misfit between the model and the data - taking into account uncertainty in both - is iteratively minimized using the genetic algorithm (GA; see Methods and Data).

We performed multiple optimization tests with combinations of parameters related to different model processes in order to identify which processes were most influential in improving the model mean annual NEE and IAV. We focus in particular on which processes are responsible for model failure to capture NEE IAV. We focus on improving NEE IAV partly because of the dominant role dryland ecosystems are thought to play in controlling global C cycle IAV, and partly because we expect that, with the exception of sites that are a strong C sink, eddy covariance estimates of mean annual NEE may be impacted by uncertainties in CO_2 flux partitioning. We identified three main groups of parameters: parameters related to 1) phenology; 2) parameters related to photosynthesis; and 3) parameters related to all process calculations that occur after gross C uptake (i.e. C allocation, autotrophic and heterotrophic respiration and biomass and soil C turnover; hereafter grouped as "post C uptake" parameters). We split the parameters into these three groups because GPP has been shown to be the dominant control on dryland NEE IAV (MacBean et al., submitted); therefore, we expect that optimizing parameters related to one of both of these main two processes controlling GPP (i.e., phenology and photosynthesis) will result in the strongest improvements in NEE IAV. However, optimizing all parameters related to processes that occur after gross C uptake can also influence NEE; therefore, we include these parameters as a third category. The parameters included in each assimilation scenario are: P1 - all parameters, including all three phenology, photosynthesis and post C uptake parameter groups; P2 - phenology and photosynthesis parameters; P3 - phenology and post C uptake; P4 - photosynthesis and post C uptake; P5 - phenology parameters only; P6 - photosynthesis only; and P7 - post C uptake only. See **Table 2** for a description of all parameters and to which category they belong. We compared the prior simulation (before parameter optimization) to the posterior simulations (after parameter optimization, with different parameter groupings for the different assimilation scenarios) by evaluating the simulations against the site data using standard goodness of fit metrics (root mean square error, RMSE and correlation coefficient, r) at daily, monthly and inter-annual timescales. We further attributed what might be causing model-data misfits by decomposing the daily mean squared deviation (MSD) into its component phase, variance and bias contributions. The bias, variance and phase indicate the mean difference in flux magnitude, the mismatch in terms of the magnitude of fluctuations, and the seasonality in flux time series, respectively (Kobayashi & Salam, 2000).

2 Methods and Data

2.1 Study sites

Twelve semi-arid sites in the southwestern US (SW US) have been utilized in this study. These sites have a range of different vegetation types, climates and have been described in detail by (Biederman et al., 2017), so we only provide a brief description here. We summarized the sites' description, dominant vegetation species, mean climate and corresponding vegetation plant functional types (PFTs), together with the observation period and disturbance history (**Table 1)**. The major regional IGBP vegetation classes represented include evergreen needleleaf forest, woody savanna, open and closed shrubland, and grassland. These sites typically experience monsoon rainfall during July to October, preceded by a hot, dry period in May and June. The SW US is characterized by water limitation at the annual scale, i.e. potential ET is greater than precipitation. The sites have large spatial gradients in mean annual precipitation (MAP 250–724 mm) and temperature (MAT 2.9 to 17.7°C) due to interactions among topography, latitude, wind patterns, and distance from oceans. For further site details, see references in **Table 1** and individual site pages on www.ameriflux.lbl.gov.

Table 1. Site descriptions, mean climate, observation years and corresponding vegetation plant functional types (PFTs) used in ORCHIDEE optimization. Simulation period corresponds to the period of available site data. PFT acronyms: BS = Bare soil (PFT=1); TeNE = Temperate Needleleaved Evergreen forest (PFT=4); TeBE = Temperate Broadleaved Evergreen forest (PFT=5); TeBD = Temperate Broadleaved Deciduous forest (PFT=6); C4G = C4 grass (PFT=11). Sites are given in order from largest mean annual C sink (US-Vcm) to mean annual C source (US-Aud).

Site ID	Description	Dominant species	IGBP class	PFT fractions	Koppen climate	Elevation (m)	MAP (mm)	MAT (°C)	Period of site data	Disturbance History	Site reference
US-Vcm	Valles Caldera mixed conifer forest	Picea engelmannii, Picea pugens, Abies lasiocarpa var. lasiocarpa, Abies concolor	Evergreen needleleaf forest	100% **TeNE**	Dfb	3042	724	2.9	2007–2012	Harvest 1960s	(Anderson-Teixeira et al., 2011)
US-Vcp	Valles Caldera ponderosa forest	Pinus ponderosa, Quercus gambeli	Evergreen needleleaf forest	100% **TeNE**	Dfb	2501	547	5.7	2007–2014	-	(Anderson-Teixeira et al., 2011)
US-Mpj	Heritage Land Conservancy pinyon-juniper	Pinus edulis, Juniperus monosperma	Savanna	20% BS; 60% **TeNE**; 20% **C4G**	Bsk	2200	423	9.6	2008–2014	-	(Anderson-Teixeira et al., 2011)
US-Fuf	Flagstaff unmanaged ponderosa	Pinus ponderosa	Evergreen needleleaf forest	100% **TeNE**	Csb	2215	607	7.1	2006–2010	Harvest 1910	(Dore et al., 2012)

US-Wjs	Tablelands juniper savanna	Juniperus monosperma, Bouteloua gracilis	Savanna	15% **TeNE**; 85% **C4G**	Bsk	1931	349	10.9	2008–2014	-	(Anderson-Teixeira et al., 2011)
US-Ses	Sevilleta creosote shrubland	Larrea tridentata, G. sarothrae	Open shrubland	20% BS; 55% **TeBE**; 25% **C4G**	Bsk	1610	252	12.6	2007–2014	-	(Petrie et al., 2015)
US-Wkg	Walnut Gulch Kendall grassland	Eragrostis lehmanniana, Bouteloua spp. Calliandra eriophylla	Grassland	60% BS; 3% TeBE; 37% **C4G**	Bsk	1529	386	15.8	2004–2013	Drought 2003-2005, non-native grass replacement 2007 onward, light grazing ongoing	(Scott, 2010)
US-SRG	Santa Rita grassland	Eragrostis lehmanniana	Savanna	45% BS; 11% **TeBD**; 44% **C4G**	Bsh	1292	494	16.7	2009–2014	Mesquite removal 1957, ongoing light grazing	(Scott et al., 2009, 2015)
US-Seg	Sevilleta grassland: burned 2009	Bouteloua eriopoda, Gutierrezia sarothrae, Ceratoides lanata	Grassland	40% BS; 60% **C4G**	Bsk	160	250	12.6	2007–2014	Burned 2009	(Petrie et al., 2015)
US-SRM	Santa Rita mesquite savanna	Prosopis velutina, Eragrostis lehmanniana	Woody savanna	50% BS; 35% **TeBD**; 15% **C4G**	Bsk	1122	421	17.7	2004–2014	Light grazing	(Scott et al., 2009)
US-Whs	Walnut Gulch Lucky Hills shrubland	Larrea tridentata, Parthenium incanum, Acacia constricta, Rhus microphylla	Open shrubland	57% BS; 40% **TeBE**; 3% C4G	Bsk	1376	352	16.8	2008–2014	Drought 2005-2006	(Scott, 2010)

US-Aud	Audubon grassland	Boutelou agracilis, B. curtipendula, Eragrostis spp.	Grassland	30% BS; 70% **C4G**	Bsk	1496	348	15.7	2004–2009	Burned 2002	(Krishnan et al., 2012)

2.1 ORCHIDEE terrestrial biosphere model

We use the ORCHIDEE (ORganizing Carbon and Hydrology In Dynamic EcosystEms) process-oriented land surface model version 2.2 that has been developed at the IPSL (Institut Pierre Simon Laplace, France). The model is a state-of-the-art mechanistic terrestrial biosphere model (Krinner et al., 2005) and is the land surface component of the IPSLCM5 Earth System Model (Dufresne et al., 2013). The model describes the exchanges of water, carbon, and energy between biosphere and atmosphere at the smallest time scale (30 min), while the slow components of the terrestrial carbon cycle (including carbon allocation, autotrophic respiration, foliar onset and senescence, mortality and soil organic matter decomposition) are computed on a daily to annual basis. Version 2.2 is virtually identical to version 2.0, which is being used in the ongoing Coupled Modeling Intercomparison Project 6 (CMIP6) simulations, but includes few recent bug corrections and code enhancements. It has been updated since the “AR5” version used in CMIP5 (see (Krinner et al., 2005)) with the following developments: i) an 11-layer mechanistic description of soil hydrology and associated modifications as described in (MacBean et al., 2020); ii) addition of a coupled carbon-nitrogen scheme (Vuichard et al., 2019); iii) an 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)), following (Yin and Struik, 2009); iv) anupdate of the soil thermal properties and extension of the soil depth for heat diffusion (Wang et al., 2016); v)a 3-layer snow scheme (Wang et al., 2013); vi)a spatially explicit observation-derived estimate for background albedo and optimized vegetation and snow albedo coefficients;; vii) a new reconstruction of global land cover history and wood harvest accounting following LUH2v2h maps (Hurtt et al., 2020) and PFT maps based the European Space Agency Climate Change Initiative Land Cover product (Poulter et al., 2015).

As in most TBMs, the vegetation is grouped into several plant functional types (PFTs), with 14 different types of vegetation plus bare soil in the case of ORCHIDEE v2.2. The original 13 PFTs are reported in (Krinner et al., 2005). Since ORCHIDEE v2.0 there are now two extra

PFTs included: C3 grasses are now split into three groups - tropical, temperate and boreal. The equations governing individual processes are generic with PFT specific parameters, except for the phenology models (see Appendix A in (MacBean et al., 2015)). In this study, ORCHIDEE is mainly used in a "grid-point mode" at each site location and forced with the corresponding local 30-minute gap-filled meteorological forcing data. Before performing the optimizations the modelled C stocks are brought to equilibrium in the spin-up phase by cycling the available site meteorological forcing over a long period (1300 years) with the default parameters of the model, which ensures a net carbon flux close to zero over annual-to-decadal time scales.

2.2 ORCHIDEE data assimilation system

The ORCHIDEE Data Assimilation System (ORCHIDAS) has been described in detail in previous studies (Bastrikov et al., 2018; Kuppel et al., 2014; MacBean et al., 2018; Peylin et al., 2016), and hence we only briefly define the method here. ORCHIDAS uses a variational data assimilation method to optimize the model parameters, accounting for uncertainties regarding the observations, the model, and the prior parameters. It relies on a Bayesian framework with the assumption of Gaussian errors, and the optimized parameters corresponds to the minimization of the following cost function *J(x)* (Tarantola, 2005):

$$\boldsymbol{J}(\boldsymbol{x}) = \frac{\mathbf{1}}{\mathbf{2}}[(\boldsymbol{H}(\boldsymbol{x}) - \boldsymbol{y})^{\boldsymbol{T}}.\boldsymbol{R}^{-\mathbf{1}}.(\boldsymbol{H}(\boldsymbol{x}) - \boldsymbol{y}) + (\boldsymbol{x} - \boldsymbol{x}^{\boldsymbol{b}})^{\boldsymbol{T}}.\boldsymbol{B}^{-\mathbf{1}}(\boldsymbol{x} - \boldsymbol{x}^{\boldsymbol{b}})] \qquad (1)$$

where x represents the parameters and *H(x)* the model contingent on the parameters, and y the observations. The cost function contains both the misfit between observations,and corresponding model outputs (first term on the right hand side of Eq. 1), and the misfit between a priori parameter values x_b and optimized parameters x *(second term on the right hand side of the Eq. 1 R* is the observation error covariance matrix (including measurement and model errors), and B is the prior parameter error covariance matrix. Both matrices (B and R) are diagonal since observation and model errors are assumed to be uncorrelated in space and time, and parameters are assumed to be independent. The cost function is iteratively minimized using the genetic algorithm (GA), which is a meta-heuristic optimization algorithm and follows the principles of genetics and natural selection (Goldberg et al., 1989; Haupt et al., 2004). The GA algorithm has been applied previously with ORCHIDAS tool and described in details by (Bastrikov et al., 2018). Briefly, the algorithm works iteratively and considers the vector of parameters as a chromosome and each parameter as a gene on that chromosome. The method fills a set of n chromosomes at every iteration, having the starting pool as a randomly perturbed parameter pool. The chromosomes at each subsequent iteration are chosen from randomly selected chromosomes of the previous iteration by either “crossover” or “mutation” process. (Santaren et al., 2014) showed that the performance of the algorithm is highly sensitive to its specific configuration and found the best configuration based on computational efficiency after testing different options. Here, we used the same configuration (i.e. number of chromosomes in the pool total number of parameters optimized; the number of iterations is 40; crossover/mutation ratio is 4:1; the number of gene blocks exchanged during crossover is 2 and the number of genes perturbed during mutation is 1) applied by (Santaren et al., 2014) and (Bastrikov et al., 2018). The algorithm does not assume prior knowledge of Gaussian PDFs for the observation and parameter uncertainties. Given we do not fully know the model uncertainty, we set the prior observation uncertainty as the RMSE between the model and the observations following (Kuppel et al., 2014). The prior parameter uncertainty is listed in Table S1.

The posterior error covariance matrix of the parameters (A) can be estimated by:

$$\boldsymbol{A} = [\boldsymbol{H}^{\boldsymbol{T}}\boldsymbol{R}^{-\mathbf{1}}\boldsymbol{H} + \boldsymbol{B}^{-\mathbf{1}}]^{-\mathbf{1}} \qquad (2)$$

This computes error correlations between parameters with the assumption of Gaussian prior errors and linearity of the model in the vicinity of the solution.

