

# 1 **Model-data fusion across ecosystems: from multi-site optimizations to** 2 **global simulations**

3

4 **S. Kuppel<sup>1,2</sup>, P. Peylin<sup>1</sup>, F. Maignan<sup>1</sup>, F. Chevallier<sup>1</sup>, G. Kiely<sup>3</sup>, L. Montagnani<sup>4</sup>, and A. Cescatti<sup>5</sup>**

5

6 [1]{Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ,  
7 91191 Gif-sur-Yvette cedex, France}

8 [2]{Grupo de Estudios Ambientales, IMASL - CONICET/Universidad Nacional de San Luis,  
9 Ejército de los Andes 950, D5700HHW San Luis, Argentina}

10 [3]{Civil and Environmental Engineering Department, and Environmental Research Institute,  
11 University College Cork, Cork, Ireland}

12 [4]{Forest Services, Autonomous Province of Bolzano, 39100 Bolzano, Italy}

13 [5]{European Commission, Joint Research Center, Institute for Environment and Sustainability,  
14 Ispra, Italy}

15 Correspondence to: S. Kuppel ([skuppel@unsl.edu.ar](mailto:skuppel@unsl.edu.ar))

16

## 17 **Abstract**

18 This study uses a variational data assimilation framework to simultaneously constrain a global  
19 ecosystem model with eddy covariance measurements of daily net ecosystem exchange (NEE) and  
20 latent heat (LE) fluxes from a large number of sites grouped in seven plant functional types (PFTs).  
21 It is an attempt to bridge the gap between the numerous site-specific parameter optimization works  
22 found in the literature and the generic parameterization used by most land surface models within  
23 each PFT. The present multi-site approach allows deriving PFT-generic sets of optimised  
24 parameters enhancing the agreement between measured and simulated fluxes at most of the sites  
25 considered, with performances often comparable to those of the corresponding site-specific  
26 optimizations. Besides reducing the PFT-averaged model-data root-mean-square difference  
27 (RMSD) and the associated daily output uncertainty, the optimization improves the simulated CO<sub>2</sub>  
28 balance at tropical and temperate forests sites. The major site-level NEE adjustments at the seasonal  
29 scale are: reduced amplitude in C3 grasslands and boreal forests, increased seasonality in temperate  
30 evergreen forests, and better model-data phasing in temperate deciduous broadleaf forests.  
31 Conversely, the poorer performances in tropical evergreen broadleaf forests points to deficiencies  
32 regarding the modeling of phenology and soil water stress for this PFT. An evaluation with  
33 data-oriented estimates of photosynthesis (GPP) and ecosystem respiration (Reco) rates indicates  
34 distinctively improved simulations of both gross fluxes. The multi-site parameter sets are then

1 tested against CO<sub>2</sub> concentrations measured at 53 locations around the globe, showing significant  
2 adjustments of the modelled seasonality of atmospheric CO<sub>2</sub> concentration, whose relevance seems  
3 PFT-dependent, along with an improved interannual variability. Lastly, a global scale evaluation  
4 with remote sensing NDVI measurements indicates an improvement of the simulated seasonal  
5 variations of the foliar cover for all considered PFTs.

6

## 7 **1. Introduction**

8 Land surface models (LSMs) have been tools of growing importance in the continuous  
9 effort to develop comprehensive Earth system models which help to understand the effects of  
10 changes in land surface processes and land-use practices upon biogeochemical (carbon, water,  
11 nutrients) and energy cycles, and more generally upon the Earth's climate (Cramer et al., 2001;  
12 Friedlingstein et al., 2006; Sitch et al., 2008). With the goal of improving accuracy and realism, the  
13 increasing amount and range of scale of the processes included in mechanistic LSMs result in a  
14 growing number of parameters associated with the corresponding model equations (Pitman, 2003).  
15 Some parameters are easily identified with a given physical process (and can sometimes be  
16 measured); others are purely empirical and account for a variety of processes embodied in a few  
17 equations, yet to be refined. In both cases, obvious computational and complexity limits have  
18 traditionally led model developers to use broad classes of soil and vegetation types, for which  
19 typical, generic parameter values are assigned (e.g., (Sellers et al., 1996)).

20 One difficulty is in scaling up the leaf- and plant-level measurements of physical parameters for  
21 ecosystem-scale simulations (Jarvis, 1995; Bonan et al., 2012). Besides, the variety of species  
22 within each of the 10 to 20 plant functional types (PFTs) typically used by most models makes the  
23 choice of a representative parameter value critical, thus adding significant uncertainty to the model  
24 outputs. In this context, parameter optimization methods have been increasingly used to calibrate  
25 model parameters and reduce the associated uncertainty. The criterion is to minimize the misfit  
26 between simulation outputs and observed data (Raupach et al., 2005). As for ecosystem models,  
27 eddy covariance measurements provide direct, near-continuous, in situ observations of carbon  
28 dioxide, water and energy exchanges between the canopy and the atmosphere (Baldocchi et al.,  
29 2001; Baldocchi, 2008). This measurement method has been applied across an extensive global  
30 network (560 sites as of October 2013), spanning a wide range of ecosystems and climates  
31 (<http://fluxnet.ornl.gov/>).

32 Over the last decade, numerous studies with various LSMs have used this available information to  
33 derive sets of parameters that significantly improve the model-data fit, with optimization  
34 approaches ranging from simple parameter adjustments to rigorous data assimilation frameworks  
35 (e.g., (Wang et al., 2001, 2007; Reichstein et al., 2003; Braswell et al., 2005; Knorr and Kattge,

1 2005; Santaren et al., 2007; Thum et al., 2008; Williams et al., 2009; Carvalhais et al., 2010;  
2 Keenan et al., 2012)). However, most of these efforts have focused on model calibration at  
3 individual sites. It often results in model parameters overly tuned to the specifics of a particular site  
4 given the small spatial footprint of each flux tower (typically a few hectares). Only recently, some  
5 studies started to assess through optimization the generic nature of model parameters within PFTs.  
6 The benefit of a set of parameters derived at one site was evaluated for simulations at a similar site  
7 (Medvigy et al., 2009; Verbeeck et al., 2011) and over the surrounding region (Medvigy and  
8 Moorcroft, 2012), with encouraging results. In parallel, two independent efforts simultaneously  
9 used data constraints from several sites to assess the degree of improvement of the simulated fluxes  
10 depending on the “generic criterion” used for the optimised parameters (Groenendijk et al., 2011;  
11 Kuppel et al., 2012). The study of (Groenendijk et al., 2011), conducted at over a hundred locations  
12 across several PFTs, found that the cross-site parameter variability after optimization explained the  
13 poorer performances of grouping sites by PFT, while no such discrepancy appeared in (Kuppel et  
14 al., 2012), a study however limited to temperate deciduous broadleaf forests.

15 Building on the optimization procedure developed by (Kuppel et al., 2012), the present work  
16 assesses the potential of the multi-site assimilation of carbon net ecosystem exchange (NEE) and  
17 latent heat (LE) flux measurements in a process-based terrestrial ecosystem model (ORCHIDEE).  
18 The objective is to improve site-scale simulations of carbon and water fluxes at a large number of  
19 flux towers sites, as well as global scale simulations of vegetation phenology and terrestrial carbon  
20 balance. Specifically, we address the following questions: 1) for each of the seven PFTs considered  
21 (out of 12 in ORCHIDEE, 5 being not covered by the measurements used here), can we find a  
22 generic set of optimised parameters that enhance the model-data fit at all sites? 2) how well does the  
23 multi-site approach compare to site-specific optimizations? 3) what are the main improvements  
24 introduced by the optimization procedure from seasonal to annual time scales: daily error,  
25 model-data bias, seasonal cycle amplitude and/or phase? 4) which processes remain poorly captured  
26 by the model after optimization? 5) have the eddy-covariance-constrained sets of multi-site  
27 parameters a notable impact on global scale simulations?

28 Section 2 presents the ecosystem model, the data assimilation system, and the eddy covariance  
29 measurements used in this study, as well the supplementary datasets and models. The results are  
30 presented and discussed in Sect. 3, successively dealing with the model-data fit at the site level  
31 (3.1), the comparison between multi- and single-site results (3.2), the uncertainties of modelled  
32 NEE and LE (3.3). Then are evaluated the impact of the derived multi-site parameterization upon  
33 the site-scale simulation of photosynthesis and ecosystem respiration rates (3.4), and at the global  
34 scale upon the simulated seasonality and interannual variability of atmospheric CO<sub>2</sub> concentration  
35 (3.5.1) and finally upon the seasonality of vegetation activity (3.5.2).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