2.3 Flux measurements

We used twelve eddy covariance flux sites (measurement period ranging between 2003 and 2014) representing the various climate, elevation, and semi-arid ecosystems of SW US. Study sites are listed consecutively based on their mean annual C balance (Table 1). Flux tower instruments collect 30-minutes measurements of meteorological forcing data and eddy covariance measurements of net surface energy and carbon exchanges, which are available from the AmeriFlux data portal (http://ameriflux.lbl.gov). Meteorological forcing data included air temperature and surface pressure, precipitation, incoming long and shortwave radiation, wind speed, and specific humidity. To run the ORCHIDEE model, we partitioned the in-situ precipitation into rain and snowfall using a temperature threshold of 0°C. The site-level meteorological forcing data were gap filled utilizing downscaled and corrected ERA-Interim data following the approach of (Vuichard & Papale, 2015). Gross primary productivity (GPP) and the ecosystem respiration (R_{eco}) were estimated from the net ecosystem exchange (NEE) via the flux partitioning method described in (Biederman et al., 2016). We acknowledge that GPP and R_{eco} are not fully independent data with respect to NEE and are essentially model-derived estimates, but these concerns have been largely discussed in previous studies e.g., (Desai et al., 2008). Note that in this study, negative NEE refers to net CO_2 uptake into the ecosystem. In order to exclude the influence of the short-term variations in the fluxes on the model optimization, the daily averaged observations smoothed with a 15-day running mean have been used for the data assimilation.

2.4 Parameters optimized

The optimized parameters are described in Table S1 with their prior values, uncertainty and bounds for different plant functional types. Prior values are the default parameter values used in all non-optimized ORCHIDEE simulations. In the most past ORCHIDAS studies with previous versions of ORCHIDEE, only subsets of ORCHIDEE C cycle parameters have been optimized (Bastrikov et al., 2018; Kuppel et al., 2012, 2014; MacBean et al., 2015; MacBean et al., 2018; Santaren et al., 2007; Verbeeck et al., 2011). In this study, we considered all possible C cycle related ORCHIDEE parameters and hence contribute towards the correct representation of net and gross CO_2 fluxes. We selected all 102 parameters and divided them into four classes, controlling the main C cycle and plant physiological processes i.e. photosynthesis, conductance, phenology and post C uptake. This resulted in 31 parameters related to photosynthesis, 42 to phenology, 16 to post C uptake (C allocation, respiration, biomass and soil turnover), and 13 related to conductance. In a preliminary study, we tested at several SW US sites (US-Vcp, US-Mpj, US-Fuf, US-Wkg, US-Whs, US-Seg) the sensitivity of the ecosystem fluxes (NEE, GPP and R_{eco}) when optimizing all model parameters and when we just optimized subsets of the parameters related to each of the main processes. This test showed no significant optimization improvement by adding the conductance related parameters (results not shown here), and thus we did not include those parameters for all final optimizations presented in this study, leaving a total of 89 optimized parameters for each site. The parameter names, descriptions, bounds and prior uncertainties used in this study (Table S1) were derived from literature analysis, parameter databases and expert knowledge of the model equations. Documentation on the parameters can be accessed via ORCHIDEE webpage (https://forge.ipsl.jussieu.fr/orchidee/wiki/Documentation/OrchideeParameters, last access: 04 January 2021). The prior uncertainty is set to 40% of the range of variation for each parameter following previous ORCHIDAS studies (Kuppel et al., 2012; MacBean et al., 2015).

2.5 Assimilation Scenarios

We conducted several different optimizations to identify which processes (and their related parameters) are potentially causing model-data discrepancies (listed in Table 2). We grouped the optimizations based on various parameters set to optimize; therefore, we tested 7 assimilation scenarios (P1 – P7): P1 included all 89 parameters, whereas each consecutive scenario (P2 – P7) optimized different subsets of parameters related to each of the main C cycle processes (Table 2). The parameters that were not optimized were set to their default (prior) value. Comparing the P1 to P7 assimilation scenarios allows us to determine which sets of parameters (i.e. specific processes) are contributing most to the improvement in fluxes as a result of the parameter optimizations. See Table S1 for groupings of model parameters according to specific processes.

Table 2. Description of the different assimilation scenarios conducted in this study. The included parameter group(s) and numbers of parameters for each assimilation scenario are given. Parameters of each subgroup are listed in Table S1.

Optimization	Parameters included	Number of parameters
P1	All parameters (Phenology, Photosynthesis and Post C uptake)	~85
P2	Phenology and Photosynthesis	~70
P3	Phenology and Post C uptake	~50
P4	Photosynthesis and Post C uptake	~60
P5	Phenology only	~30
P6	Photosynthesis only	~45
P7	Post C uptake only	~15

2.6 Post-optimization analysis

We assessed the goodness of fit of the different assimilation scenarios by the mean square deviation (MSD) (in addition root mean squared error, RMSE or correlation coefficients, R or slope of linear least-square regression). Model evaluation metrics are presented in one of three ways: i) for each site; ii) grouped across all sites; and iii) sites grouped according to their mean net annual CO_2 flux characteristics across the time period identified in (Biederman et al., 2017). For the latter, the net CO_2 "sink" sites are US-Vcm, US-Vcp, US-Mpj, US-Fuf and US-Wjs; the "pivot" sites are US-Ses, US-Wkg, US-SRG, US-SRM, US-Whs, US-Seg; and the "source" site is US-Aud. We follow the approach of (Kobayashi & Salam, 2000) to quantify the differences between the simulations and observation in terms of bias, variance and phasing. We calculate the MSD between daily observation and each of the simulations and decompose it following the below equation:

$$\boldsymbol{MSD} = \frac{\mathbf{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{x_i} - \boldsymbol{y_i})^{\mathbf{2}} = (\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}})^{\mathbf{2}} + (\boldsymbol{\sigma_x} - \boldsymbol{\sigma_y})^{\mathbf{2}} + \mathbf{2}\boldsymbol{\sigma_x}\boldsymbol{\sigma_y}(\mathbf{1} - \boldsymbol{R}) \qquad (\mathbf{3})$$

where x is the model and y is the observations, σ is the standard deviation and R is the correlation coefficient.

The first term specifies the bias between model simulation and observation (squared). The second term "variance" measures their differences in terms of variability (i.e., the difference between the magnitude of the modeled and observed fluctuations). And the third term generally demonstrates the lack of correlation between them weighted by their standard deviations, which can be deemed as a measure of their disagreement in terms of phase (Bacour et al., 2019; Gauch et al., 2003). However, the second and third terms are partially dependent on each other (Bacour et al., 2019; Gauch et al., 2003). We further calculated the contribution of each component (bias, variance and phase) to the overall MSD by dividing each component by the total MSD.

3 Results

3.1 Impact of optimization of all parameters on model net and gross CO_2 fluxes

Across all sites, the prior ORCHIDEE simulations (i.e. before parameter optimization) fail to capture both the mean annual NEE and its IAV (**Figure 1a**) - as also seen for all TRENDY TBMs in MacBean et al (submitted). Across all sites, optimizing all C cycle-related parameters (phenology, photosynthesis and post C uptake - assimilation scenario P1) with NEE data dramatically increases the ability of the model to capture both the mean C source/sink behavior and the IAV (**Figure 1b**). C sink and source sites show significant improvement in terms of both mean annual NEE and IAV. There is not a strong bias in the model simulations at pivot sites whose mean annual NEE is close to zero; therefore, the optimization results in an improvement mainly in IAV (as represented by the correlation and slope values shown in inset figures in Figures 1 a and b). Improvement of the model-data fit resulting from the assimilation of NEE and optimizing all parameters (P1) is evident across all sites, with a reduction of daily NEE RMSE between 0.05 to 0.65 $gCm^{-2}d^{-1}$ (**Figure S1)**, with a similar reductions in daily GPP and R_{eco} RMSE (Table S2).Moreover, the temporal dynamics are well captured for all the sites: when optimizing all parameters, the median pearson correlation coefficients (R) increase by 0.45, 0.45, and 0.25 for daily, monthly and annual modeled NEE, respectively and posterior median slope values $\geq$0.5 (**Figure S2a and d)**. GPP temporal dynamics are also much improved by the P1 assimilation with a higher and tighter range in posterior R and slope values than NEE (**Figure S2b and e**). In contrast, there is less improvement in R_{eco} temporal dynamics although the median R and slope values are higher after the optimization (**Figure S2c and f**).

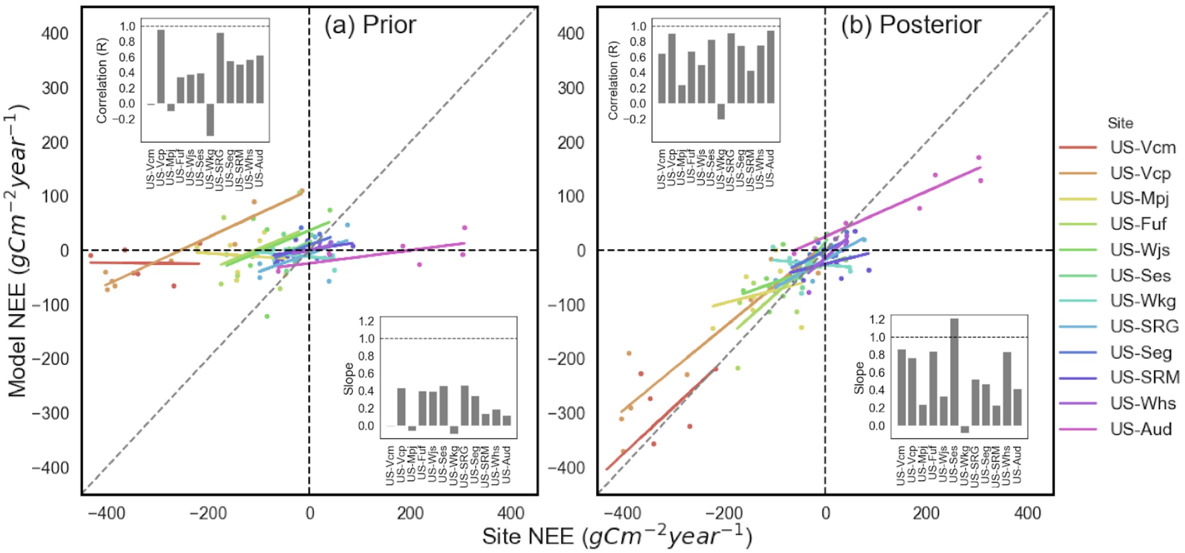

Figure 1. Comparison between modeled and observed annual NEE when assimilating NEE data and optimizing all phenology, photosynthesis and post C uptake parameters (P1) in the same assimilation. (**a**) Prior annual NEE simulation before parameter optimization, and (**b**) Posterior annual NEE after optimization. The trendline and slope value for the linear regression between the model and observations (bottom right inset figures) is shown for each site, together with their Pearson correlation coefficient, r (top left inset figures). The middle of the trend line should sit on the 1:1 line if the accurate mean annual source/sink behavior for a site is well captured by the model. A slope value close to or equal to 1 demonstrates the model is better at capturing the IAV. Colored points and trend lines represent all twelve sites, ordered from the largest mean sink (US-Vcm) to the largest mean source (US-Aud). The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the only source site is: US-Aud.

Across the majority of SW US dryland sites, the prior model simulates a depressed seasonal NEE amplitude and/or is unable to capture the observed bi-modal seasonality (**Figure 2**). The NEE amplitude and bi-modal seasonality generally improve when optimizing all parameters (P1 - blue curves in **Figure 2**), although the posterior simulations struggles to reach the exact magnitude of the spring and monsoon NEE troughs (net CO_2 uptake) for several sites (e.g. US-Mpj, US-Wjs, US-Ses, US-Seg, US-Wkg and US-Whs). Accurately capturing the seasonal peaks and troughs is important for NEE IAV because summer monsoon season fluxes are the dominant driver of NEE IAV (MacBean et al., submitted). While posterior seasonal NEE peaks and troughs are generally well captured, the assimilation of NEE alone often fails to capture the correct peaks in gross CO_2 fluxes (**Figure S3**), likely due to compensating errors in both GPP and R_{eco}. At the C source site (US-Aud) the model also fails to simulate the accurate peaks in springtime net carbon release (**Figure 2**). As also identified in MacBean et al (submitted), this is due to the fact that at US-Aud, TBMs tend to overestimate spring GPP and underestimate the earlier rise in spring R_{eco} (**Figure S3**). The optimization only partially corrects these model biases, suggesting that other missing processes may ultimately be responsible for the model-data misfit (such as disturbance following a fire that occurred at the site in 2002, which is not implemented in the current version of ORCHIDEE).

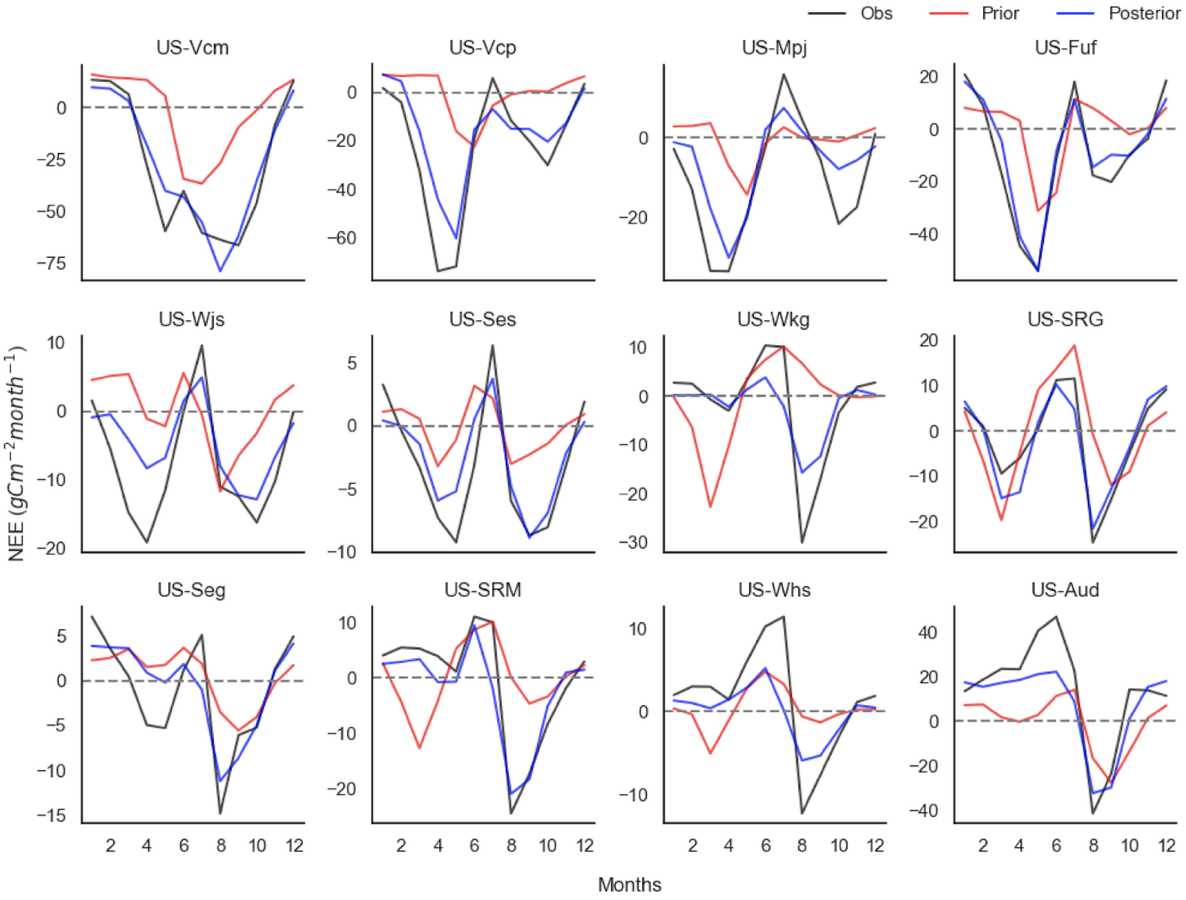

Figure 2. Mean monthly NEE seasonal cycles for each site comparing prior (red curve) and posterior (blue curve) ORCHIDEE simulations with observations (black curve). Posterior

simulation after assimilation of NEE data and optimization of all parameters: phenology, photosynthesis and post C uptake (P1). The sites are listed in order from largest mean annual C sink (US-Vcm) to mean annual C source (US-Aud).

Decomposing the daily NEE mean square deviations (MSD) between model and observations into bias, variance and phase components shows that across all sites, bias, variance and phase all contribute to prior NEE model-data discrepancies (**Figure 3a** left of vertical dashed line). The prior daily NEE MSD at the C sink sites are dominated by both phase and bias components (**Figure 3a** top panel). The fact that sink site MSD is also dominated by bias is unsurprising given that at those sites the prior model does not capture the mean annual C sink (**Figure 1a**). Note that, if we decompose the annual NEE MSD into the constituent bias, phase and variance components then bias overwhelmingly dominates the MSD at sink (and source) sites given their large underestimate of mean annual NEE (**Figure S4** top and bottom rows). In contrast, at the C pivot and source sites, the highest contribution to the prior daily NEE MSD is from the phase component (**Figure 3a** middle and bottom panel), indicating the default model does a poor job of representing the timing of dryland C cycle related processes. Across all sites, optimizing all parameters (P1) dramatically reduces the bias, variance and phase components of the daily NEE MSD, with phase remaining the strongest contributor to daily NEE MSD (**Figure 3a** right of dashed line).

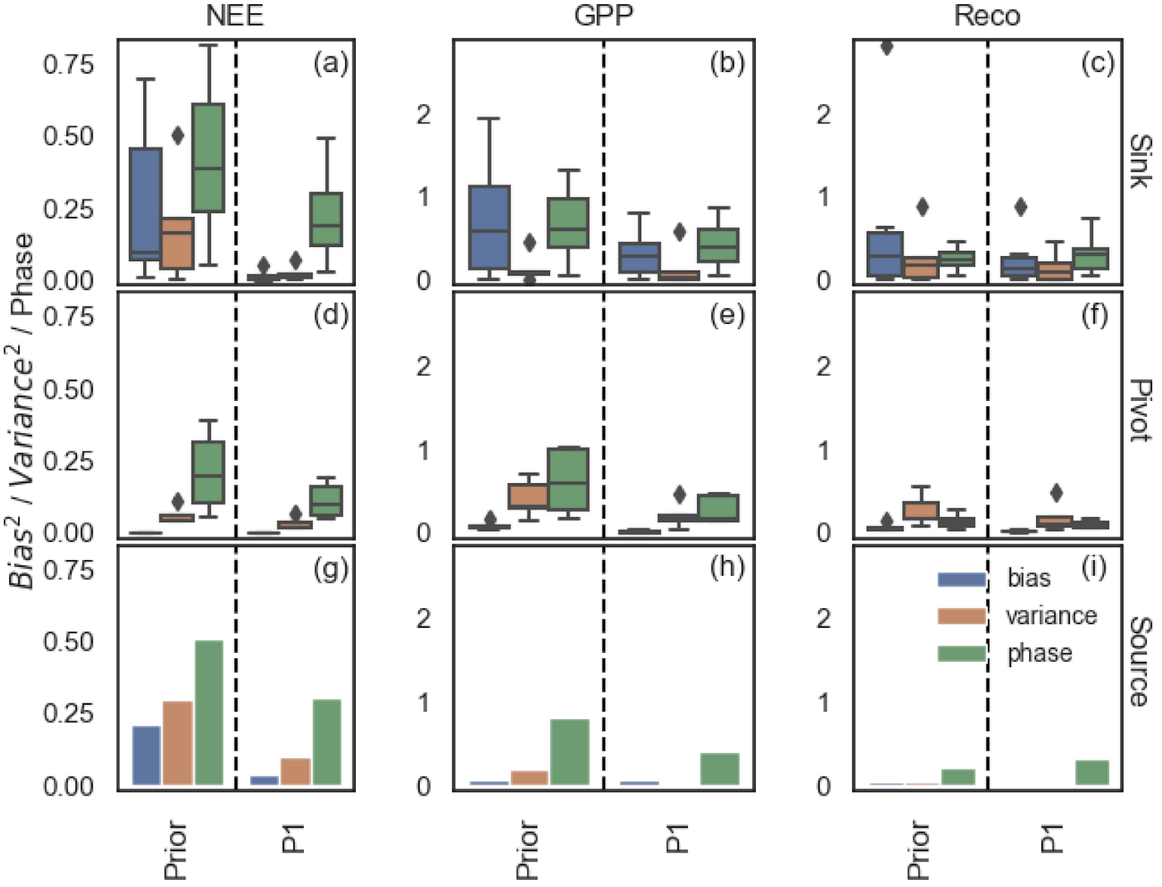

Figure 3. Daily NEE, GPP and R_{eco} mean square deviation (MSD) decomposition into bias, variance, and phase between simulations and observations for assimilating NEE observations and optimizing all phenology, photosynthesis and post C uptake parameters (P1). Blue, orange and green boxplots for bias, variance and phase components, respectively. Different rows separate the sites as sink (a-c), pivot (d-f) and source (g-i) based on total annual C flux. The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the source site ia: US-Aud. The x axes display the optimization scenarios (Prior and P1). The box whiskers show the spread of bias, variance and phase for all 12 sites considered in this study. The bias, variance and phase indicate the mean difference in flux magnitude, the mismatch in terms of flux fluctuation magnitude scales with the mean seasonal amplitude, and the seasonality in flux time series, respectively. Note that the y axis limits for both gross fluxes (GPP and R_{eco}) are the same.

As for the NEE, bias and phase are the dominant contributors to prior daily GPP MSD for the sink sites (left of vertical dashed line in **Figure 3b**), and phase only for the pivot and source sites (**Figures 3e and h**) For R_{eco}, a different MSD component is dominant depending on the mean C behavior of a site: bias dominates the prior daily R_{eco} MSD at the sink sites, variance at the pivot sites, and phase at the source sites (**Figures 3c, f and i**). Overall, assimilating NEE data in the P1 scenario reduces all gross CO_2 flux MSD components (right of dashed line in **Figure 3** middle and left columns), with phase remaining the strongest contributor to daily gross CO_2 flux MSD. However, unlike for the NEE, at the C sink sites phase *and* bias remain strong contributors to posterior GPP MSD (**Figure 3b**).

3.2 Impact of different processes (assimilation scenarios) on optimization results

Across all sites, modeled annual and seasonal NEE are improved the most in the P1 assimilation scenario compared to the other assimilation scenarios (P2 to P7), although all scenarios result in some improvement (**Figures S5, S6a and d,** and seasonal cycles in **Figure S7**). In general, there is less improvement in R_{eco} compared to NEE and GPP (**Figure S6**).

Comparing the MSD decomposition results for the various assimilation scenarios (P1-P7) can help to identify which processes may be causing the prior model-discrepancies in mean annual NEE and NEE IAV. At the source and sink sites, the bias component (blue bars in **Figure 4a and c**) is reduced dramatically by all optimization tests that include the post C uptake parameters related to C allocation, respiration, and biomass and soil C turnover (P1, P3, P4 and P7). For the sink sites, assimilation scenarios that also include photosynthesis parameters (P2 and P6) also result in a strong reduction in bias. This decrease in mean bias is also shown by the fact that the midpoints of the linear regression trendline between model and observations at forested sink sites (US-Vcm, US-Vcp, US-Mpj, and US-Fuf) and low-elevation source site (US-Aud) with optimization scenarios P1 to P4, P6 and P7 parameters all lie much closer to the 1:1 (grey dashed) line compared to P5 (**Figure S5**).

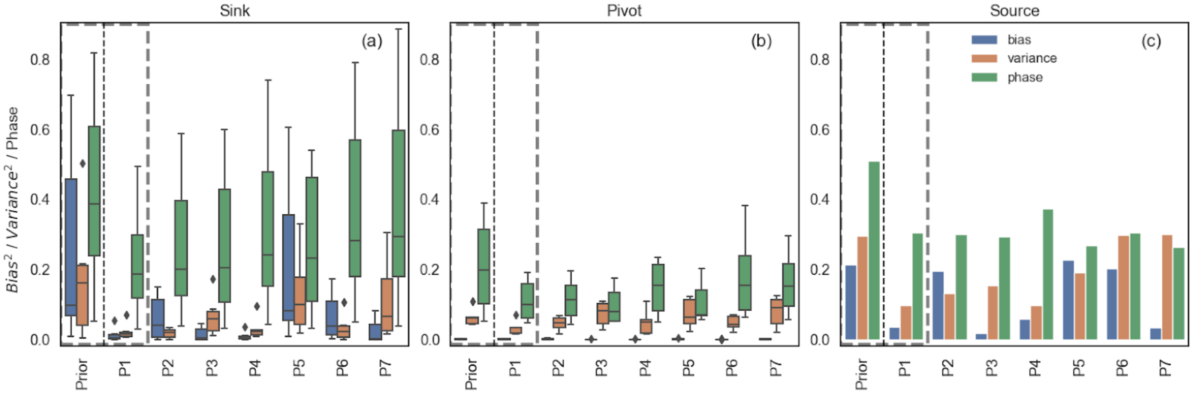

Figure 4. Daily NEE MSD decomposition into bias, variance, and phase components when assimilating NEE observations for different assimilation scenarios (P1-P7). Different panels separate the sites as sink (a), pivot (b) and source (c) based on total annual C flux. The C sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the C pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the C source site is: US-Aud. The grey dashed boxes highlight results repeated from **Figure 3(a,d,g)** to have better comparison of different process parameters side-by-side. The parameters included in each optimization are: P1: all parameters; P2: phenology and photosynthesis; P3: phenology and post C uptake; P4: photosynthesis and post C uptake; P5: phenology; P6: photosynthesis and P7: post C uptake. The boxplots show the median and interquartile range of the bias, variance and phase across all 12 sites considered in this study. US-Aud is the only C source site; therefore, the barplots in (c) show the bias, phase, and variance components of the MSD for that one site. The bias, variance and phase indicate the mean difference in flux magnitude, the difference in the magnitude of flux variations, and the difference in the correlations weighted by the standard deviations, respectively (see Methods).

Across all sites the difference in phase between the model and observations (green bars in **Figure 4**), which, as already noted, is the largest contribution to the prior NEE MSD across all sites, is mostly reduced by assimilation scenarios that include phenology parameters (i.e. P1, P2, P3 and P5). However, the phase component is not reduced as much as the bias in any of the assimilation scenarios; thus, for all sites and all assimilation scenarios the phase remains the largest component of the posterior daily NEE MSD (**Figure 4**). Including parameters related to photosynthesis or post C uptake with the phenology parameters (i.e. assimilation scenarios P2 and P3) helps to slightly reduce the phase discrepancy at sink sites compared with phenology parameters alone (P5) (as seen above for the improvement in slope values at the sink sites). Examining the spread in slope and R values across all sites, we see that the annual variability (median slope and R values) is improved the most for assimilation scenarios with at least two parameter sets (P1 to P4 - **Figure S6a and d**). The persistence of phase as the dominant component of the posterior daily NEE suggests further model improvement in dryland phenology models is needed before TBMs can correctly reproduce NEE seasonality and IAV.

The variance component of the daily NEE MSD (orange bars in **Figure 4**), which also shows a modest contribution to daily NEE MSD at the sink and source sites, is mostly reduced at the sink sites with assimilation scenarios that include photosynthesis parameters (i.e. P1, P2, P4 and P6). At US-Aud the variance component was reduced most by assimilation scenarios that included two or more sets of parameters (i.e. P1 - P4) (**Figure 4c**).

While the post C uptake parameters are key for reducing bias in forested sink site NEE, biases in GPP and R_{eco} at these sites are reduced by optimizing photosynthesis parameters (P1, P2, P4, and P6 - blue boxes **Figure S8b and c**). The GPP and R_{eco} bias components at the sink sites are not reduced as strongly as NEE biases for any assimilation scenario; thus, bias remains a key contributor to posterior gross CO_2 flux MSD. Similarly to NEE, parameter subsets that include phenology parameters (P1, P2, P3 and P5) are key for reducing the daily GPP MSD phase component at pivot sites (green boxes in **Figure S8e**; however, in contrast with the NEE results, at sink sites the GPP phase component tends to be reduced by all assimilation scenarios except P7 (see also median GPP slope and R values in **Figures S6b and e**). With the exception of P1 and P2 for GPP, the GPP and R_{eco} variance components are not reduced much by any of the assimilation scenarios and remain a considerable component of the MSD for both GPP and R_{eco} at the pivot sites, and for R_{eco} at the sink sites (**Figures S8b,c,e,f**). We note that the GPP and R_{eco} reductions in MSD components tend to be similar, suggesting model-deficiencies in R_{eco} are mainly influenced by those in GPP. Addressing GPP model-data deficiencies is therefore a high priority.