## 2. Materials & Methods

### 2.1 Vegetation model and optimised parameters

The biogeochemical vegetation model used in this study is ORCHIDEE (Organizing Carbon and Hydrology in Dynamic Ecosystems, (Krinner et al., 2005)). It calculates the water, energy and carbon fluxes between the land surface and the atmosphere at a half-hourly time step. The exchange of carbon and water during photosynthesis and the energy balance are treated every 30 minutes, while carbon allocation, autotrophic respiration, foliar onset and senescence, mortality and soil organic matter decomposition are computed on a daily time step. The soil hydrology follows a double-bucket scheme (Ducoudré et al., 1993) and its impact on stomatal conductance is described in (Krinner et al., 2005). The reader is referred to previous publications for the standard equations of ORCHIDEE (e.g., (Kuppel et al., 2012)). As in most biogeochemical models, the vegetation is grouped into several PFTs, 12 in the case of ORCHIDEE, excluding bare soil. Except for the modelled phenology (Botta et al., 2000), the equations governing the different processes are generic across PFTs, but with specific parameter values for each vegetation class. When used in “grid-point mode” at a given site, we force the model with the corresponding half-hourly gap-filled meteorological data measured at the flux towers. At the global scale, the global ERA-Interim meteorology (Dee et al., 2011) is used as forcing and the model outputs are calculated at a 0.72 x 0.72 deg<sup>2</sup> resolution. In this case the global PFT map is computed at the spatial resolution of the forcing fields, from an original vegetation map available at 5km, which is derived from a high-resolution IGBP AVHRR global land dataset (Eidenshink and Faundeen, 1994) and uses 94 ecosystem classes (Olson, 1994). Importantly, the modelled carbon pools are initially brought to equilibrium before both site- and global-scale simulations by cycling the available meteorological forcing over several centuries (spin up procedure), with the prior parameterization of the model. This procedure ensures a net carbon flux close to zero over annual-to-decadal timescales. Table 1 presents the PFT-generic parameters used in this study. As our emphasis is on adjusting the carbon cycle, there are significantly more optimised parameters leveraging on photosynthesis and respiration processes than, for instance, on the energy balance. We also included two additional parameters to optimise the initial state of the model provided by the spin up procedure, which are 1) a common multiplier of the initial carbon pool content, by default equal to one, and 2) the initial leaf area index (LAI) of non-deciduous PFTs, by default taken from the spin up outputs. Both parameters are considered as site-specific, since the soil organic carbon content is closely related to the local land-use history, while the foliar cover of evergreen and herbaceous species directly relates to vegetation history at the site level. One consequence is that they cannot be spatially extrapolated,

1 thus the global simulations performed for evaluation (see sect. 2.4) use the default value of these  
2 last two parameters, i.e., using the initial carbon pool content and foliar cover provided by the spin  
3 up procedure.

4

## 5 **2.2 Data assimilation system**

6 The model parameters are optimised using the variational data assimilation method described in  
7 (Kuppel et al., 2012). Assuming a Gaussian distribution for errors on the parameters, the model  
8 outputs and the measured data, the optimised set of parameters corresponds to the minimization of  
9 the following Bayesian cost function  $J(\mathbf{x})$  (Tarantola, 2005):

10

$$11 \quad J(\mathbf{x}) = \frac{1}{2} \left[ (\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y} - H(\mathbf{x})) + (\mathbf{x} - \mathbf{x}_b)^T \mathbf{P}_b^{-1} (\mathbf{x} - \mathbf{x}_b) \right], \quad (1)$$

12

13 which quantifies both the misfit between modelled and observed fluxes, and the misfit between a  
14 priori and optimised parameters.  $\mathbf{x}$  is the vector of unknown parameters,  $\mathbf{x}_b$  the vector of background  
15 (i.e., here, prior) parameter values,  $H(\mathbf{x})$  the model output,  $\mathbf{y}$  the vector of observed fluxes, while  $\mathbf{P}_b$   
16 and  $\mathbf{R}$  are the prior covariance matrices of parameter errors and observation errors, respectively.

17 The cost function is iteratively minimized with the gradient-based algorithm L-BFGS-B, which  
18 allows prescribing boundaries for each variable to optimise (Byrd et al., 1995). At each iteration, the  
19 gradient of the cost function  $J(\mathbf{x})$  is computed with respect to all parameters, mostly using the  
20 Tangent Linear (TL) model of ORCHIDEE generated with the automatic differentiator tool TAF  
21 (Transformation of Algorithms in Fortran, see (Giering et al., 2005)). Exceptions concern two  
22 phenological parameters,  $K_{\text{pheno,crit}}$  and  $c_{\text{Tsenes}}$  (see Table 1), where the threshold functions prevent the  
23 use of a linear approximation. In these cases we use a finite-difference approach with prescribed  
24 perturbation steps respectively equal to 4% and 2% of the allowed variation range. The recent work  
25 of (Santaren et al., 2013) with the same ecosystem model highlighted the risk of converging  
26 towards a local minimum within a site-specific variational optimization. In our case, preliminary  
27 tests within three PFTs (tropical and temperate evergreen broadleaf forests, and temperate  
28 deciduous broadleaf forests) allowed us to verify that the convergence of our multi-site approach  
29 barely depends on the choice of the first-guess values assigned to the optimised parameters.  
30 However, such robustness is not guaranteed with the site-specific approach, and potential  
31 convergence issues are discussed in the results section.

32 Once the cost function reaches the minimum, the posterior parameter error variance/covariance  
33 matrix  $\mathbf{P}_a$  is explicitly calculated from the prior error covariance matrices ( $\mathbf{P}_b$  and  $\mathbf{R}$ ) and the

1 Jacobian of the model  $H$  at the minimum of the cost function ( $H_\infty$ ), using the linearity assumption  
2 (Tarantola, 2005):

$$P_a = [H_{\infty}^T R^{-1} H_\infty + P_{b^{-1}}]^{-1} \quad (2)$$

3  
4  
5  
6 The prior parameter error covariance matrix  $\mathbf{P}_b$  is diagonal as prior uncertainties are supposed to be  
7 uncorrelated between parameters. The prior standard deviation for each parameter is equal to 1/6 of  
8 the range between lower and higher boundaries. The latter have been carefully specified following  
9 the physical and empirical expertise of the ORCHIDEE modelers, based on literature reviews or  
10 databases (such as TRY, (Kattge et al., 2011)).

11 In the prior observation error covariance matrix  $\mathbf{R}$ , we include both the random error on the  
12 measurements and the model error, the latter stemming from missing/inadequate process  
13 representation in the structural equations of the ecosystem model. Although the measurements error  
14 is known not to be constant (e.g. (Richardson et al., 2008)), a previous study using the ORCHIDEE  
15 model suggested that the model component dominates the observation error budget (Kuppel et al.,  
16 2013). The variances in  $\mathbf{R}$  are chosen constant at a given site for each type of data (NEE and LE),  
17 equal to the mean square difference between the prior model and the observations. We also choose  
18 for simplicity to keep  $\mathbf{R}$  diagonal, based on the rapid decline of the model error autocorrelation  
19 beyond the first lag day (Kuppel et al., 2013).

20

### 21 **2.3 Assimilated eddy covariance flux data**

22 We use the eddy covariance data provided by 78 flux towers of the FLUXNET global network  
23 (Baldocchi, 2008), representative of seven of the 12 vegetated PFTs defined in the ORCHIDEE  
24 model (supplementary Table S2). All the sites of a given dominant ecosystem are located in the  
25 same geographical hemisphere, which makes seasonal analyses easier. These observations derive  
26 from standard flux data processing methodologies (correction, gap-filling and partitioning) of the  
27 *La Thuile* dataset (Papale, 2006). From the large amount of available site-years in this dataset, our  
28 selection was driven by several requirements, the first of these being a minimum vegetation cover  
29 of 70% by the dominant PFT within each tower footprint, based on site-level information. Then  
30 were discarded the sites where measurements show a significant disagreement with the prior  
31 simulation outputs, as it suggests strong model structural deficiencies that make the parameter  
32 optimization pointless. Lastly, we selected at each site the longest data segment of consecutive years  
33 without gaps larger than a few weeks. Where measurements of the ground heat flux (G) were  
34 available, the monthly energy balance was closed with a correction factor then

1 half-hourly-interpolated and applied to the latent heat (LE) and sensible heat (usually called H)  
2 fluxes, according to the Bowen ratio technique (Twine et al., 2000). The half-hourly, gap-filled  
3 measured fluxes of net ecosystem exchange (NEE) and LE are then used to compute daily means.  
4 We chose to assimilate daily-averaged observations and not half hourly measurements so as to focus  
5 the optimization on time scales ranging from seasonal to annual variations, and to take advantage of  
6 the rapidly-decreasing autocorrelation of gap-filled half-hourly fluxes (Lasslop et al., 2008). In  
7 order not to give too much weight to data estimated from gap-filling as compared to measured data,  
8 each daily observation error is inflated by a factor  $1+0.5k$ , where  $k$  is the daily fraction of  
9 half-hourly data estimated from gap-filling. We also checked that the gaps still remaining after the  
10 gap-filling were distributed evenly over the course of the day. The individual days with more than  
11 20% of these “ultimate” gaps were not included in the assimilation.  
12 The eddy covariance data are compared to the simulated fluxes in terms of RMSD and bias. In  
13 addition, for the six non-tropical PFTs we use a curve fitting procedure (composed of a polynome of  
14 degree 2 and four harmonics) to decompose the fluxes into their trends and mean seasonal cycles  
15 following (Thoning et al., 1989). The detrended smooth seasonal cycle is used to estimate the ratio  
16 between the average annual amplitude of the simulated and observed fluxes, as well as a model  
17 phase coefficient defined as

$$C_{phase} = 1 - \frac{|e_{sim} - e_{obs}| + |b_{sim} - b_{obs}|}{e_{obs} - b_{obs}} \quad (3)$$

18  
19  
20  
21 Here,  $b_i$  and  $e_i$  are respectively defined as the days when the detrended smooth curve crosses the  
22 zero line. In tropical evergreen broadleaf forests, the phase and amplitude diagnostics presented  
23 above are not applied, due to the lack of a marked seasonal cycle. Instead, the predictive power of  
24 the simulations is evaluated using the Nash-Sutcliffe model efficiency coefficient (Nash and  
25 Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum (F_{sim}^t - F_{obs}^t)^2}{\sum (F_{obs}^t - \overline{F_{obs}})^2}, \quad (4)$$

26  
27  
28  
29 where  $F_i^t$  is the value of the simulated or observed flux at the time step  $t$ , and  $\overline{F_{obs}}$  the mean  
30 observed flux.

31

## 32 2.4 Evaluation tools

1 The model is evaluated at the sites using the two components of NEE: the gross primary  
2 productivity (GPP) and the ecosystem respiration rate ( $R_{eco}$ ), both estimated via the flux-partitioning  
3 method described in (Reichstein et al., 2005). This method extrapolates nighttime measurement of  
4 NEE, representing nighttime  $R_{eco}$ , into daytime  $R_{eco}$  using a short-term-calibrated temperature  
5 response function. GPP is then derived as the difference between  $R_{eco}$  and NEE. We acknowledge  
6 that GPP and  $R_{eco}$  are not fully independent data (with respect to the assimilated NEE) and are  
7 essentially model-derived estimates somewhat conditional on our underlying assumptions, and it  
8 will be kept in mind during the analysis.

9 Additionally, measurements of the Normalized Difference Vegetation Index (NDVI) made by the  
10 MODIS instrument are used to evaluate the simulated phenology at the global scale. From 2000 to  
11 2010, the calculated reflectances (from measured irradiances) have been corrected for atmospheric  
12 absorption and scattering (Vermote et al., 2002) and directional effects (Vermote et al., 2009) in  
13 order to obtain a daily NDVI product with a 5-km spatial resolution. Observations contaminated  
14 with snow cover were removed from the analysis using MODIS' quality filter, and we discarded  
15 NDVI observations below 0.2 in order to minimize the impact of bare soil reflectance. Spatial  
16 averaging is used to match the ERA-Interim resolution ( $0.72^\circ \times 0.72^\circ$ ) used for the global scale  
17 simulations. Because it is directly derived from surface reflectances, we preferred NDVI to other  
18 satellite products such as FAPAR or LAI, the latter requiring intermediate processing steps usually  
19 involving radiative transfer models, and thus possibly adding uncertainty to the retrieved data  
20 (Garrigues et al., 2008). Following (Maignan et al., 2011), we then calculate the Pearson correlation  
21 factor between the times series of measured NDVI and the Fraction of Absorbed Photosynthetically  
22 Active Radiation (FAPAR) modelled by ORCHIDEE, at the weekly time scale during the period  
23 2000-2010. FAPAR has been estimated from modelled LAI with a simple Beer's law:

$$FAPAR = 1 - \exp(-0.5 \times LAI) \quad (5)$$

24  
25  
26  
27 The link between simulations and measurements is made by spatially averaging the latter to reach  
28 the resolution of the vegetation model (i.e. that of the ERA-Interim forcing). For each of the seven  
29 PFTs considered, we restrict our correlation computation to the model boxes where the dominant  
30 PFT's cover fraction exceeds 50% and where both NDVI and FAPAR time-series exhibit a visible  
31 seasonal cycle (i.e. with a standard deviation larger than 0.04).

32 Lastly, the simulated global NEE fluxes are output at the daily timescale and spatially averaged  
33 from the ERA-Interim grid ( $0.72^\circ \times 0.72^\circ$ ) to a  $2.5^\circ \times 3.75^\circ$  resolution (latitude, longitude). The  
34 LMDz atmospheric transport model (Hourdin et al., 2006) was used at this resolution to convert  
35 these terrestrial fluxes into monthly atmospheric  $CO_2$  concentrations, during the period 1989 to



1 2009. In order to complete the carbon balance at the planetary scale, we also transport the global  
2 oceanic and fossil net carbon fluxes respectively taken from a climatology (Takahashi et al., 2009)  
3 and from the EDGAR database (<http://edgar.jrc.ec.europa.eu>). The contribution of biomass burning  
4 is neglected, because re-growth of burnt vegetation is not accounted for in this version of  
5 ORCHIDEE, and so are the evasion of CO<sub>2</sub> from aquatic bodies and emissions from harvested  
6 wood and agricultural products. The transported fluxes are evaluated using 53 smoothed record of  
7 atmospheric CO<sub>2</sub> concentrations ( $C_{CO_2}$ ) over the globe (supplementary Table S3)  
8 (GLOBALVIEW-CO<sub>2</sub>, 2013). As the optimization of the initial soil carbon content cannot be  
9 spatially extrapolated for global simulations (see sect. 2.1), the modelled trend of  $C_{CO_2}$  is not  
10 evaluated. Rather, we focus on the seasonal analysis and use the curve-fitting procedures of  
11 (Thoning et al., 1989) to extract the detrended seasonal signal of  $C_{CO_2}$ . In addition, we identify the  
12 contributions of 11 sub-continental regions to the simulated atmospheric CO<sub>2</sub> concentration at each  
13 station by independently transporting the fluxes from each following area: boreal North America,  
14 temperate North America, tropical America, South America, Europe, northern Africa, southern  
15 Africa, boreal Asia, temperate Asia, tropical Asia, and Australia (e.g., Fig. 1 in (Gurney et al.,  
16 2003)). The simulated interannual variability of the  $C_{CO_2}$  is evaluated using the model-data RMSD  
17 of monthly anomaly, from the detrended smooth seasonal signal calculated above:

$$18 \quad C_{CO_2,anom} = C_{CO_2,month} - \langle C_{CO_2,month} \rangle_{allyears}, \quad (6)$$

19 where  $\langle C_{CO_2,month} \rangle_{allyears}$  is the all-time average, for each month of the year.

20

## 21 **3 Results & discussion**

22

### 23 **3.1 Site-level simulation of carbon and water fluxes**

24 Figure 1 shows the average corrections brought by the optimization to the modelled NEE fluxes  
25 (with negative values meaning carbon uptake), grouped by dominant PFT (see acronyms in Table  
26 1), in terms of RMSD and bias between simulations and measured data, also showing the  
27 PFT-averaged mean seasonal cycles. The largest reductions of model-data RMSD are found in  
28 temperate and boreal broadleaf forests (TempEBF, TempDBF and BorDBF), where the two  
29 optimization scenarios (single- and multi-site) decrease the misfit by more than 25% compared with  
30 the prior (unoptimized) model. In temperate needleleaf forests (TempENF) and C3 grasslands  
31 (C3grass), the RMSD reduction exceeds 30% for single-site optimizations, but the corresponding  
32 multi-site sets of parameters reduce this value to less than 20%. The improvements are less  
33 significant in tropical evergreen broadleaf forests (TropEBF) and boreal evergreen needleleaf  
34 forests (BorENF), where the reductions of misfit is between 9 and 15%. Figure 1B shows that the

1 NEE is on average overestimated by the prior model for all PFTs. This feature is even more striking  
2 in ecosystems which are marked sinks of carbon (according to the average measured carbon  
3 balance, not shown), here tropical and temperate forests. This positive bias is an artifact from the  
4 model initialization procedure, which brings each simulated site to a near equilibrium (see sect.  
5 2.1). It is significantly corrected by the optimization, notably via the scaling of the initial carbon  
6 pool content at each site (parameter  $K_{\text{soilC}}$  in Table 1), one consequence being a clear reduction of  
7 the respiration during the winter of temperate and boreal ecosystems and grasslands sites in  
8 agreement with the measured data (Fig. 1C).

9  
10 Figure 2 shows that the simulation of the latent heat flux (LE) is overall less improved by the  
11 optimizations than that of the net carbon flux, keeping however in mind the problem of energy  
12 balance closure discussed in sect. 2.3. The reduction of RMSD is the highest on average at  
13 TempDBF sites with values 24% below the prior value, while decreases of 15 to 19% are found at  
14 TempEBF and BorENF sites. The effect of the optimization is the weakest on average at sites  
15 located in TempENF and C3grass ecosystems. These weaker performances regarding LE flux  
16 indicate that the energy and water cycles in the ORCHIDEE model involve other relevant  
17 parameters not optimised here, and possibly that the structural equations bear a significant error.  
18 Notably, we include in the optimization only one parameter that directly controls the soil  
19 evaporation ( $ZO_{\text{overheight}}$ , see Table 1), and there is for example no constraint on the calculation of the  
20 surface temperature, a key component of the energy balance.

21  
22 At the seasonal scale, Fig. 3A shows that large reductions (in relative value) of the simulated  
23 mean seasonal NEE amplitude are found in boreal evergreen needleleaf and deciduous broadleaf  
24 forests and C3 grasslands. The average correction is somewhat exaggerated in the two former cases  
25 and relatively accurate in the latter case. Conversely, the seasonal NEE variations are consistently  
26 amplified by the optimization in temperate evergreen needleleaf and broadleaf forests. However, the  
27 averaged model-data phasing is only weakly modified for the five aforementioned PFTs, with the  
28 exception of the site-specific improvements at TempENF and C3grass sites. Besides, considering  
29 the mild correction of the model-data biases in BorENF, BorDBF and C3grass (Fig. 1B), one can  
30 deduce that most of the reduction of RMSD discussed earlier is for these three PFTs due to an  
31 improvement of the simulated NEE amplitude after the optimization.

32 In temperate deciduous broadleaf forests, the simulated pattern of NEE is chiefly improved via a  
33 better phased seasonal cycle, as shown by the increased phase score, which was already close to one  
34 before optimization. An earlier study at a similar set of sites of the same PFT showed that the  
35 optimization scheme tends to correct the overall prior model overestimation of the growing season

1 length (Kuppel et al., 2012). On the other hand, the simulated seasonal amplitude of NEE is barely  
2 changed after optimization, as the corrected flux overestimations in winter and summer tend to  
3 cancel out, with a PFT-averaged seasonal amplitude remaining smaller than that of the observed  
4 data (Figs. 1C and 3).