3.3 Constraint on parameters

For all assimilation scenarios, we found significant parameter deviations from prior values for numerous phenology, photosynthesis and post C uptake related parameters (**Figure 5**), which is consistent with the fact that we find all parameter subsets are needed to improve model mean annual NEE and IAV. We do not find that parameters deviate more, or the uncertainty reduction much different, when only one subset or two parameter subsets are included in the optimization instead of all three (e.g. cf. P2 with P1), although posterior values are different for each assimilation scenario (**Figure 5**). In particular, most of the post C uptake parameters deviate strongly from the prior median deviations (>20% of total parameter bound). There are also significant uncertainty reductions (>50%) for most of the parameters which show strong deviations from their prior value: 10 for phenology (out of 42), 7 for photosynthesis (out of 31) and 7 for post C uptake (out of 16) (**Figure 5**). The error correlations between the estimated parameters are usually minimal except between post C uptake parameters (see example for one site in **Figure S9**).

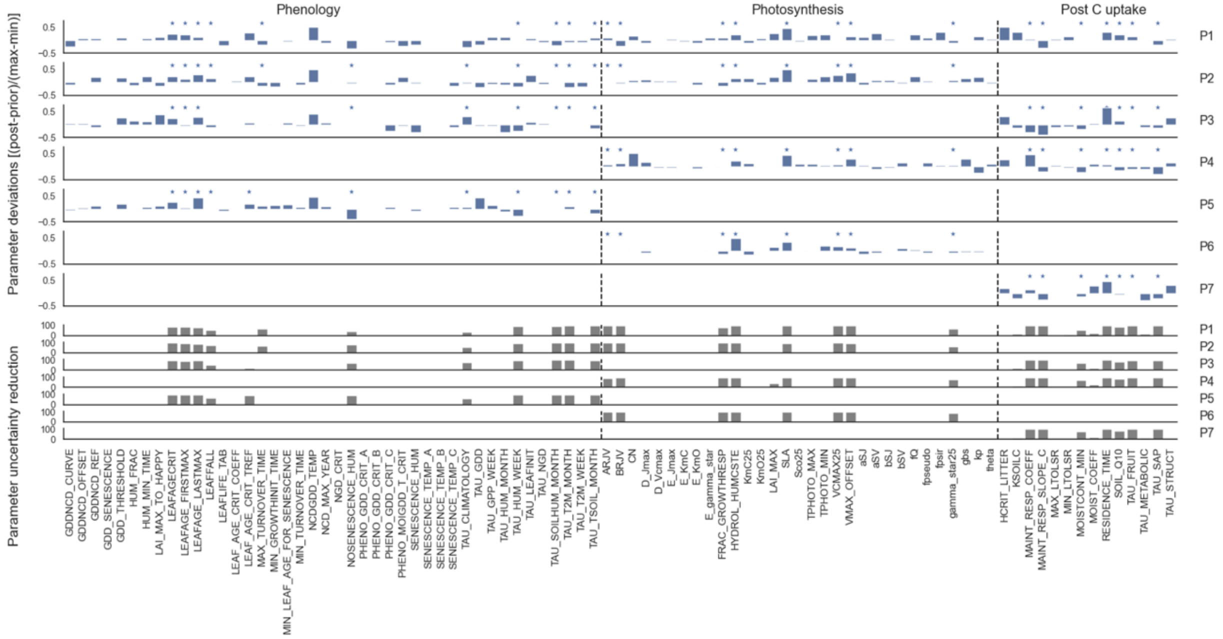

Figure 5. Optimized median parameter deviations [(posterior - prior) / (max - min)] (blue bars) and associated median parameter uncertainty reductions (grey bars) for all parameters controlling phenology, photosynthesis and post C uptake assimilating NEE data (P1-P7). Bars represent the median across all 12 sites. The asterisks above blue bars indicate the parameters that have larger than 50% uncertainty reduction. Each line corresponds to a specific optimization test (shown on the right axis). The parameters are given on the bottom axis. The vertical dashed lines separate

the different parameter subsets (phenology, photosynthesis and post C uptake). **Table S1** details the prior and posterior parameter values and their uncertainty for all parameters together with the maximum and minimum bounds used in the optimizations.

Looking at the individual parameter plots for the P1 assimilation scenario (**Figure S10)**, we find that at some sites several posterior parameters are “edge-hitting” (e.g. soil Q10). Given we chose weak prior constraints (wide bounds) in all optimizations, the fact that some posterior parameters are hitting their bounds suggests that the optimization may be aliasing model structural error onto the parameters (MacBean et al., 2016) and/or that the model cannot improve further via parameter optimization. Instead, further model developments are likely needed to address structural uncertainties and missing processes. Hypotheses as to what (missing) processes might be responsible for these remaining model errors have been discussed in MacBean et al (submitted). For example, it is possible that model inability to capture peak GPP fluxes for some sites (see Section 3.1) is due to too weak a response of model vegetation growth to pulses of moisture availability, an issue which is not fully alleviated by using only NEE to optimize model parameters.

4 Conclusions

While improvement in modeled NEE is best when all parameters are optimized, our assimilation scenario experiments have shown that post C uptake and phenology parameters are key for correcting model underestimates in dryland mean annual NEE and IAV, respectively. Remaining model-data discrepancies indicate that additional model developments are needed. . The specific DA configuration (e.g. the number of parameters, and the processes with which they are related) can lead to different posterior values and degree of improvement in model-data fit. Therefore, further tests of different DA configurations and optimizations at other locations are needed to explore the potential of Bayesian DA systems for quantifying and reducing error in dryland ecosystem C fluxes. However, our study demonstrates that dryland C cycle parameters and processes in TBMs need optimizing and/or further development before we can reliably use these models to accurately simulate dryland ecosystem contributions to global C cycle variability.

Acknowledgments

This work was supported by the Department of Geography at Indiana University.

Code availability

The ORCHIDEE model is under a free software license (CeCILL; see http://www.cecill.info/index.en.html) and the source code is visible here: https://forge.ipsl.jussieu.fr/orchidee/browser/tags/ORCHIDEE. The ORCHIDEE model code is written in Fortran 90 and is maintained and developed under an SVN version control system at the Institute Pierre Simon Laplace (IPSL) in France. The ORCHIDAS code is currently in the process of being put on a GitHub repository but for now it is available on request to vladislav.bastrikov@lsce.ipsl.fr.

Data availability

Meteorological forcing data and eddy covariance measurements of net surface energy and carbon exchanges at 30-minutes intervals are available from the AmeriFlux data portal (http://ameriflux.lbl.gov). The model outputs from ORCHIDEE simulations and post-processing python scripts for manuscript figures and tables are freely available in a Git repository (https://github.com/kashifmahmud/SW_US_semiarid).

References

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., Jain, A. K., Kato, E., Poulter, B., Sitch, S., Stocker, B. D., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S., & Zeng, N. (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.

Anderson-Teixeira, K. J., Delong, J. P., Fox, A. M., Brese, D. A., & Litvak, M. E. (2011). Differential responses of production and respiration to temperature and moisture drive the carbon balance across a climatic gradient in New Mexico. In *Global Change Biology* (Vol. 17, Issue 1, pp. 410–424). https://doi.org/10.1111/j.1365-2486.2010.02269.x

Bacour, C., Maignan, F., Peylin, P., MacBean, N., Bastrikov, V., Joiner, J., Köhler, P., Guanter, L., & Frankenberg, C. (2019). Differences Between OCO-2 and GOME-2 SIF Products From a Model-Data Fusion Perspective. In *Journal of Geophysical Research: Biogeosciences* (Vol. 124, Issue 10, pp. 3143–3157). https://doi.org/10.1029/2018jg004938

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). In *Geoscientific Model Development* (Vol. 11, Issue 12, pp. 4739–4754). https://doi.org/10.5194/gmd-11-4739-2018

Biederman, J. A., Scott, R. L., Bell, T. W., Bowling, D. R., Dore, S., Garatuza-Payan, J., Kolb, T. E., Krishnan, P., Krofcheck, D. J., Litvak, M. E., Maurer, G. E., Meyers, T. P., Oechel, W. C., Papuga, S. A., Ponce-Campos, G. E., Rodriguez, J. C., Smith, W. K., Vargas, R., Watts, C. J., … Goulden, M. L. (2017). CO exchange and evapotranspiration across dryland ecosystems of southwestern North America. *Global Change Biology*, *23*(10), 4204–4221.

Biederman, J. A., Scott, R. L., Goulden, M. L., Vargas, R., Litvak, M. E., Kolb, T. E., Yepez, E. A., Oechel, W. C., Blanken, P. D., Bell, T. W., Garatuza-Payan, J., Maurer, G. E., Dore, S., & Burns, S. P. (2016). Terrestrial carbon balance in a drier world: the effects of water availability in southwestern North America. In *Global Change Biology* (Vol. 22, Issue 5, pp. 1867–1879). https://doi.org/10.1111/gcb.13222

Cleverly, J., Eamus, D., Luo, Q., Coupe, N. R., Kljun, N., Ma, X., Ewenz, C., Li, L., Yu, Q., & Huete, A. (2016). The importance of interacting climate modes on Australia's contribution to global carbon cycle extremes. In *Scientific Reports* (Vol. 6, Issue 1). https://doi.org/10.1038/srep23113

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., & Luke, C. M. (2013). Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, *494*(7437), 341–344.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M., & Stauch, V. J. (2008). Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. In *Agricultural and Forest Meteorology* (Vol. 148, Issues 6-7, pp. 821–838). https://doi.org/10.1016/j.agrformet.2007.11.012

Dore, S., Montes-Helu, M., Hart, S. C., Hungate, B. A., Koch, G. W., Moon, J. B., Finkral, A. J., & Kolb, T. E. (2012). Recovery of ponderosa pine ecosystem carbon and water fluxes from thinning and stand-replacing fire. *Global Change Biology*, *18*(10), 3171–3185.

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., … Vuichard, N. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. In *Climate Dynamics* (Vol. 40, Issues 9-10, pp. 2123–2165). https://doi.org/10.1007/s00382-012-1636-1

Fu, Z., Dong, J., Zhou, Y., Stoy, P. C., & Niu, S. (2017). Long term trend and interannual variability of land carbon uptake—the attribution and processes. In *Environmental Research Letters* (Vol. 12, Issue 1, p. 014018). https://doi.org/10.1088/1748-9326/aa5685

Gauch, H. G., Gene Hwang, J. T., & Fick, G. W. (2003). Model Evaluation by Comparison of Model-Based Predictions and Measured Values. In *Agronomy Journal* (Vol. 95, Issue 6, pp. 1442–1446). https://doi.org/10.2134/agronj2003.1442

Goldberg, D. E., David Edward, G., Goldberg, D. E. G., & Visiting Assistant Professor of History David E Goldberg. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Company.

Haupt, R. L., Haupt, S. E., & Haupt, S. E. A. (2004). *Practical Genetic Algorithms*. Wiley.

Haverd, V., Ahlström, A., Smith, B., & Canadell, J. G. (2017). Carbon cycle responses of semi-arid ecosystems to positive asymmetry in rainfall. *Global Change Biology*, *23*(2), 793–800.

Haverd, V., Raupach, M. R., Briggs, P. R., Canadell, J. G., Isaac, P., Pickett-Heaps, C., Roxburgh, S. H., van Gorsel, E., Viscarra Rossel, R. A., & Wang, Z. (2013a). Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles. In *Biogeosciences* (Vol. 10, Issue 3, pp. 2011–2040). https://doi.org/10.5194/bg-10-2011-2013

Haverd, V., Raupach, M. R., Briggs, P. R., Canadell, J. G., Isaac, P., Pickett-Heaps, C., Roxburgh, S. H., van Gorsel, E., Viscarra Rossel, R. A., & Wang, Z. (2013b). Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles. In *Biogeosciences* (Vol. 10, Issue 3, pp. 2011–2040). https://doi.org/10.5194/bg-10-2011-2013

Hogue, T. S., Bastidas, L., Gupta, H., Sorooshian, S., Mitchell, K., & Emmerich, W. (2005). Evaluation and Transferability of the Noah Land Surface Model in Semiarid Environments. In *Journal of Hydrometeorology* (Vol. 6, Issue 1, pp. 68–84). https://doi.org/10.1175/jhm-402.1

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., … Zhang, X. (2020). Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geoscientific Model Development*, *13*(11), 5425–5464.

Kobayashi, K., & Salam, M. U. (2000). Comparing Simulated and Measured Values Using Mean Squared Deviation and its Components. In *Agronomy Journal* (Vol. 92, Issue 2, p. 345). https://doi.org/10.1007/s100870050043

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., & Colin Prentice, I. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. In *Global Biogeochemical Cycles* (Vol. 19, Issue 1). https://doi.org/10.1029/2003gb002199

Krishnan, P., Meyers, T. P., Scott, R. L., Kennedy, L., & Heuer, M. (2012). Energy exchange and evapotranspiration over two temperate semi-arid grasslands in North America. In *Agricultural and Forest Meteorology* (Vol. 153, pp. 31–44). https://doi.org/10.1016/j.agrformet.2011.09.017

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. In *Biogeosciences* (Vol. 9, Issue 10, pp. 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. In *Geoscientific Model Development* (Vol. 7, Issue 6, pp. 2581–2597). https://doi.org/10.5194/gmd-7-2581-2014

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., Köhler, P., Gómez-Dans, J., & Disney, M. (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. In *Biogeosciences* (Vol. 12, Issue 23, pp. 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. In *Geoscientific Model Development* (Vol. 9, Issue 10, pp. 3569–3588). https://doi.org/10.5194/gmd-9-3569-2016

MacBean, N., R. L. Scott, J. A. Biederman, P. Peylin, T. Kolb, M. Litvak, P. Krishnan, T. Meyers, V. Arora, V. Bastrikov, D. Goll, D. L. Lombardozzi, J. Nabel, J. Pongratz, S. Sitch, A. P. Walker, S. Zaehle, and D. J. P. Moore. (n.d.). Dynamic Global Vegetation Models Underestimate Net CO2 Flux Mean and Inter-Annual Variability in Semiarid Ecosystems. *Submitted. Preprint Available upon Request.*

MacBean, N., Scott, R. L., Biederman, J. A., Ottlé, C., Vuichard, N., Ducharne, A., Kolb, T., Dore, S., Litvak, M., & Moore, D. J. P. (2020). Testing water fluxes and storage from two hydrology configurations within the ORCHIDEE land surface model across US semi-arid sites. In *Hydrology and Earth System Sciences* (Vol. 24, Issue 11, pp. 5203–5230). https://doi.org/10.5194/hess-24-5203-2020

Peng, S., Ciais, P., Chevallier, F., Peylin, P., Cadule, P., Sitch, S., Piao, S., Ahlström, A., Huntingford, C., Levy, P., Li, X., Liu, Y., Lomas, M., Poulter, B., Viovy, N., Wang, T., Wang, X., Zaehle, S., Zeng, N., Zhao, H. (2015). Benchmarking the seasonal cycle of CO2fluxes simulated by terrestrial ecosystem models. In *Global Biogeochemical Cycles* (Vol. 29, Issue 1, pp. 46–64). https://doi.org/10.1002/2014gb004931

Petrie, M. D., Collins, S. L., Swann, A. M., Ford, P. L., & Litvak, M. E. (2015). Grassland to shrubland state transitions enhance carbon sequestration in the northern Chihuahuan Desert. *Global Change Biology*, *21*(3), 1226–1235.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., & Prunet, P. (2016). A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle. In *Geoscientific Model Development* (Vol. 9, Issue 9, pp. 3321–3346). https://doi.org/10.5194/gmd-9-3321-2016

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P. E., Li, J., Lin, X., Lomas, M. R., Lu, M., Luo, Y., … Zeng, N. (2013). Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends. *Global Change Biology*, *19*(7), 2117–2132.