5 Regarding the latent heat flux, Fig. 3B shows that the optimization has generally a weaker effect  
6 on the simulated LE average phase and amplitude than in the case of NEE. In most cases the  
7 correction brought by the optimization barely affects the modelled phase, but improves the seasonal  
8 amplitude. We notice that the LE seasonal cycle is most often flattened as compared to the prior  
9 model in agreement with the observations, except for the inconsistent amplification at TempEBF  
10 sites and the over-reduction after the site-specific optimization in C3 grasslands. The weak phase  
11 correction might be related to the soil evaporation component of the latent heat flux, on which the  
12 optimization has a limited leverage as mentioned earlier in this section, while the transpiration rate  
13 is tightly linked to GPP. It would also explain the generally lower phase coefficient in deciduous  
14 ecosystems (Fig. 3), where soil evaporation is a potentially significant component of LE during leaf  
15 onset and senescence.

16

17 Besides, applying the Nash-Sutcliffe model efficiency coefficient (NSE, see Eq. 4) to all  
18 sites shows that TropEBF is the only PFT studied here where the PFT-averaged value of this metric  
19 remains below zero after optimization for both fluxes, with  $NSE_{NEE}=[-2.77, -1.99, -1.83]$  and  
20  $NSE_{LE}=[-0.64, -0.35, -0.52]$  in prior, multi-site and site-specific cases, respectively (other PFT  
21 values not shown). It means that after optimization the model-data mean square error is still larger  
22 than the variance of the observations, or, in other words, that the observed mean is on average a  
23 better predictor than the model outputs. Figure 1 shows that for TropEBF the prior model simulates  
24 an unobserved increase of NEE from sink to source around July, and the simultaneous decrease of  
25 LE (Fig. 2) points towards an unrealistic simulated drought stress during this period of the year, the  
26 driest at most sites of this PFT. The optimization barely corrects the NEE variations during the dry  
27 season, although a more realistic LE flux is simulated after the multi-site optimization. An earlier  
28 optimization study at a site of the same PFT highlighted the need for a much deeper soil water  
29 reservoir along with a more linear root profile than that parameterized in the prior model, in order to  
30 account for the ability of tropical evergreen forests to maintain high photosynthesis and  
31 transpiration during the dry season (Verbeeck et al., 2011). Our multi-site parameterization of the  
32 processes dealing with soil water availability goes in that direction, with values of soil water depth,  
33 root profile and water stress coefficient respectively adjusted from 2 to  $2.38 \pm 0.065$  m, from 0.8 to  
34  $0.72 \pm 0.095$  m<sup>-1</sup> and from 6 to  $6.5 \pm 1.06$  (Table 1). These corrections from the prior  
35 parameterization remain however insufficient, as shown by the poorly realistic optimised seasonal

1 cycle of NEE in Fig. 1. On the other hand, (Verbeeck et al., 2011) also pointed at the structural  
2 inconsistency in the standard ORCHIDEE model for tropical evergreen forests: the phenological  
3 scheme notably neglects the leaf renewal at the transition between wet and dry season (Chave et al.,  
4 2010) and the hydric stress calculation ignores the role of groundwater (Miguez-Macho and Fan,  
5 2012), while these mechanisms possibly explain the high subsequent photosynthesis and  
6 transpiration rates often observed (Weirdt et al., 2012). Concerning the LE flux, the optimization  
7 brings somewhat limited, yet consistent changes, while the reduction of daily uncertainty is modest,  
8 indicating a poor level of constraint by the observations used. It suggests either significant errors in  
9 the model equations, or that relevant, poorly known parameters, have not been considered.

10

### 11 **3.2 Single-site versus multi-site**

12

13 It can be noticed in Figs. 1, 2 and 3 that there is a general consistency across PFTs between  
14 RMSD reductions introduced by multi-site and site-specific optimizations, with some exceptions in  
15 TempENF and most notably C3grass where the average site-specific RMSD reduction is twice as  
16 large for NEE, while there is almost no average multi-site RMSD decrease for LE. Although the  
17 large number of sites selected for this last PFT and the associated inter-site variability calls for  
18 prudence when considering average seasonal flux variations, it is worth noting that C3 grasslands  
19 are here the only PFT generically spanning such a diversity of climates. The reported discrepancy  
20 might thus indicate a need for additional classes of C3 grasslands in the model, at least with a  
21 climatic regionalization and ideally taking also into account pedologic conditions and management  
22 practices.

23 More generally, one would expect better efficiency from a site-specific scheme than with a  
24 multi-site approach, given that grouping sites with different characteristics introduces conflicting  
25 constraints on the model equations, along with the fact that the RMSD is the criterion used in the  
26 optimization procedure (as the prior covariance matrix in the cost function of Eq. (1) is chosen  
27 diagonal, see sect. 2.2). It is true most of the time, except notably for NEE in boreal deciduous  
28 broadleaf forests and LE at TropEBF, TempEBF and BorENF sites where the multi-site  
29 optimization results on average in larger RMSD decreases than the site-specific approach. In these  
30 cases, Figs. 1 and 2 show that it stems from unchanged local RMSD after the site-specific  
31 optimization at a few sites of these particular PFTs. As found by (Santaren et al., 2013), it may point  
32 to a failure of the single-site inversion algorithm to converge towards the global minimum of the  
33 cost function, possibly due to the presence of local minima. Our hypothesis is that the  
34 corresponding multi-site cost functions avoid this pitfall because they are made more regular by the  
35 larger amount of simultaneous constraints on the parameters, “smoothing out” some of the

1 problematic local minimums. Preliminary multi-site optimization tests, using a few tens of random  
2 starting points, support this hypothesis, and further investigations will be needed to evaluate if this  
3 behavior is valid for all PFTs. Indeed, we acknowledge some uncertainty regarding whether or not  
4 the optimised sets of parameters correspond to the very minimum of the cost function, as the  
5 efficiency of the variational optimization approach employed is conditional on a reasonable  
6 compliance with the linearity hypothesis.

7

### 8 **3.3 Site-scale uncertainty**

9 In addition to improving the agreement between modelled and measured fluxes, the  
10 optimization procedure is also useful to reduce the total uncertainty associated with the modelled  
11 output variables at each site, defined as:

12

$$13 \quad \sigma_{total} = \sqrt{\sigma_{observations}^2 + \sigma_{parameters}^2} . \quad (7)$$

14

15  $\sigma_{total}$  represents the summed contribution of two errors arising in the observation space: the  
16 measurements error and the error of the equations of the model (see sect. 2.2). It is not directly  
17 altered through parameter optimization, although the model component may in principle vary with  
18 the parameter values. Following (Desrozier et al., 2005),  $\sigma_{total}$  is diagnosed at each site as the square  
19 root of the covariance between the time series of prior and posterior flux residuals (model minus  
20 observations).  $\sigma_{parameters}$  is the parameter error contribution to the simulated fluxes, calculated at each  
21 site, before optimization, as the average daily standard deviation of the projection in observation  
22 space of the prior error covariance matrix  $\mathbf{P}_b$ , using the model's Jacobian matrix  $\mathbf{H}$ , based on the  
23 definitions of sect 2.2. The same is done after optimization, respectively using  $\mathbf{P}_a$  and  $\mathbf{H}_\infty$ .

24 Figure 4 reports the average value of  $\sigma_{total}$  for simulated daily NEE and LE, showing individual sites  
25 values and the corresponding PFT means as in Figs. 1 and 2. The reduction of the total NEE  
26 uncertainty varies from one PFT to another, ranging from 6% in tropical evergreen broadleaf forests  
27 to 33% in boreal evergreen needleleaf forests. As  $\sigma_{parameters}$  is reduced by 65 to 95% (not shown), we deduce  
28 from Eq. (7) that the weak relative decrease of  $\sigma_{total}$  indicates a dominance of the observation error in the  
29 total uncertainty budget. This is for example consistent with the reported inaccuracies in the model  
30 structure for TropEBF ecosystems discussed in the previous section.

31 Regarding the LE flux, the mild changes from prior to posterior uncertainty means that we might  
32 face a potentially large observation error component (model + measurements) –the latter being  
33 insensitive to parameter optimization, see sect. 2.3– and overall that little statistical information has  
34 been gained from the optimization of the selected water cycle parameters.

1 Besides, one can notice that the posterior uncertainties are always slightly lower for multi-site  
2 optimization than with the site-specific approach, which is consistent with the fact that the number  
3 of assimilated data is larger in the former case than in the latter. Finally, we also found that the  
4 optimization suppresses at each site much of the temporal correlation of the flux error, which are  
5 large in the prior ORCHIDEE model (see for instance the time correlogram in the Fig. 1 of (Kuppel  
6 et al., 2013)). It results in a large decrease of the total yearly uncertainty from the prior model for all  
7 PFTs, by 77 to 86% and 43 to 80% for simulated NEE and LE flux, respectively (not shown).