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y., Running, S. W., Sitch, S., & van der Werf, G. R. (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., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., & Peylin, P. (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.

Raczka, B. M., Davis, K. J., Huntzinger, D., Neilson, R. P., Poulter, B., Richardson, A. D., Xiao, J., Baker, I., Ciais, P., Keenan, T. F., Law, B., Post, W. M., Ricciuto, D., Schaefer, K., Tian, H., Tomelleri, E., Verbeeck, H., & Viovy, N. (2013). Evaluation of continental carbon cycle simulations with North American flux tower observations. In *Ecological Monographs* (Vol. 83, Issue 4, pp. 531–556). https://doi.org/10.1890/12-0893.1

Renwick, K. M., Fellows, A., Flerchinger, G. N., Lohse, K. A., Clark, P. E., Smith, W. K., Emmett, K., & Poulter, B. (2019). Modeling phenological controls on carbon dynamics in dryland sagebrush ecosystems. In *Agricultural and Forest Meteorology* (Vol. 274, pp. 85–94). https://doi.org/10.1016/j.agrformet.2019.04.003

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. In *Biogeosciences* (Vol. 11, Issue 24, pp. 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. In *Global Biogeochemical Cycles* (Vol. 21, Issue 2). https://doi.org/10.1029/2006gb002834

Schaefer, K., Schwalm, C. R., Williams, C., Altaf Arain, M., Barr, A., Chen, J. M., Davis, K. J., Dimitrov, D., Hilton, T. W., Hollinger, D. Y., Humphreys, E., Poulter, B., Raczka, B. M., Richardson, A. D., Sahoo, A., Thornton, P., Vargas, R., Verbeeck, H., Anderson, R., … Zhou, X. (2012). A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. In *Journal of Geophysical Research: Biogeosciences* (Vol. 117, Issue G3). https://doi.org/10.1029/2012jg001960

Scott, R. L. (2010). Using watershed water balance to evaluate the accuracy of eddy covariance evaporation measurements for three semiarid ecosystems. In *Agricultural and Forest Meteorology* (Vol. 150, Issue 2, pp. 219–225). https://doi.org/10.1016/j.agrformet.2009.11.002

Scott, R. L., Biederman, J. A., Hamerlynck, E. P., & Barron-Gafford, G. A. (2015). The carbon balance pivot point of southwestern U.S. semiarid ecosystems: Insights from the 21st century drought. In *Journal of Geophysical Research: Biogeosciences* (Vol. 120, Issue 12, pp. 2612–2624). https://doi.org/10.1002/2015jg003181

Scott, R. L., Darrel Jenerette, G., Potts, D. L., & Huxman, T. E. (2009). Effects of seasonal drought on net carbon dioxide exchange from a woody-plant-encroached semiarid grassland. In *Journal of Geophysical Research* (Vol. 114, Issue G4). https://doi.org/10.1029/2008jg000900

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. https://doi.org/10.1137/1.9780898717921

Traore, A., Ciais, P., Vuichard, N., MacBean, N., Dardel, C., Poulter, B., Piao, S., Fisher, J., Viovy, N., Jung, M., & Myneni, R. (2014). 1982–2010 Trends of Light Use Efficiency and

Inherent Water Use Efficiency in African vegetation: Sensitivity to Climate and Atmospheric CO2 Concentrations. In *Remote Sensing* (Vol. 6, Issue 9, pp. 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. In *Biogeosciences* (Vol. 13, Issue 23, pp. 6363–6383). https://doi.org/10.5194/bg-13-6363-2016

Unland, H. E., Houser, P. R., Shuttleworth, W. J., & Yang, Z.-L. (1996). Surface flux measurement and modeling at a semi-arid Sonoran Desert site. In *Agricultural and Forest Meteorology* (Vol. 82, Issues 1-4, pp. 119–153). https://doi.org/10.1016/0168-1923(96)02330-1

Verbeeck, H., Peylin, P., Bacour, C., Bonal, D., Steppe, K., & Ciais, P. (2011). Seasonal patterns of CO2fluxes in Amazon forests: Fusion of eddy covariance data and the ORCHIDEE model. In *Journal of Geophysical Research* (Vol. 116, Issue G2). https://doi.org/10.1029/2010jg001544

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V., & Peylin, P. (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. In *Geoscientific Model Development* (Vol. 12, Issue 11, pp. 4751–4779). https://doi.org/10.5194/gmd-12-4751-2019

Vuichard, N., & Papale, D. (2015). Filling the gaps in meteorological continuous data measured at FLUXNET sites with ERA-Interim reanalysis. In *Earth System Science Data* (Vol. 7, Issue 2, pp. 157–171). https://doi.org/10.5194/essd-7-157-2015

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.

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S., & Peng, S. (2013). Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model: ORCHIDEE SNOW MODEL EVALUATION. *Journal of Geophysical Research*, *118*(12), 6064–6079.

Whitley, R., Beringer, J., Hutley, L. B., Abramowitz, G., De Kauwe, M. G., Duursma, R., Evans, B., Haverd, V., Li, L., Ryu, Y., Smith, B., Wang, Y.-P., Williams, M., & Yu, Q. (2016a). A model inter-comparison study to examine limiting factors in modelling Australian tropical savannas. In *Biogeosciences* (Vol. 13, Issue 11, pp. 3245–3265). https://doi.org/10.5194/bg-13-3245-2016

Whitley, R., Beringer, J., Hutley, L. B., Abramowitz, G., De Kauwe, M. G., Duursma, R., Evans, B., Haverd, V., Li, L., Ryu, Y., Smith, B., Wang, Y.-P., Williams, M., & Yu, Q. (2016b). A model inter-comparison study to examine limiting factors in modelling Australian tropical savannas. In *Biogeosciences* (Vol. 13, Issue 11, pp. 3245–3265). https://doi.org/10.5194/bg-13-3245-2016

Yin, X., and Struik, P. C.: C3 and C4 photosynthesis models: An overview from the perspective of crop modelling, NJAS - Wageningen Journal of Life Sciences, 57, 27-38, https://doi.org/10.1016/j.njas.2009.07.001, 2009.

AGU Advances

Supporting Information for

Optimizing Phenology Parameters Drastically Improves Terrestrial Biosphere Model Underestimates of Dryland Net CO_2 Flux Inter-Annual Variability

K. Mahmud[1], J. A. Biederman[2], R. L. Scott[2], M. E. Litvak[3], T. Kolb[4], T. P. Meyers[5], P. Krishnan[5,6], V. Bastrikov[7,8], and N. MacBean[1]

[1]Department of Geography, Indiana University, Bloomington, IN 47405, USA

[2]Southwest Watershed Research Center, United States Department of Agriculture, Agricultural Research Service, Tucson, AZ 85719, USA

[3]Department of Biology, University of New Mexico, Albuquerque, NM, 87131, USA

[4]School of Forestry, Northern Arizona University, Flagstaff, AZ, 86011, USA

[5]NOAA/ARL Atmospheric Turbulence and Diffusion Division, Oak Ridge, TN, 37830, USA

[6]Oak Ridge Associated Universities, Oak Ridge, TN, 37830, USA

[7]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, F-91191, France

[8]Now at: Science Partners, Paris, 75010, France

Contents of this file

Tables S1 to S2
Figures S1 to S10

Parameter Name	Description (Unit)	Plant functional type			
		TeNE	TeBE	TeBD	GC4
Photosynthesis parameters					
ARJV	a coefficient of the linear regression (a+bT) defining the Jmax25/Vcmax25 ratio (mu mol e- (mu mol CO_2)-1)	2.59 ± 0.4 2, 3	2.59 ± 0.4 2, 3	2.59 ± 0.4 2, 3	1.715 ± 0.48 1, 2.2
aSJ	a coefficient of the linear regression (a+bT) defining the Entropy term for Jmax (J K-1 mol-1)	659.7 ± 264 330, 990	659.7 ± 264 330, 990	659.7 ± 264 330, 990	630 ± 252 315, 945
aSV	a coefficient of the linear regression (a+bT) defining the Entropy term for Vcmax (J K-1 mol-1)	668.39 ± 267.6 334, 1003	668.39 ± 267.6 334, 1003	668.39 ± 267.6 334, 1003	641.64 ± 256.4 321, 962
BRJV	b coefficient of the linear regression (a+bT) defining the Jmax25/Vcmax25 ratio (mu mol e- (mu mol CO_2)-1)	-0.035 ± 0.028 -0.07, 0	-0.035 ± 0.028 -0.07, 0	-0.035 ± 0.028 -0.07, 0	-0.01 ± 0.028 -0.035, 0.035
bSJ	b coefficient of the linear regression (a+bT) defining the Entropy term for Jmax (J K-1 mol-1 C-1)	-0.75 ± 0.6 -1.5, 0	-0.75 ± 0.6 -1.5, 0	-0.75 ± 0.6 -1.5, 0	0.01 ± 0.6 -0.75, 0.75
bSV	b coefficient of the linear regression (a+bT) defining the Entropy term for Vcmax (J K-1 mol-1 C-1)	-1.07 ± 0.8 -2, 0	-1.07 ± 0.8 -2, 0	-1.07 ± 0.8 -2, 0	0.1 ± 0.856 -1.07, 1.07
CN	C/N ratio	40 ± 32 20, 100	40 ± 32 20, 100	40 ± 32 20, 100	-
D_Jmax	Energy of deactivation for Jmax (J/mol)	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	192000 ± 15200 173000, 211000
D_Vcmax	Energy of deactivation for Vcmax (J/mol)	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	192000 ± 15200 173000, 211000
E_gamma_star	Energy of activation for gamma_star (J mol-1)	37830 ± 8000 27830, 47830	37830 ± 8000 27830, 47830	37830 ± 8000 27830, 47830	37830 ± 8000 27830, 47830
E_Jmax	Energy of activation for Jmax (J mol-1)	49884 ± 8000 39884, 59884	49884 ± 8000 39884, 59884	49884 ± 8000 39884, 59884	77900 ± 8000 67900, 87900
E_KmC	Energy of activation for KmC (J mol-1)	79430 ± 8000 69430, 89430	79430 ± 8000 69430, 89430	79430 ± 8000 69430, 89430	79430 ± 8000 69430, 89430
E_KmO	Energy of activation for KmO (J mol-1)	36380 ± 8000 26380, 46380	36380 ± 8000 26380, 46380	36380 ± 8000 26380, 46380	36380 ± 8000 26380, 46380
fpseudo	Fraction of electrons at PSI that follow pseudocyclic transport	-	-	-	0.1 ± 0.032 0.06, 0.14
fpsir	Fraction of PSII e−transport rate partitioned to the C4 cycle	-	-	-	0.4 ± 0.16 0.4, 0.6
FRAC_GROWTHRESP	Fraction of GPP which is lost as growth respiration	0.28 ± 0.064 0.2, 0.36	0.28 ± 0.064 0.2, 0.36	0.28 ± 0.064 0.2, 0.36	0.28 ± 0.064 0.2, 0.36