8

### 9 **3.4 Simulated GPP and respiration**

10 Figure 5 shows the mean seasonal cycle, averaged over each PFT, for the gross carbon fluxes:  
11 photosynthesis (GPP, Fig. 5A) and ecosystem respiration ( $R_{\text{eco}}$ , Fig. 5B). The “observed” values are  
12 estimates based on a partition of the measured NEE (see sect. 2.4). These gross carbon fluxes have  
13 not been used as constraints in the optimization procedure, but are useful as indicators of the model  
14 performance. One can first notice that the average increases of GPP at TropEBF and TempEBF sites  
15 are responsible for the NEE decrease observed in Fig. 1. It is worth noting that the results reported  
16 in Fig. 5 also confirm the inability of the model to simulate a sustained high photosynthesis rate  
17 during the dry season at TropEBF sites (see Sect. 3.1), while this feature appears in the observations  
18 estimates. At TempENF sites, the remarkable adjustment of the NEE cycle primarily derives from a  
19 reduced  $R_{\text{eco}}$  at the peak of the growing season. Both GPP and  $R_{\text{eco}}$  are consistently decreased in  
20 boreal forests and C3 grasslands sites, although the reduction is still lower than what would be  
21 needed to match the estimates. In addition, because the respiration rate is the sole reducing  
22 component in winter and because the photosynthesis rate is more largely decreased than  $R_{\text{eco}}$  during  
23 the growing season, the net result is the reduction of the seasonal amplitude of NEE for these three  
24 PFTs. Finally there is a large, yet insufficient, decrease of  $R_{\text{eco}}$  after the optimization in temperate  
25 deciduous broadleaf forests, notably related to the scaling of the initial carbon pool content (sect 3.1  
26 and (Kuppel et al., 2012)), while GPP is less drastically reduced, in close agreement with the  
27 observations. This evaluation at each site with gross carbon fluxes shows that the optimization  
28 procedure is able to provide a set of parameters which improves the simulation of both assimilation  
29 and respiration processes in the ORCHIDEE model for six out of the seven PFTs considered here,  
30 suggesting a partial distinction of both gross contributions from the constraint provided by the net  
31 carbon flux.

32

### 33 **3.5 Global-scale evaluation**

34 One of the objectives of assimilating flux data from a large number of sites, spanning a wide  
35 range of ecosystems, is to identify generic sets of parameters that improve the simulation of carbon

1 and water balance at the regional-to-global scale. Indeed, there is no guarantee that a set of  
2 parameters improving the simulations at an ensemble of individual sites sharing broadly common  
3 biogeochemical and biophysical characteristics, but with a limited spatial footprint, will also be  
4 beneficial for simulations at much larger scales. In this context, global simulations allow evaluation  
5 of how the constraint of eddy covariance data is propagated from one spatial scale to another, and  
6 how transferable the optimised parameterization is from grouped in situ optimizations to gridded  
7 simulations.

8

### 9 **3.5.1 Seasonality of atmospheric CO<sub>2</sub> concentrations**

10

11 Regarding the simulated mean seasonal cycle of atmospheric C<sub>CO2</sub>, the optimised set of  
12 parameters yields a median reduction of the model-data RMSD of 5.2%. Among the 53 samples  
13 locations used in this study, there is a significant improvement at 27 of them with a RMSD decrease  
14 larger than 5%, a notable degradation at 20 sites with a RMSD increase larger than 5%, and less  
15 than a 5% shift at the remaining 6 locations. In addition, a latitudinal clustering can be identified, as  
16 a large median improvement by 42.2% (RMSD-wise) is found at the 3 northernmost sites (Alert,  
17 Ny-Alesund, and Barrow) and by 33.5% at the 18 locations of the Southern Hemisphere, while  
18 there is a median degradation by 5.6% in the rest of the Northern Hemisphere.

19 Figure 6 shows the mean seasonal cycle of the simulated C<sub>CO2</sub>, compared to the extended record  
20 at three locations, one in each of the latitudinal areas defined above: Alert, South Pole, and Mauna  
21 Loa, respectively. We note that using the optimised parameters sets tends to reduce the seasonal  
22 amplitude of C<sub>CO2</sub>, with in the Northern Hemisphere an earlier phasing for the “breathing of the  
23 biosphere”. At station Alert, there is a significant adjustment of the simulated seasonal cycle, when  
24 changing from the default to the multi-site parameterization of the ORCHIDEE model. This  
25 correction chiefly benefits the seasonal amplitude, which is decreased and becomes remarkably  
26 close to that observed. The analysis of the contribution of the 11 sub-continental regions in the  
27 simulated atmospheric signal (see sect. 2.4), grouped in Fig. 6D in larger regions, indicates that the  
28 major terrestrial contribution to this result are changes in C<sub>CO2</sub> due to the boreal Northern  
29 Hemisphere fluxes. It is consistent with the decrease of the NEE seasonal amplitude produced by  
30 the multi-site optimization at sites in boreal evergreen needleleaf forests, boreal deciduous  
31 broadleaf forests and C3 grasslands (Figs 1C and 3A), dominant in this region. Separate global  
32 simulations using optimised parameterization for one PFT at a time show that the degraded phasing  
33 at Alert produced by the multi-site approach in Fig. 6A mainly stems from the contributions of  
34 BorENF and C3 grasslands ecosystems (not shown). While a mild multi-site phase deterioration  
35 from the prior parameterization is found at the site level in BorENF (Fig. 3A), it is not the case for

1 C3 grassland and it may thus question the representativeness of the flux measurements sites used,  
2 with respect to high-latitudes ecosystems in general.

3 At station South Pole, the model-data fit is also mostly enhanced after the optimization by a  
4 significant decrease of the seasonal amplitude of  $C_{CO_2}$ , which is more than twice as large in the prior  
5 simulation as in the measured data. Besides, the “regionalized” analysis indicates that the  
6 corrections are primarily due to the reduced seasonal amplitude of  $C_{CO_2}$  components from temperate  
7 South America and southern Africa, and contributions from the boreal Northern Hemisphere are  
8 also noticeable (not shown). We therefore deduce that the optimization of C3 grasslands parameters  
9 is the most influential factor explaining the improved simulation of  $C_{CO_2}$  at this station, but also that  
10 the influence of boreal needleleaf evergreen forests cannot be neglected here.

11 The reduction of the simulated cycle amplitude is too strong at station Mauna Loa, which,  
12 combined with earlier seasonality, leads to a poorer model-data fit after optimization. The remote  
13 location of Mauna Loa (north Pacific) makes it sensitive to influences from most of the Northern  
14 Hemisphere. We find that the main drivers of the simulated correction are flattened, earlier  $C_{CO_2}$   
15 variations in temperate and boreal regions of North America and Asia, and Europe (not shown).  
16 These results thus reflect part of the reduction of the seasonal amplitude of NEE in boreal  
17 ecosystems and C3 grasslands noticed at Alert and South Pole stations. The degraded model-data fit  
18 between optimised  $C_{CO_2}$  and data from Mauna Loa also suggests that the boreal correction of NEE  
19 amplitude is too strong or insufficiently compensated at large scale by the amplification of the  
20 seasonal cycle at temperate latitudes visible at TempENF and TempDBF sites (Figs. 1C and 3A).

21  
22 Finally, using the multi-site-optimised model overall brings a small improvement of the  
23 modelled interannual variability of  $C_{CO_2}$ , with a median reduction of 3.9% for the RMSD between  
24 modelled and measured monthly anomaly (not shown). Forty-five locations, out of the 53 used in  
25 this study, display an improvement with RMSD decreases up to 27%, while at the remaining 8 sites  
26 the degradation of the simulated interannual variability remains small with RMSD increases always  
27 smaller than 1.5% (not shown). These results suggest that despite the relative shortness (one to three  
28 years) of most of the FluxNet datasets selected to optimise the ORCHIDEE model, the diversity of  
29 the covered weather situations gives a modest, yet consistent source of information to better  
30 reproduce interannual variations of carbon fluxes at the global scale.

31

### 32 **3.5.2 Global scale phenology index**

33

34 Figure 7 reports for each optimised PFT the correlation factor between weekly values of  
35 measured NDVI and modelled FAPAR during the period 2000-2010 (see sect. 2.4), for both the



1 prior and optimised model. There is no result for BorDBF whose vegetation fraction never exceeds  
2 40% in our case. All remaining six PFTs exhibit a higher median correlation factor when using the  
3 multi-site parameterization, which means that the modelled leaf seasonal cycle better matches the  
4 global scale observations. This median improvement seems to accurately reflect the overall trend  
5 for TempDBF-, BorENF- and C3grass-dominated pixels, while a larger inter-pixel variability is  
6 introduced in the case of temperate evergreen forests. The improved modelled seasonality is related  
7 to the more accurately simulated GPP at FluxNet sites after multi-site optimization, the latter being  
8 in turn partly driven by the improvement of the seasonal variations of simulated LAI. The dominant  
9 feature seems to be a shorter growing season length for TempDBF, which is consistent with the  
10 site-level simulations of GPP seasonality for this PFT (Fig. 5A), and an earlier beginning of the  
11 growing season for C3 grasses (not shown). Note that this improvement also explains most of the  
12 increased correlation factors in temperate and boreal evergreen forests, since these PFTs do not  
13 present a climate-driven leaf phenology in the current formulation of the ORCHIDEE model.  
14 Consequently, deciduous and herbaceous PFTs are the only significant contributors to the seasonal  
15 cycle at such a coarse resolution, even when these ecosystems are secondary and/or the understory  
16 within an evergreen-dominated forest. Lastly, the score for TropEBF remains poor because the  
17 model wrongly simulates the leaf renewal and the hydric stress during the dry season, as discussed  
18 in Sects. 3.1 and 3.4.