fQ	Fraction of electrons at reduced plastoquinone that follow the Q-cycle	-	-	-	1 ± 0.24 0.7, 1.3
gamma_star25	Ci-based CO_2 compensation point in the absence of Rd at 25C (ubar)	42.75 ± 8 22.75, 62.75	42.75 ± 8 22.75, 62.75	42.75 ± 8 22.75, 62.75	42.75 ± 8 22.75, 62.75
gbs	Bundle-sheath conductance (mol m-2 s-1 bar-1)	-	-	-	0.003 ± 0.0008 0.001, 0.005
HYDROL_HUMCSTE	Root profile (m) in empirical plant water stress function calculation	1 ± 1.5 0.25, 4	0.8 ± 1.12 0.2, 3	0.8 ± 1.12 0.2, 3	1 ± 1.5 0.25, 4
KmC25	Michaelis-Menten constant of Rubisco for CO_2 at 25C (ubar)	404.9 ± 160 204.9, 604.9	404.9 ± 160 204.9, 604.9	404.9 ± 160 204.9, 604.9	650 ± 160 450, 850
KmO25	Michaelis-Menten constant of Rubisco for O_2 at 25C (ubar)	278400 ± 80000 178400, 378400	278400 ± 80000 178400, 378400	278400 ± 80000 178400, 378400	450000 ± 80000 350000, 550000
kp	Initial carboxylation efficiency of the PEP carboxylase (mol m-2 s-1 bar-1)	-	-	-	0.7 ± 0.24 0.4, 1
LAI_MAX	Maximum LAI (m2/m2)	5 ± 2 3, 8	5 ± 2 3, 8	5 ± 2 3, 8	2.5 ± 0.8 4, 10
Sco25	Relative CO_2 /O_2 specificity factor for Rubisco at 25C (bar bar^{-1})	2800 ± 800 1800, 3800	2800 ± 800 1800, 3800	2800 ± 800 1800, 3800	2590 ± 800 1590, 3590
SLA	Specific leaf area (m2/gC)	0.00926 ± 0.005 0.004, 0.02	0.02 ± 0.012 0.01, 0.04	0.026 ± 0.0148 0.013, 0.05	0.026 ± 0.0148 0.013, 0.05
theta	Convexity factor for response of J to irradiance	0.7 ± 0.18 0.5, 0.95	0.7 ± 0.18 0.5, 0.95	0.7 ± 0.18 0.5, 0.95	0.7 ± 0.18 0.5, 0.95
TPHOTO_MAX	Maximum photosynthesis temperature (deg C)	55 ± 4 50, 60	55 ± 4 50, 60	55 ± 4 50, 60	55 ± 4 50, 60
TPHOTO_MIN	Minimum photosynthesis temperature (deg C)	-4 ± 4 -9, 1	-4 ± 4 -9, 1	-4 ± 4 -9, 1	-4 ± 4 -9, 1
VCMAX25	Maximum rate of Rubisco activity-limited carboxylation at 25C (micromol/m2/s)	35 ± 10 19, 51	45 ± 16 25, 65	55 ± 20 30, 80	70 ± 25.6 38, 102
VMAX_OFFSET	Offset (minimum relative vcmax)	0.3 ± 0.048 0.24, 36			
Post C uptake parameters - autotrophic and heterotrophic respiration, C allocation, biomass and soil C turnover					
HCRIT_LITTER	Scaling depth for litter humidity (m)	0.08 ± 0.192 0.02, 0.5			
KSOILC	Scalar on the active soil C pool content (to account for uncertainty in spin-up)	1 ± 0.6 0.5, 2			
MAINT_RESP_COEFF	Coefficient to calculate maintenance respiration as a fraction of biomass	1.4 ± 0.84 0.7, 2.4			
MAINT_RESP_SLOPE_C	Slope of maintenance respiration coefficient (1/K), constant c of aT^2+bT+c, tabulated	0.16 ± 0.064 0.08, 0.24	0.16 ± 0.064 0.08, 0.24	0.16 ± 0.064 0.08, 0.24	0.12 ± 0.06 0.04, 0.2
MAX_LTOLSR	Extrema of leaf allocation fraction	0.5 ± 0.08 0.4, 0.6			
MIN_LTOLSR	Extrema of leaf allocation fraction	0.2 ± 0.08 0.1, 0.3			
MOIST_COEFF__1	Coefficient to calculate moisture control for litter and soil C decomposition	1.1 ± 0.24 0.8, 1.4			
MOIST_COEFF__2	Coefficient to calculate moisture control for litter and soil C decomposition	2.4 ± 0.24 2.1, 2.7			
MOIST_COEFF__3	Coefficient to calculate moisture control for litter and soil C decomposition	0.29 ± 0.232 0.01, 0.59			
MOISTCONT_MIN	Minimum soil wetness to limit the heterotrophic respiration	0.25 ± 0.2 0.1, 0.6			

RESIDENCE_TIME	Residence time of trees (years)	40 ± 24 30, 90	40 ± 24 30, 90	40 ± 24 30, 90	0 ± 0 0,0
SOIL_Q10	Temperature dependency factor for heterotrophic respiration (Note: actual Q10 = exp$^{\mathrm{SOIL\_Q10}}$.	0.69 ± 0.44 0, 1.1			
TAU_FRUIT	Fruit lifetime (days)	90 ± 24 60, 120	90 ± 24 60, 120	90 ± 24 60, 120	-
TAU_METABOLIC	A coefficient to calculate residence times in metabolic litter pools (days)	0.066 ± 0.0112 0.052, 0.08			
TAU_SAP	Sapwood heartwood conversion time (days)	730 ± 144 550, 910	730 ± 144 550, 910	730 ± 144 550, 910	-
TAU_STRUCT	A coefficient to calculate residence times in structural litter pools (days)	0.245 ± 0.04 0.2, 0.3			
Phenology parameters					
GDD_THRESHOLD	Temperature threshold used in the calculation of number of growing degree day, GDD (days)	5 ± 0.8 4, 6			
GDDNCD_CURVE	Constant in the computation of critical GDD	0.0091 ± 0.00112 0.0072, 0.01			
GDDNCD_OFFSET	Constant in the computation of critical GDD (days)	64 ± 11.2 50, 78			
GDDNCD_REF	Reference value used in the computation of critical GDD (days)	603 ± 96.8 482, 724			
HUM_FRAC	Critical humidity (relative to min/max) for phenology (%)	-	-	-	0.5 ± 0.2 0.25, 0.75
HUM_MIN_TIME	Minimum time elapsed since moisture minimum (days)	-	-	-	35 ± 12 20, 50
LAI_MAX_TO_HAPPY	Threshold of LAI below which plant uses carbohydrate reserves	0.5 ± 0.14 0.35, 0.7	0.5 ± 0.14 0.35, 0.7	0.5 ± 0.14 0.35, 0.7	0.5 ± 0.14 0.35, 0.7
LEAF_AGE_CRIT_COEF_1	A coefficient to calculate critical leaf age (days)	1.5 ± 0.24 1.2, 1.8			
LEAF_AGE_CRIT_COEF_2	A coefficient to calculate critical leaf age (days)	0.75 ± 0.12 0.6, 0.9			
LEAF_AGE_CRIT_COEF_3	A coefficient to calculate critical leaf age (days)	10 ± 1.6 12, 8			
LEAF_AGE_CRIT_TREF	Reference temperature used to calculate of critical leaf age (days)	20 ± 4 15, 25			
LEAFAGE_FIRSTMAX	Leaf age at which vmax attains vcmax_opt (in fraction of critical leaf age)	0.03 ± 0.0048 0.024, 0.036			
LEAFAGE_LASTMAX	Leaf age at which vmax falls below vcmax_opt (in fraction of critical leaf age)	0.5 ± 0.08 0.4, 0.6			
LEAFAGECRIT	Critical leaf age, tabulated (days)	910 ± 200 610, 1210	730 ± 192 490, 970	180 ± 60 120, 240	120 ± 60 30, 180
LEAFFALL	Length of death of leaves, tabulated (days)	-	-	10 ± 4 5, 15	10 ± 4 5, 15
LEAFLIFE_TAB	Leaf longevity (years)	0.33 ± 0.1 0.2, 0.75	1 ± 0.668 0.33, 2	2 ± 0.9 0.75, 3	2 ± 0.9 0.75, 3
MAX_TURNOVER_TIME	Maximum turnover time for grass (days)	-	-	-	80 ± 4 75, 85
MIN_GROWTHINIT_TIME	Minimum time since last beginning of a growing season (days)	300 ± 24 270, 330			
MIN_LEAF_AGE_FOR_SENESCENCE	minimum leaf age to allow senescence (days)	-	-	90 ± 8 80, 100	30 ± 4 25, 35
MIN_TURNOVER_TIME	Minimum turnover time for grass (days)	-	-	-	10 ± 4 5, 15

NCD_MAX_YEAR	A coefficient to calculate maximum possible number of chilling days (NCD)	3 ± 0.8 2, 4			
NCDGDD_TEMP	Critical temperature for the ncd vs. gdd function in phenology (C)	-	-	5 ± 4 0, 10	-
NOSENESCENCE_HUM	Relative moisture availability above which there is no humidity-related senescence	-	-	-	0.3 ± 0.12 0.15, 0.45
PHENO_GDD_CRIT_A	Critical gdd tabulated constant a	-	-	-	0 ± 0 0, 0
PHENO_GDD_CRIT_B	Critical gdd constant b	-	-	-	0 ± 0 0, 0
PHENO_GDD_CRIT_C	Critical gdd constant c	-	-	-	400 ± 64 320, 480
PHENO_MOIGDD_T_CRIT	Average temperature threshold for C4 grass used (C)	-	-	-	22 ± 8 12, 32
SENESCENCE_HUM	Critical relative moisture availability for senescence	-	-	-	0.2 ± 0.08 0.1, 0.3
SENESCENCE_TEMP_A	Critical temperature for senescence (C), constant a of aT^2+bT+c, tabulated	-	-	0 ± 0 0, 0	0 ± 0 0, 0
SENESCENCE_TEMP_B	Critical temperature for senescence (C), constant b of aT^2+bT+c, tabulated	-	-	0 ± 0 0, 0	0 ± 0 0, 0
SENESCENCE_TEMP_C	Critical temperature for senescence (C), constant c of aT^2+bT+c, tabulated	-	-	12 ± 8 2, 22	5 ± 4.8 -1, 11
TAU_CLIMATOLOGY	tau for climatologic variables (days)	20 ± 8 10, 30			
TAU_GDD	Time scales for phenology and other processes (days)	40 ± 16 20, 60			
TAU_GPP_WEEK	Time scales for phenology and other processes (days)	6 ± 1 5, 7			
TAU_HUM_MONTH	Time scales for phenology and other processes (days)	20 ± 8 10, 30			
TAU_HUM_WEEK	Time scales for phenology and other processes (days)	6 ± 1 5, 7			
TAU_LEAFINIT	Time to attain the initial foliage using the carbohydrate reserve (days)	10 ± 10 5, 30	10 ± 10 5, 30	10 ± 10 5, 30	10 ± 10 5, 30
TAU_NGD	Time scales for phenology and other processes (days)	50 ± 20 25, 75			
TAU_SOILHUM_MONTH	Time scales for phenology and other processes (days)	20 ± 8 10, 30			
TAU_T2M_MONTH	Time constant for the "monthly" 2-meter temperature (days)	20 ± 8 10, 30			
TAU_T2M_WEEK	Time constant for the "weekly" 2-meter temperature (days)	6 ± 1 5, 7			
TAU_TSOIL_MONTH	Time constant for the "monthly" soil temperature (days)	20 ± 8 10, 30			
Conductance parameters - included in initial optimization sensitivity test but not in final optimizations					
A1	empirical factor involved in the calculation of fvpd	0.85 ± 0.04 0.8, 0.9	0.85 ± 0.04 0.8, 0.9	0.85 ± 0.04 0.8, 0.9	0.85 ± 0.04 0.8, 0.9
B1	empirical factor involved in the calculation of fvpd	0.14 ± 0.032 0.1, 0.18	0.14 ± 0.032 0.1, 0.18	0.14 ± 0.032 0.1, 0.18	0.2 ± 0.032 0.15, 0.25
CHOISNEL_RSOL_CSTE	Constant in the computation of resistance for bare soil evaporation (s/m2)	3.3E4 ± 19400 1.75E4, 6.6E4			
CONDVEG_Z0	Surface roughness (m)	0.15 ± 0.12 0, 0.3			
DEFC_MULT	Constant in the computation of surface resistance (KW^{-1})	1.5 ± 0.9 0.75, 3			
DEFC_PLUS	Constant in the computation of surface resistance (KW^{-1})	0.023 ± 0.016 0.003, 0.043			

g0	Residual stomatal conductance when irradiance approaches zero (mol m-2 s-1 bar-1)	0.00625 ± 0.00048 0.00565, 0.00685	0.00625 ± 0.00048 0.00565, 0.00685	0.00625 ± 0.00048 0.00565, 0.00685	0.01875 ± 0.0016 0.01675, 0.02075
GB_REF	Leaf bulk boundary layer resistance (s m-1)	0.04 ± 0.032 0, 0.08			
KZERO	A vegetation dependent constant used in the calculation of the surface resistance (kg/m^2/s)	0.00012 ± 0.000016 0.0001, 0.00014	0.00012 ± 0.000016 0.0001, 0.00014	0.00025 ± 0.00004 0.0002, 0.0003	0.0003 ± 0.00004 0.00025, 0.00035
RATIO_Z0M_Z0H	Ratio between z0m and z0h	1 ± 0.4 0.5, 1.5	1 ± 0.4 0.5, 1.5	1 ± 0.4 0.5, 1.5	1 ± 0.4 0.5, 1.5
Z_DECOMP	Scaling depth for soil activity	0.2 ± 0.6 0, 1.5			
Z0_BARE	Bare soil roughness length (m)	0.01 ± 0.0016 0.008, 0.012			
Z0_OVER_HEIGHT	To get z0 from height	0.0625 ± 0.032 0.02, 0.1			

Table S1. Prior information for all ORCHIDEE parameters optimized in this study: prior value, uncertainty and maximum and minimum bounds for the different plant functional types (temperate needleleaf/broadleaf evergreen (TeNE, TeBE) forests, temperate broadleaf deciduous (TeBD) forest, C4 grassland (GC4)).

Site	Daily root mean square error (RMSE)			
	GPP		R_{eco}	
	Prior	Posterior	Prior	Posterior
US-Vcm	1.392	1.366	2.009	1.315
US-Vcp	1.822	1.116	0.979	0.927
US-Mpj	1.312	0.903	1.051	0.867
US-Fuf	1.074	0.828	0.534	0.629
US-Wjs	0.995	0.73	0.778	0.672
US-Ses	0.259	0.229	0.233	0.255
US-Wkg	1.177	0.809	0.617	0.49
US-SRG	1.345	0.982	0.95	0.825
US-Seg	0.699	0.449	0.441	0.324
US-SRM	1.158	0.617	0.8	0.572
US-Whs	0.698	0.531	0.507	0.437
US-Aud	1.028	0.707	0.561	0.602

Table S2. Daily GPP and ecosystem respiration (R_{eco}) model-data fit when assimilating NEE observations with all parameters (P1) in terms of prior and posterior root mean square error (RMSE) for all twelve sites. The reduction in daily GPP RMSE varies between 0.03 to 0.7 $gCm^{-2}d^{-1}$, and the reduction in daily R_{eco} RMSE varies between 0 to 0.7 $gCm^{-2}d^{-1}$. The sites are listed in order from largest mean annual C sink (US-Vcm) to mean annual C source (US-Aud).

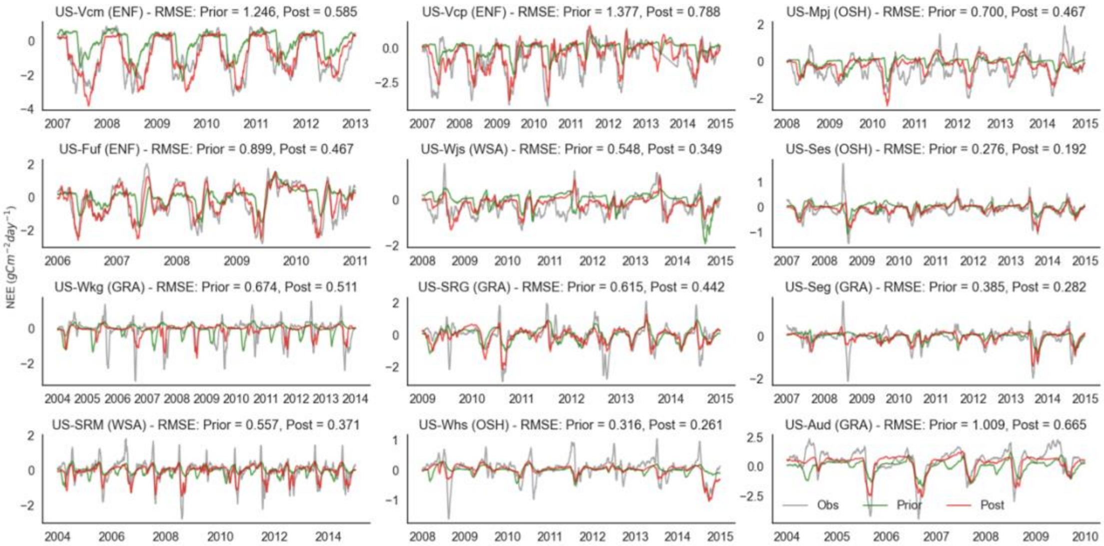

Figure S1. Comparison of NEE observations (grey) with corresponding ORCHIDEE model simulations before (green line) and after assimilation (red line) for assimilating NEE observations with all parameters (P1). The vegetation types are listed within brackets for each site. The RMSE measures the fit of the model prior and posterior simulations with the corresponding observations. Across all sites, the prior and posterior NEE RMSEs vary between 0.291-1.377 and 0.196-0.788, respectively. The sites are listed in order from largest mean annual C sink (US-Vcm) to mean annual C source (US-Aud).

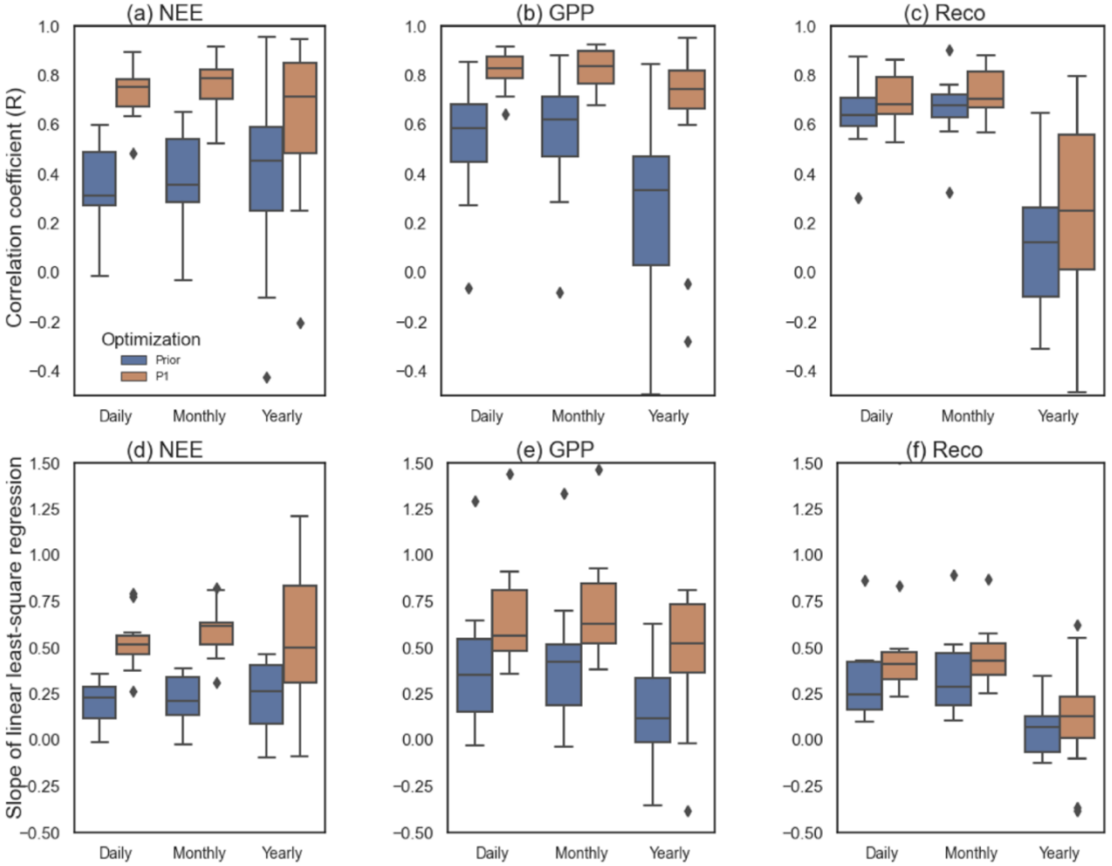

Figure S2. Daily, monthly and annual NEE (a, d), GPP (b, e) and R_{eco} (c, f) prior and posterior Pearson correlation coefficients (R) and slope values for the linear regression between model and observed fluxes for assimilating NEE observations and optimizing all phenology, photosynthesis and post C uptake parameters (P1). The R between observed and modeled NEE at daily, monthly and annual timescales for optimizing all parameters (P1) increase by up to 0.50, 0.55, 0.65 respectively. Note that the y axis limits for both R and slope are the same and therefore 3 sites fall outside the y-axis upper limit for the R_{eco} slope.

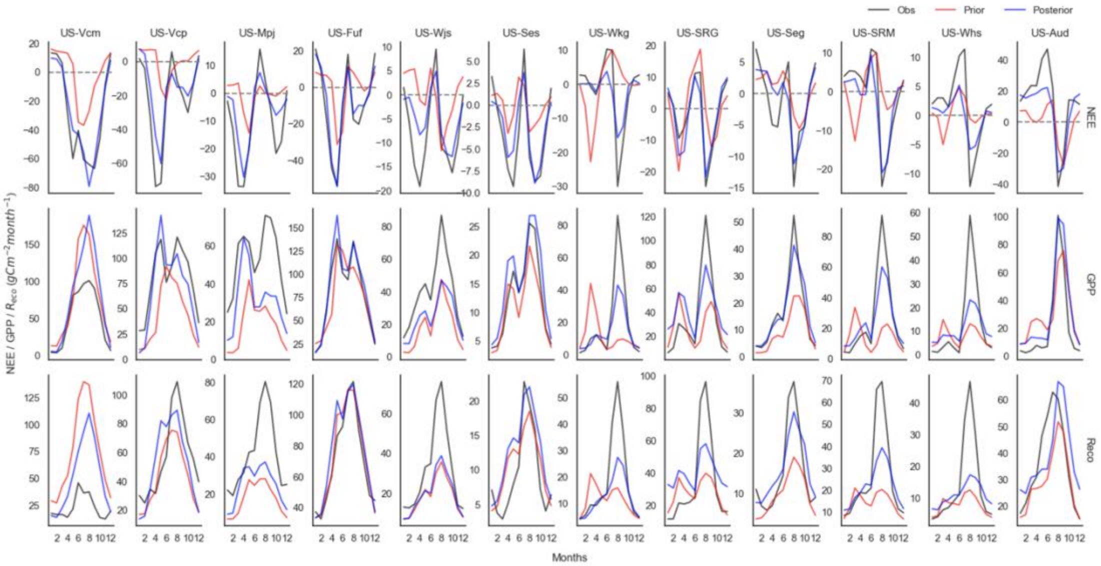

Figure S3. Seasonal cycle with mean monthly total fluxes. Comparison of flux observations with corresponding ORCHIDEE model simulations (prior and posterior) for assimilating NEE observations and optimizing all phenology, photosynthesis and post C uptake parameters (P1). The sites are listed from left to right according to C sink to source.

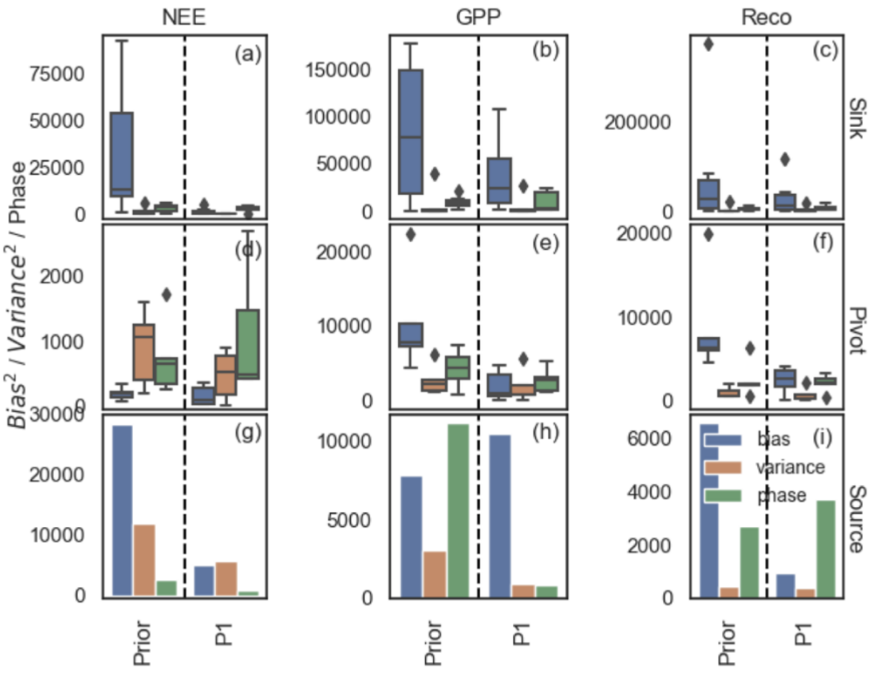

Figure S4. Annual NEE, GPP and R_{eco} mean square deviation (MSD) decomposition into bias, variance, and phase between simulations and observations for assimilating NEE observations and optimizing all phenology, photosynthesis and post C uptake parameters (P1). Different rows separate the sites as sink (a-c), pivot (d-f) and source (g-i) based on total annual C flux. The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the source site ia: US-Aud. The x axes display the optimization scenarios (Prior and P1). The box whiskers show the spread of bias, variance and phase for all 12 sites considered in this study. The bias, variance and phase indicate the mean difference in flux magnitude, the mismatch in terms of flux fluctuation magnitude scales with the mean seasonal amplitude, and the seasonality in flux time series, respectively. Note that the y axis limits are different for all fluxes and site types.

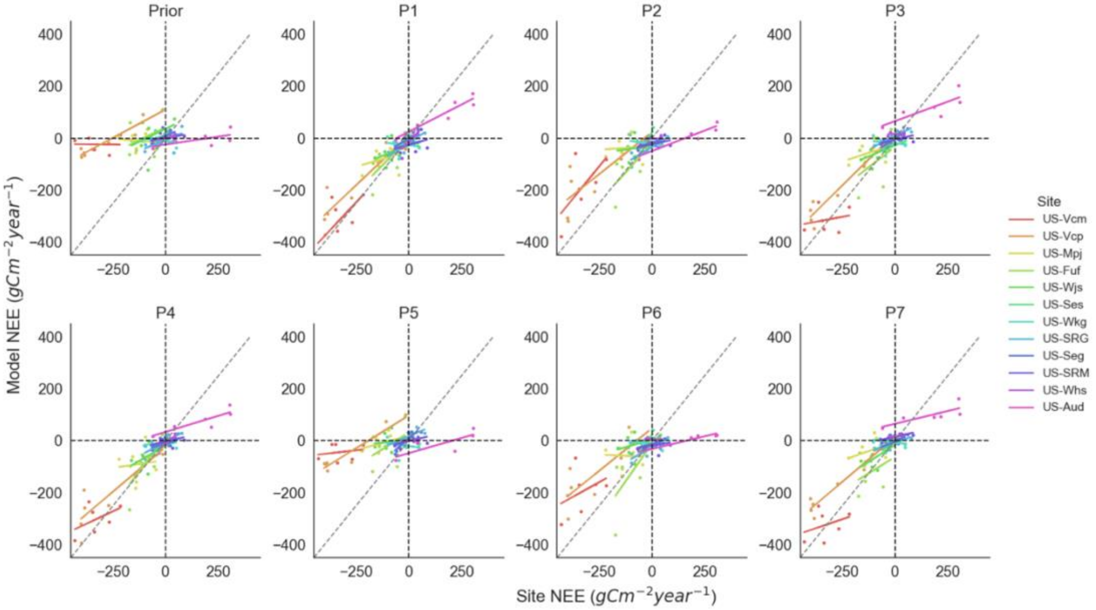

Figure S5. Annual NEE scatter plots for prior and all posterior simulations for assimilating NEE observations with various parameter sets (P1-P7). Different colour legends represent various sites, ordered from the largest mean sink (US-Vcm) to the largest mean source (US-Aud). The middle of the trend line should sit on the 1:1 line if the accurate mean annual source/sink behavior for a site is well captured by the model. A slope value close to or equal to 1 demonstrates the model is better at capturing the IAV. The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the source site is: US-Aud.