19

### 20 **3.6 Limitations of the current approach: summary and discussion**

21

22 The limitations to our model-data fusion method highlighted throughout the results section  
23 are of three kinds, somewhat interlocked: 1) within the limits of the model structure, 2) how  
24 adequate the chosen set of optimised parameters was and 3) how close to the optimal values the  
25 optimization algorithm tuned these parameters.

26 Taking these items in reverse order, we first acknowledge that using a variational  
27 optimization algorithm with a model with non-linearities might expose to miss the global minimum  
28 of the cost function, and indeed a few obvious convergence failures cases have been found for some  
29 single-site optimizations in TropEBF, TempENF, and boreal forests. Some functions of the  
30 ORCHIDEE model could potentially be linearized to generate a more accurate tangent linear model  
31 –and to advantageously avoid to use finite-differences for some phenological parameters (see Sect.  
32 2.1)–, while remaining coherent with the model's philosophy. It might imply a demanding effort of  
33 model recoding, but it has already been done for another LSM (Knorr et al., 2010). Alternatively,  
34 stochastic optimization approaches could yield better convergence, as they can circumvent the  
35 linearity constraint. While a single-site model-data fusion study with the same LSM showed

1 advantageous results for a genetic Monte-Carlo-based technique over its variational counterpart  
2 (Santaren et al., 2013), no major difference was found by (Ziehn et al., 2012) between Monte-Carlo  
3 and gradient-based approaches when optimizing a simpler LSM with atmospheric CO<sub>2</sub>  
4 observations. In the case of a multi-site optimization efforts, we suggest that the cost function  
5 “smoothing” discussed in Sect. 3.2 could make the convergence efficiency less sensitive to the  
6 choice of the minimization approach, while keeping in mind the much lower computational time  
7 required in the variational case.

8         Second, the number of optimised parameters remains somewhat modest as compared to the  
9 diversity of processes modelled in the ORCHIDEE model. Our choice was partly driven by a model  
10 sensitivity criterion, while the actual leverage of an optimised parameter on model outputs also  
11 depends on the uncertainty associated to this very parameter (Dietze et al., 2014). It can result in  
12 selecting some parameters that are already reasonably well known but that have medium-to-high  
13 model sensitivity and thus with low overall leverage, while poorly known parameters with  
14 mild-to-low model sensitivity could have a comparatively higher value for the optimization. In  
15 addition, as our focus was on the carbon cycle, only a few water-and-energy-related parameters  
16 were considered. Notably, the correction of LE partly benefited from that of NEE via transpiration,  
17 but the soil evaporation optimization was neglected despite being a significant -and debated- player  
18 of the terrestrial water cycle (Schlesinger and Jasechko, 2014).

19         The third hindering factor to simulating carbon and water fluxes close to their true value is  
20 the “observation error”, i.e. the uncertainty arising from the simplification needed to make  
21 ecosystem functioning fit within explicit equations plus the error made associated to the  
22 measurements, fluxes and meteorological forcing included. Although this error is rarely quantified  
23 in model-data fusion efforts, model-data fit analyses and uncertainty budgets showed in this study  
24 that the relative importance of this observation error greatly varies from one PFT to another –and is  
25 potentially dominated by the model error component in the case simulations at flux towers sites  
26 (Kuppel et al., 2013). It is the highest in tropical evergreen broadleaf forests, where parameter  
27 optimization will likely be of limited help until a more realistic phenological scheme is  
28 implemented. Regarding the simulations of LE in general, the small amount of related parameters  
29 optimised makes it difficult to assess to which extent the nearly-unchanged flux uncertainty comes  
30 from the parameter scarcity or structural inaccuracies in the model, stressing again the need for a  
31 better consideration of water and energy cycles together with that of carbon in future model-data  
32 fusion efforts.

33

#### 34 **4 Conclusions**

35

1 Generalizing the results of (Kuppel et al., 2012) across ecosystems, this study has shown that a  
2 significant degree of improvement is introduced to the simulation of carbon and water fluxes,  
3 through a generic optimization approach with in situ measurements of NEE and LE fluxes, relying  
4 on the traditional PFT classification used in many land surface models. At the global scale, this  
5 optimization method allows first a better simulation of the seasonal foliar cover. Second, the  
6 multi-site parameter set has a significant leverage upon the simulated seasonality of atmospheric  
7  $C_{CO_2}$ , with performances somewhat spatially heterogeneous and depending on the PFT considered,  
8 while a small, yet encouraging improvement of the simulated interannuality of  $C_{CO_2}$  is found. The  
9 remaining discrepancies in  $C_{CO_2}$  indicate that combining atmospheric CO<sub>2</sub> concentration and a  
10 larger number of flux towers observations, in a Carbon Cycle multi-Data Assimilation Systems  
11 (CCDASs, e.g., (Kaminski et al., 2013)), would be beneficial. Using more site-years of flux data  
12 will also allow a systematic in situ evaluation of the multi-site parameters across time periods,  
13 regions and climate regimes by separating training sites from evaluation sites. Such procedure was  
14 not applied in this study due to the small number of sites for some PFTs, but remains essential to  
15 test a LSM used for climate projections. More generally, we suggest that the assimilation of  
16 FluxNet data should be considered as a baseline for the development of multi-data assimilation  
17 systems where more complementary data streams are combined. In particular, daytime and  
18 nighttime NEE could replace the daily values used here, and adding measurements of leaf area  
19 index, soil respiration fluxes (e.g. chamber measurements), biomass and litter/soil carbon pools,  
20 would help better separating the processes and constraining environmental drivers, as would a  
21 simultaneous parameter optimization of both over- and understory PFT fractions. The FluxNet  
22 multi-site approach can also be used to characterize the structural, parametric and total uncertainties  
23 associated with the simulated annual biospheric carbon balance at regional-to-global scales, and to  
24 compare it with 1) the discrepancies of results between global ecosystem models (Sitch et al.,  
25 2008), and 2) the error carried by the terrestrial carbon fluxes estimated via inverse modeling with  
26 atmospheric transport models (e.g., (Chevallier et al., 2010)). The underlying problem is thus to  
27 evaluate what would be gained from simultaneously assimilating various data streams covering  
28 different spatial and temporal scales into a terrestrial ecosystem model, and how the PFT  
29 classification should be refined to maximize this improvement. In parallel, by using a diagonal prior  
30 covariance matrix for parameter error, within a same PFT and across PFTs, we implicitly assumed  
31 that all parameters could in principle be efficiently corrected as independent random distributions. It  
32 ignores the fact that a covariance structure interlinking the optimised parameterization would be  
33 necessary to translate the interconnectedness of ecophysiological processes within a given PFT. For  
34 instance, the allocation of carbon within the plant reservoirs depends on specific allometric relations  
35 and on photosynthesis rate; these relations would need to be embedded in the prior parameter error

1 covariance matrix. Additionally, the influence of nearby individuals of other PFTs (e.g., the  
2 understory) should be accounted for when correcting parameters of a given PFT. Together with a  
3 simultaneous optimization of several PFTs, building standard spatialized parameter covariance  
4 tables from databases of plant traits and soil characteristics (e.g., (Kattge et al., 2011)) and  
5 'preliminary' posterior multi-site parameter error covariance matrices (e.g., supplementary material  
6 of (Kuppel et al., 2012)) might soon become necessary to consistently apply model-data fusion to  
7 more sophisticated mechanistic ecosystem models.

8

## 9 **5 Code availability**

10

11 The source code of the data assimilation system is available at  
12 <https://pypi.python.org/pypi/ORCHISM>. Regarding the ORCHIDEE vegetation model, the source  
13 files of the Tag version 1.9.5.2 used for this study can be obtained upon request (see  
14 <http://labex.ipsl.fr/orchidee/index.php/contact>), while the associated documentation can be found at  
15 <https://forge.ipsl.jussieu.fr/orchidee/wiki/Documentation>. Note that the tangent linear version of the  
16 ORCHIDEE model has been generated using a commercial software (TAF, see sect. 2.2). For this  
17 reason, only the 'forward' version of the ORCHIDEE model is available for sharing, to which only  
18 the finite differences method is employed for parameter optimization. Finally, the source code of the  
19 LMDz atmospheric transport model can be found at <http://web.lmd.jussieu.fr/trac>.

20

## 21 **Acknowledgements**

22

23 This work has been supported by the CARBONES project, within the EU's 7th Framework  
24 Program for Research and Development. E. Vangorsel, J. Beringer, H.R. da Rocha, J. Grace, B.  
25 Kruijt, M. Goulden, A. Black, L. Flanagan, B. Amiro, J.H. McCaughey, H. Margolis, A. Barr, A.  
26 Arain, T. Kato, T. Foken, A. Knohl, A. Don, C. Rebmann, M.-J. Sanz, M.T. Sebastiá, T. Vesala, T.  
27 Laurila, E. Dufrêne, A. Granier, D. Loustau, N. Zoltan, T. Hirano, G. Kiely, R. Valentini, D.  
28 Gianelle, A. Raschi, H. Kondo, A.J. Dolman, T. Johansson, M. Nilsson, A. Lindroth, A. Grelle, J.  
29 Moncrieff, M. Wilkinson, M. Torn, T. Meyers, A.D. Richardson, J. Randers, J.W. Munger, D.  
30 Hollinger, R. Matamala, W.C. Oechel, J. Hadley, B. Law, L. Gu, A. Noormets, P. Blanken, G.  
31 Bohrer, D.D. Baldocchi, P. Bolstad, and K. Bible are warmly thanked for making their flux  
32 measurements available. So are M. Grünwald, E.J. Moors, D. Schimel, M. Smith, G. Baldi and  
33 three anonymous reviewers for their useful comments on earlier versions of the manuscript, and G.  
34 Lasslop, M. Forkel, N. Carvalhais and M. Reichstein for applying the energy-balance corrections to  
35 the processed data. We are grateful to C. Bacour for the help while developing the optimization and

1 the useful comments on the manuscript, to N. MacBean for setting up the global simulations, and to  
2 N. Viovy for the general discussions on the ORCHIDEE model. Finally, we would like to thank the  
3 computer team at LSCE for the computational resources provided.