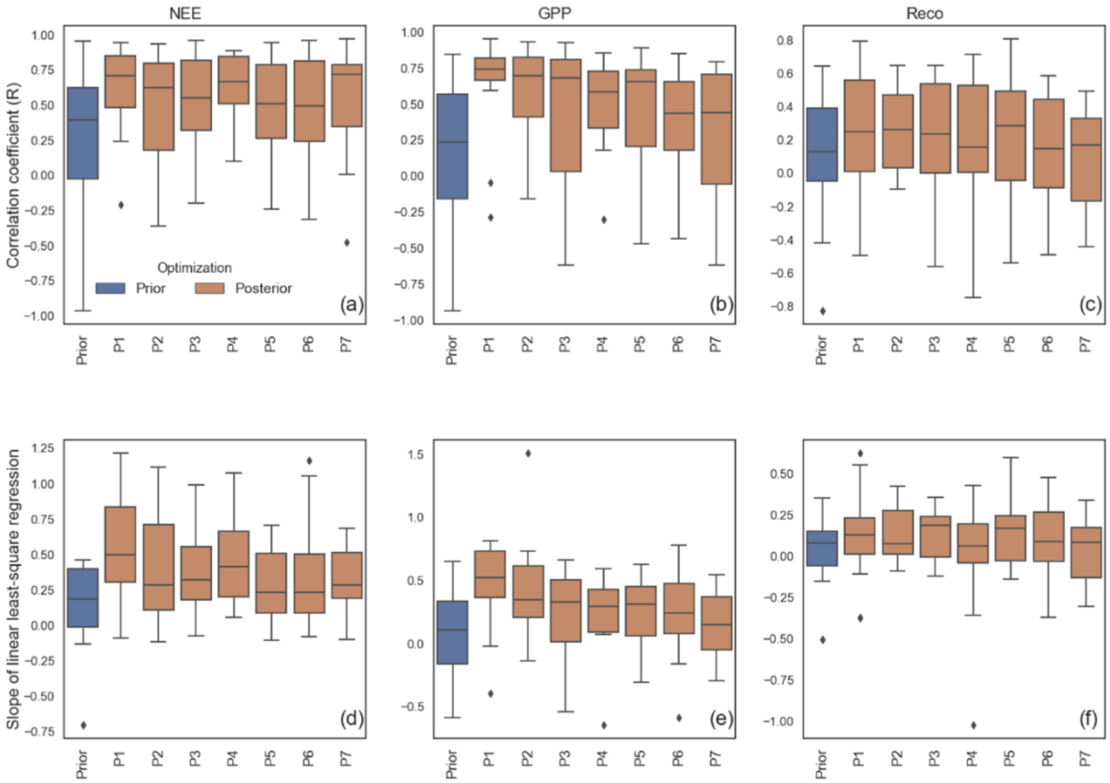

Figure S6. NEE (a, d), GPP (b, e) and R_{eco} (c, f) annual anomaly prior and posterior Pearson correlation coefficients (R) and slope values for the linear regression between model and observed fluxes across all assimilation scenarios with different parameter combinations (P1-P7). The legend represents various assimilation scenarios (Prior - blue bars, and posterior P1-P7 - orange bars).

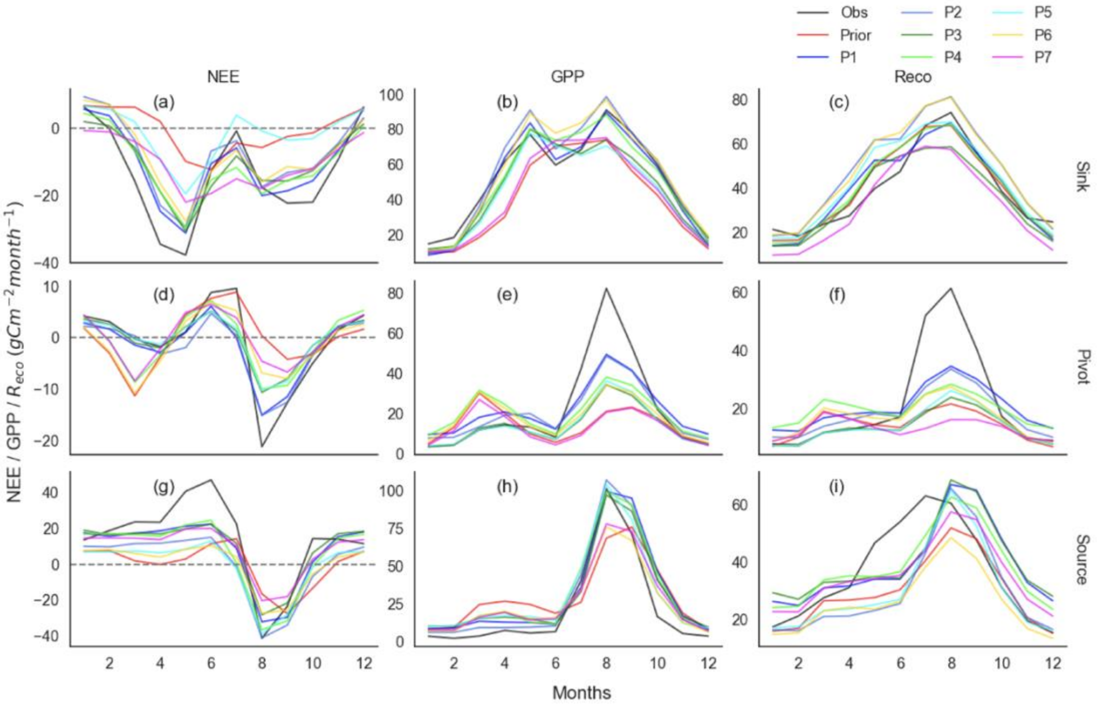

Figure S7. Mean monthly seasonal cycles comparing observations (black curve), prior (red curve), and posterior simulations for assimilation scenarios (P1 to P7 - blue to magenta curves) for NEE (left column), GPP (middle column), and R_{eco} (right column) averaged across site C balance types (sink - top row; pivot - middle row; and the source site, US-Aud, on the bottom row).

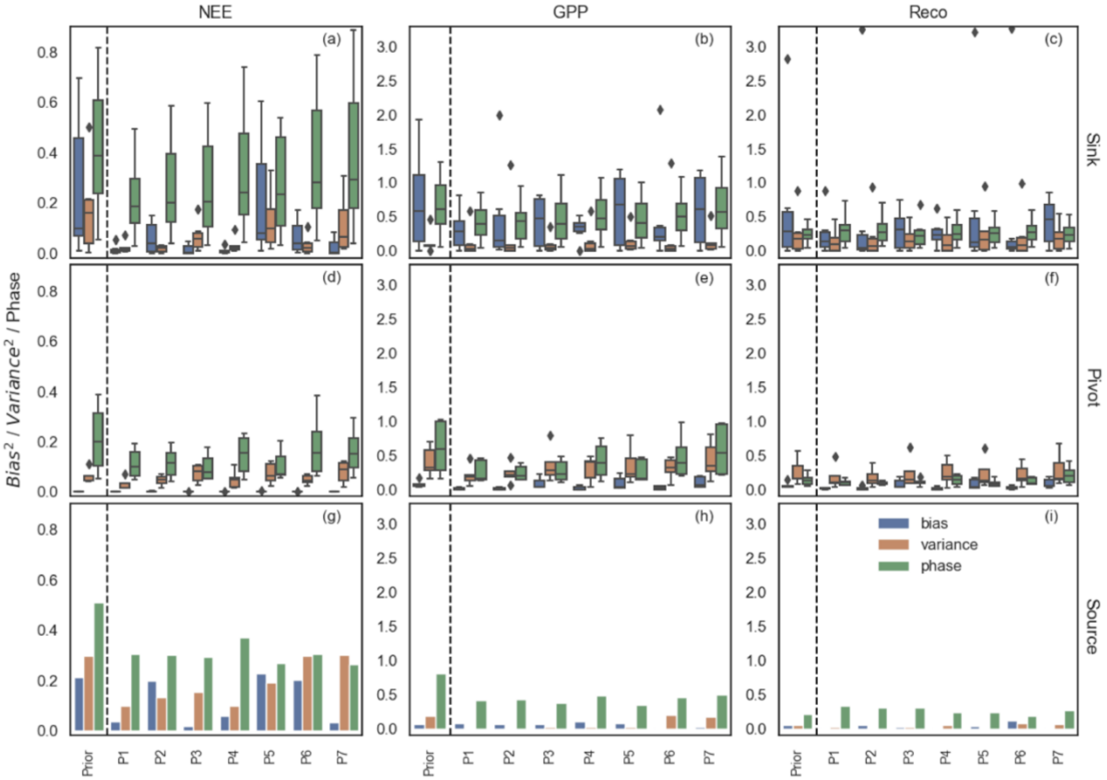

Figure S8. Daily NEE, GPP and R_{eco} mean square deviation (MSD) decomposition into bias, variance, and phase between simulations and observations for assimilating NEE observations with various parameter sets (P1-P7). Different rows separate the sites as sink (a-c), pivot (d-f) and source (g-i) based on total annual C flux. The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the only source site is: US-Aud. The x axes display various optimization scenarios (Prior, P1-P7). The parameters included in each optimization are: P1: all parameters; P2: phenology and photosynthesis; P3: phenology and post C uptake; P4: photosynthesis and post C uptake; P5: phenology; P6: photosynthesis and P7: post C uptake. The box whiskers show the spread of bias, variance and phase for all 12 sites considered in this study. The bias, variance and phase indicate the mean difference in flux magnitude, the mismatch in terms of flux fluctuation magnitude scales with the mean seasonal amplitude, and the seasonality in flux time series, respectively. Note that the y axis limits for both gross fluxes are the same.

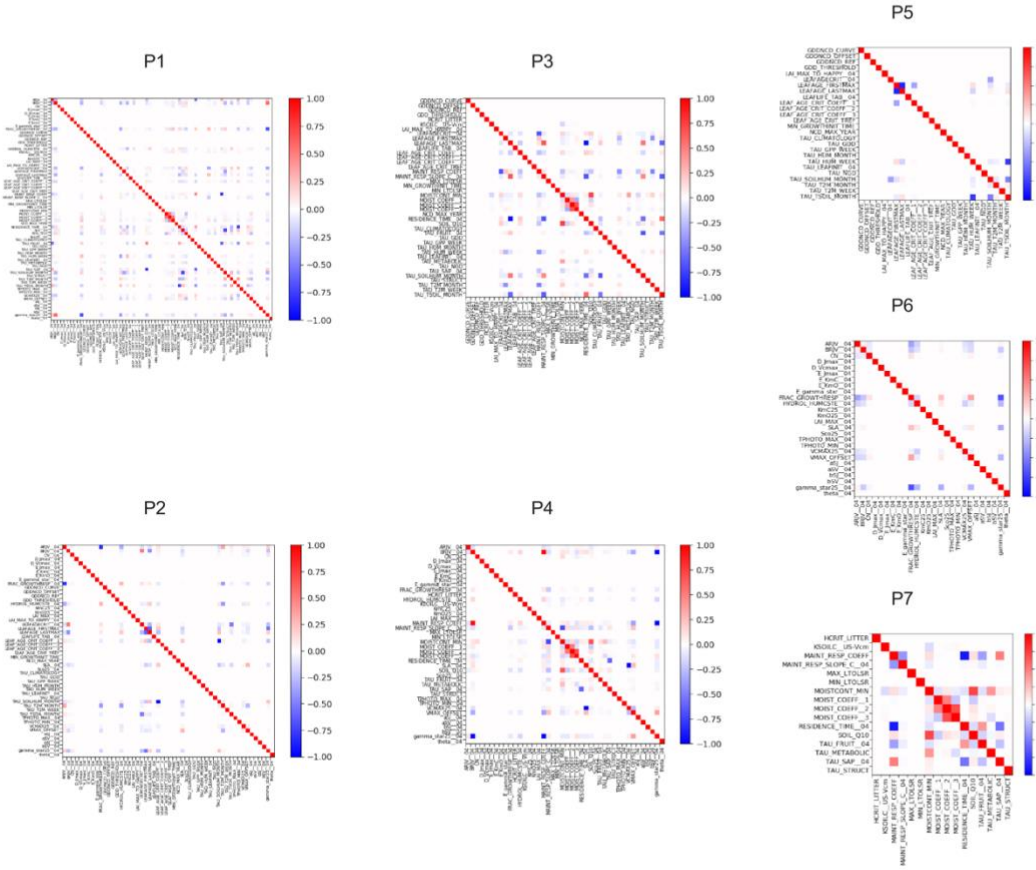

Figure S9. Parameter posterior error covariance matrix for US-Vcm for various assimilation scenarios (P1-P7).

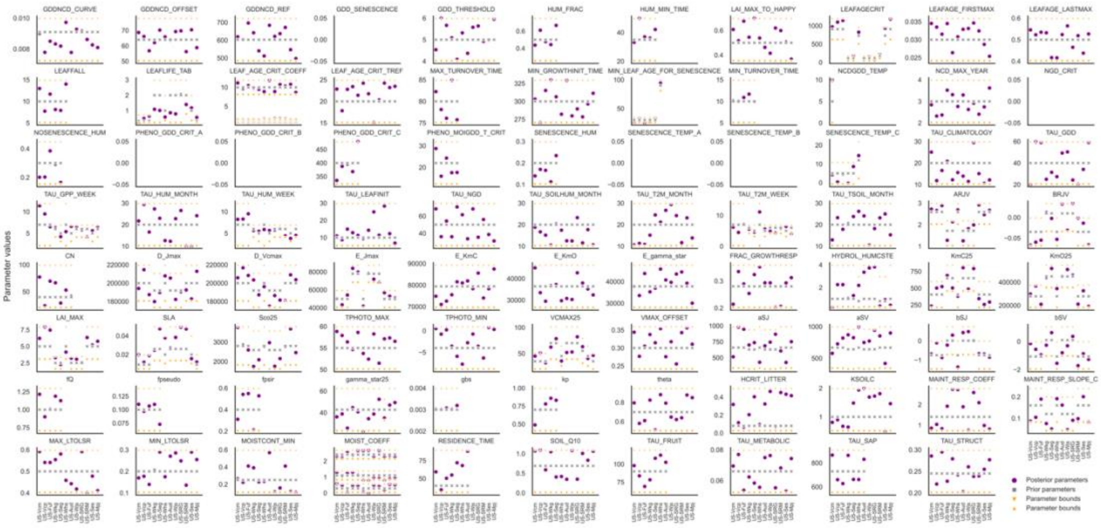

Figure S10. Values of all optimized parameters related to phenology, photosynthesis and post C uptake when assimilating NEE and optimizing all parameters (P1) for all 12 sites. For each parameter, the range of variation (corresponding to yellow arrows), the prior and the posterior values are provided for all sites. For the mixed-PFT sites, only the parameters for the majority PFT fraction are presented, although the other PFT parameters are also optimized. For example, this figure shows the parameters associated with PFT=4 (TeNE) for site US-Mpj, however the optimization is performed with all the parameters of both PFT=4 (TeNE) and 11 (C4G). Note that the soil Q10 parameter is the exponent of the actual Q10 value used to calculate heterotrophic respiration temperature sensitivity (see **Table S1** for more information).