4

## 5 **References**

6

Baldocchi, D.: TURNER REVIEW No. 15. 'Breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems, *Aust. J. Bot.*, 56(1), 1–26, 2008.

Botta, A., Viovy, N., Ciais, P., Friedlingstein, P. and Monfray, P.: A global prognostic scheme of leaf onset using satellite data, *Glob. Change Biol.*, 6(7), 709–725, 2000.

Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C.: A limited memory algorithm for bound constrained optimization, *SIAM J. Sci. Comput.*, 16(5), 1190–1208, 1995.

Chave, J., Navarrete, D., Almeida, S., Alvarez, E., Aragão, L. E., Bonal, D., Châtelet, P., Silva-Espejo, J. E., Goret, J.-Y., Hildebrand, P. von and others: Regional and seasonal patterns of litterfall in tropical South America, *Biogeosciences*, 7(1), 43–55, 2010.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M. and others: CO<sub>2</sub> surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, *J. Geophys. Res. Atmospheres* 1984–2012, 115(D21), 2010.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P. and others: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, 137(656), 553–597, 2011.

Desroziers, G., Berre, L., Chapnik, B. and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, *Q. J. R. Meteorol. Soc.*, 131(613), 3385–3396, 2005.

Dietze, M. C., Serbin, S. P., Davidson, C., Desai, A. R., Feng, X., Kelly, R., Kooper, R., LeBauer, D., Mantooh, J., McHenry, K. and others: A quantitative assessment of a terrestrial biosphere model's data needs across North American biomes, *J. Geophys. Res. Biogeosciences*, 119(3), 286–300, 2014.

Ducoudré, N. I., Laval, K. and Perrier, A.: SECHIBA, a new set of parameterizations of the hydrologic exchanges at the land-atmosphere interface within the LMD atmospheric general circulation model, *J. Clim.*, 6(2), 248–273, 1993.

Eidenshink, J. C. and Faundeen, J. L.: The 1 km AVHRR global land data set: first stages in implementation, *Int. J. Remote Sens.*, 15(17), 3443–3462, 1994.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., Myneni, R. B. and others: Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, *J. Geophys. Res. Biogeosciences* 2005–2012, 113(G2), 2008.

Giering, R., Kaminski, T. and Slawig, T.: Generating efficient derivative code with TAF: adjoint and

tangent linear Euler flow around an airfoil, *Future Gener. Comput. Syst.*, 21(8), 1345–1355, 2005.

GLOBALVIEW-CO2: Cooperative Global Atmospheric Data Integration Project. Multi-laboratory compilation of synchronized and gap-filled atmospheric carbon dioxide records for the period 1979-2012 (obspack\_co2\_1\_GLOBALVIEW-CO2\_2013\_v1.0.4\_2013-12-23). Compiled by NOAA Global Monitoring Division: Boulder, Colorado, U.S.A. Data product accessed at <http://dx.doi.org/10.3334/OBSPACK/1002>, 2013.

Groenendijk, M., Dolman, A. J., van der Molen, M. K., Leuning, R., Arneeth, A., Delpierre, N., Gash, J. H. C., Lindroth, A., Richardson, A. D., Verbeeck, H. and Wohlfahrt, G.: Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data, *Agric. For. Meteorol.*, 151(1), 22–38, doi:10.1016/j.agrformet.2010.08.013, 2011.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., CHEN, Y.-H., Ciais, P., Fan, S. and others: TransCom 3 CO2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus B*, 55(2), 555–579, 2003.

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y. and others: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, *Clim. Dyn.*, 27(7-8), 787–813, 2006.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R. and others: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, *J. Geophys. Res. Biogeosciences*, 118(4), 1414–1426, 2013.

Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J. and others: TRY—a global database of plant traits, *Glob. Change Biol.*, 17(9), 2905–2935, 2011.

Keenan, T. F., Davidson, E., Moffat, A. M., Munger, W. and Richardson, A. D.: Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling, *Glob. Change Biol.*, 18(8), 2555–2569, 2012.

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R. and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, *J. Geophys. Res. Biogeosciences* 2005–2012, 115(G4), 2010.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Glob. Biogeochem. Cycles*, 19(1), 2005.

Kuppel, S., Chevallier, F. and Peylin, P.: Quantifying the model structural error in carbon cycle data assimilation systems, *Geosci. Model Dev.*, 6(1), 45–55, 2013.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F. and Richardson, A. D.: Constraining a global ecosystem model with multi-site eddy-covariance data, *Biogeosciences*, 9(10), 3757–3776, 2012.

Lasslop, G., Reichstein, M., Kattge, J. and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, *Biogeosciences*, 5, 1311–1324, 2008.

Maignan, F., Bréon, F.-M., Chevallier, F., Viovy, N., Ciais, P., Garrec, C., Trules, J. and Mancip, M.: Evaluation of a Global Vegetation Model using time series of satellite vegetation indices, *Geosci. Model Dev.*, 4(4), 1103–1114, 2011.

Medvigy, D. and Moorcroft, P. R.: Predicting ecosystem dynamics at regional scales: an evaluation of a terrestrial biosphere model for the forests of northeastern North America, *Philos. Trans. R. Soc. B Biol. Sci.*, 367(1586), 222–235, 2012.

Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y. and Moorcroft, P. R.: Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2, *J. Geophys. Res. Biogeosciences* 2005–2012, 114(G1), 2009.

Miguez-Macho, G. and Fan, Y.: The role of groundwater in the Amazon water cycle: 2. Influence on seasonal soil moisture and evapotranspiration, *J. Geophys. Res. Atmospheres* 1984–2012, 117(D15), 2012.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I—A discussion of principles, *J. Hydrol.*, 10(3), 282–290, 1970.

Olson, J. S.: Global ecosystem framework-definitions, USGS EROS Data Cent. Intern. Rep. Sioux Falls SD, 37, 1994, 1994.

Papale, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, 2006.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A. and others: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, *Glob. Change Biol.*, 11(9), 1424–1439, 2005.

Richardson, A. D., Mahecha, M. D., Falge, E., Kattge, J., Moffat, A. M., Papale, D., Reichstein, M., Stauch, V. J., Braswell, B. H., Churkina, G., Kruijt, B. and Hollinger, D. Y.: Statistical properties of random CO<sub>2</sub> flux measurement uncertainty inferred from model residuals, *Agric. For. Meteorol.*, 148(1), 38–50, doi:10.1016/j.agrformet.2007.09.001, 2008.

Santaren, D., Peylin, P., Bacour, C., Ciais, P. and Longdoz, B.: Ecosystem model optimization using in-situ flux observations: benefit of monte-carlo vs. variational schemes and analyses of the year-to-year model performances, *Biogeosciences Discuss.*, 10(11), 18009–18064, 2013.

Schlesinger, W. H. and Jasechko, S.: Transpiration in the global water cycle, *Agric. For. Meteorol.*, 189, 115–117, 2014.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P. and others: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), *Glob. Change Biol.*, 14(9), 2015–2039, 2008.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C. and others: Climatological mean and decadal change in surface ocean pCO<sub>2</sub>, and net sea–air CO<sub>2</sub> flux over the global oceans, *Deep Sea Res. Part II Top. Stud. Oceanogr.*, 56(8), 554–577, 2009.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, Siam, 2005.

- Thoning, K. W., Tans, P. P. and Komhyr, W. D.: Atmospheric carbon dioxide at Mauna Loa Observatory: 2. Analysis of the NOAA GMCC data, 1974–1985, *J. Geophys. Res. Atmospheres* 1984–2012, 94(D6), 8549–8565, 1989.
- Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, Pr., Meyers, T. P., Prueger, J. H., Starks, P. J. and Wesely, M. L.: Correcting eddy-covariance flux underestimates over a grassland, *Agric. For. Meteorol.*, 103(3), 279–300, 2000.
- Verbeeck, H., Peylin, P., Bacour, C., Bonal, D., Steppe, K. and Ciais, P.: Seasonal patterns of CO<sub>2</sub> fluxes in Amazon forests: fusion of eddy covariance data and the ORCHIDEE model, *J. Geophys. Res. Biogeosciences* 2005–2012, 116(G2), 2011.
- Vermote, E. F., El Saleous, N. Z. and Justice, C. O.: Atmospheric correction of MODIS data in the visible to middle infrared: first results, *Remote Sens. Environ.*, 83(1), 97–111, 2002.
- Vermote, E., Justice, C. O. and Breon, F.-M.: Towards a generalized approach for correction of the BRDF effect in MODIS directional reflectances, *Geosci. Remote Sens. IEEE Trans. On*, 47(3), 898–908, 2009.
- Weirdt, M. D., Verbeeck, H., Maignan, F., Peylin, P., Poulter, B., Bonal, D., Ciais, P. and Steppe, K.: Seasonal leaf dynamics for tropical evergreen forests in a process-based global ecosystem model, *Geosci. Model Dev.*, 5(5), 1091–1108, 2012.
- Xu, T., White, L., Hui, D. and Luo, Y.: Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction, *Glob. Biogeochem. Cycles*, 20(2), 2006.
- Ziehn, T., Scholze, M. and Knorr, W.: On the capability of Monte Carlo and adjoint inversion techniques to derive posterior parameter uncertainties in terrestrial ecosystem models, *Glob. Biogeochem. Cycles*, 26(3), 2012.



Table 1. Parameters of ORCHIDEE optimised in this study. The prior values are given for each PFT, and multi-site posterior values are in bold font. A hyphen means that the parameter is not optimised, *spinup* that the spinup value is taken, and *site* that the posterior value is site-specific.

Parameter	Description	Plant functional type <sup>a</sup>						
		Trop	Temp	Temp	Temp	Bor	Bor	C3
		EBF	ENF	EBF	DBF	ENF	DBF	grass
<b>Photosynthesis</b>								
$V_{cmax}$	Maximum carboxylation rate	65	35	45	55	35	45	70
	( $\mu\text{mol.m}^{-2}.\text{s}^{-1}$ )	<b>70.28</b>	<b>31.94</b>	<b>47.84</b>	<b>55.83</b>	<b>32.36</b>	<b>32.97</b>	<b>51.10</b>
$G_{s,slope}$	Ball-Berry slope	9	9	9	9	9	9	9
		<b>8.756</b>	<b>8.841</b>	<b>10.99</b>	<b>6.000</b>	<b>7.961</b>	<b>7.714</b>	<b>9.970</b>
$c_{Tmax}$	Offset controlling the maximum	55	38	48	38	38	38	41.13
	photosynthesis temperature ( $^{\circ}\text{C}$ )	<b>55.31</b>	<b>40.41</b>	<b>49.66</b>	<b>36.09</b>	<b>36.42</b>	<b>36.70</b>	<b>40.20</b>
$c_{Topt}$	Offset controlling the optimal	37	25	32	26	25	25	27.25
	photosynthesis temperature ( $^{\circ}\text{C}$ )	<b>35.93</b>	<b>17.49</b>	<b>28.82</b>	<b>28.44</b>	<b>26.48</b>	<b>28.71</b>	<b>29.76</b>
$c_{Tmin}$	Offset controlling the minimal	2	-4	-3	-2	-4	-4	-3.25
	photosynthesis temperature ( $^{\circ}\text{C}$ )	<b>1.356</b>	<b>-7.536</b>	<b>-6.062</b>	<b>-0.219</b>	<b>-6.167</b>	<b>-2.563</b>	<b>-3.403</b>
<b>Phenology</b>								
$SLA$	Specific leaf area (foliar surface per dry	0.0154	0.0093	0.02	0.026	0.0093	0.0260	0.0260
	matter content, $\text{m}^2.\text{g}^{-1}$ )	<b>0.0169</b>	<b>0.0200</b>	<b>0.0252</b>	<b>0.0400</b>	<b>0.0090</b>	<b>0.0233</b>	<b>0.0345</b>
$LAI_{MAX}$	Maximum LAI ( $\text{m}^2.\text{m}^{-2}$ )	7	5	5	5	4.5	4.5	4.5
		<b>7.000</b>	<b>5.000</b>	<b>5.000</b>	<b>3.949</b>	<b>4.500</b>	<b>4.960</b>	<b>2.349</b>
$K_{lai, happy}$	Minimum fraction of $LAI_{MAX}$ to stop	0.5	0.5	0.5	0.5	0.5	0.5	0.5
	carbohydrate use	<b>0.500</b>	<b>0.500</b>	<b>0.500</b>	<b>0.321</b>	<b>0.500</b>	<b>0.547</b>	<b>0.408</b>
$K_{pheno, crit}$	Multiplicative factor for growing season	-	-	-	1	-	1	1
	start threshold				<b>1.510</b>		<b>0.758</b>	<b>0.729</b>
$c_{T, senes}$	Offset controlling the temperature	-	-	-	12	-	7	-
	threshold for senescence ( $^{\circ}\text{C}$ )				<b>14.36</b>		<b>7.899</b>	
$L_{agecrit}$	Critical age for leaves (days)	730	910	730	180	910	180	120
		<b>717.9</b>	<b>1084</b>	<b>709.2</b>	<b>165.1</b>	<b>790.5</b>	<b>163.3</b>	<b>113.9</b>
$LAI_{init}$	Initial LAI ( $\text{m}^2.\text{m}^{-2}$ )	<i>spinup</i>	<i>spinup</i>	<i>spinup</i>	-	<i>spinup</i>	-	<i>spinup</i>
		<b>site</b>	<b>site</b>	<b>site</b>		<b>site</b>		<b>site</b>
<b>Soil water availability</b>								
$f_{stressh}$	Parameter reducing the hydric limitation	6	6	6	6	6	6	6
	of photosynthesis	<b>6.507</b>	<b>7.146</b>	<b>7.135</b>	<b>5.039</b>	<b>4.881</b>	<b>5.505</b>	<b>5.131</b>
$Dpu_{cste}$	Total depth of the soil water reservoir (m)	2	2	2	2	2	2	2
		<b>2.377</b>	<b>2.387</b>	<b>1.536</b>	<b>0.959</b>	<b>2.012</b>	<b>2.303</b>	<b>1.865</b>
$Hum_{cste}$	Parameter describing the exponential root	0.8	1	0.8	0.8	1	1	4
	profile ( $\text{m}^{-1}$ )	<b>0.718</b>	<b>1.102</b>	<b>0.743</b>	<b>1.577</b>	<b>1.874</b>	<b>0.676</b>	<b>2.800</b>
<b>Autotrophic respiration</b>								
$MR_a$	Slope of the temperature dependence	0.16	0.16	0.16	0.16	0.16	0.16	0.16
		<b>0.105</b>	<b>0.127</b>	<b>0.156</b>	<b>0.094</b>	<b>0.185</b>	<b>0.178</b>	<b>0.174</b>
$MR_b$	Offset of the temperature dependence of	1	1	1	1	1	1	1

	maintenance respiration	<b>0.929</b>	<b>0.772</b>	<b>0.928</b>	<b>0.622</b>	<b>0.710</b>	<b>1.212</b>	<b>1.140</b>
$GR_{\text{frac}}$	Fraction of biomass available for growth	0.28	0.28	0.28	0.28	0.28	0.28	0.28
	respiration	<b>0.269</b>	<b>0.250</b>	<b>0.265</b>	<b>0.206</b>	<b>0.303</b>	<b>0.301</b>	<b>0.317</b>
<b><i>Heterotrophic respiration</i></b>								
$K_{\text{soilC}}$	Scaling factor for all initial soil carbon	1	1	1	1	1	1	1
	stocks after spinup	<b>site</b>	<b>site</b>	<b>site</b>	<b>site</b>	<b>site</b>	<b>site</b>	<b>site</b>
$Q_{10}$	Factor of the temperature control function	1.994	1.994	1.994	1.994	1.994	1.994	1.994
		<b>2.119</b>	<b>1.676</b>	<b>2.067</b>	<b>2.182</b>	<b>2.879</b>	<b>2.663</b>	<b>2.778</b>
$HR_{\text{H,b}}$	Parameter of the soil/litter moisture	2.4	2.4	2.4	2.4	2.4	2.4	2.4
	control function	<b>2.356</b>	<b>2.387</b>	<b>2.343</b>	<b>2.191</b>	<b>2.503</b>	<b>2.457</b>	<b>2.489</b>
$HR_{\text{H,c}}$	Offset of the soil/litter moisture control	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29
	function	<b>-0.332</b>	<b>-0.272</b>	<b>-0.329</b>	<b>-0.544</b>	<b>-0.192</b>	<b>-0.252</b>	<b>-0.304</b>
<b><i>Decomposition</i></b>								
$h_{\text{crit,litter}}$	Total litter height (m)	0.08	0.08	0.08	0.08	0.08	0.08	0.08
		<b>0.0697</b>	<b>0.0434</b>	<b>0.0613</b>	<b>0.0200</b>	<b>0.0213</b>	<b>0.114</b>	<b>0.0358</b>
$Z_{\text{decomp}}$	Factor of the exponential profile of soil	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	temperature and moisture	<b>0.371</b>	<b>0.649</b>	<b>0.175</b>	<b>0.142</b>	<b>0.662</b>	<b>0.474</b>	<b>0.448</b>
<b><i>Energy balance</i></b>								
$K_{\text{albedo,veg}}$	Multiplicative factor of surface albedo	1	1	1	1	1	1	1
		<b>1.031</b>	<b>1.042</b>	<b>0.930</b>	<b>1.110</b>	<b>1.076</b>	<b>1.048</b>	<b>0.989</b>
$Z0_{\text{overheight}}$	Reference roughness length (m)	0.0625	0.0625	0.0625	0.0625	0.0625	0.0625	0.0625
		<b>0.0648</b>	<b>0.0877</b>	<b>0.0359</b>	<b>0.0200</b>	<b>0.0200</b>	<b>0.0513</b>	<b>0.0200</b>

<sup>a</sup> TropEBF = tropical evergreen broadleaf forest; TempENF = temperate evergreen needleleaf forest; TempEBF = temperate evergreen broadleaf forest; TempDBF = temperate deciduous broadleaf forest; BorENF = boreal evergreen needleleaf forest; BorDBF = boreal deciduous broadleaf forest; C3grass = C3 grassland.

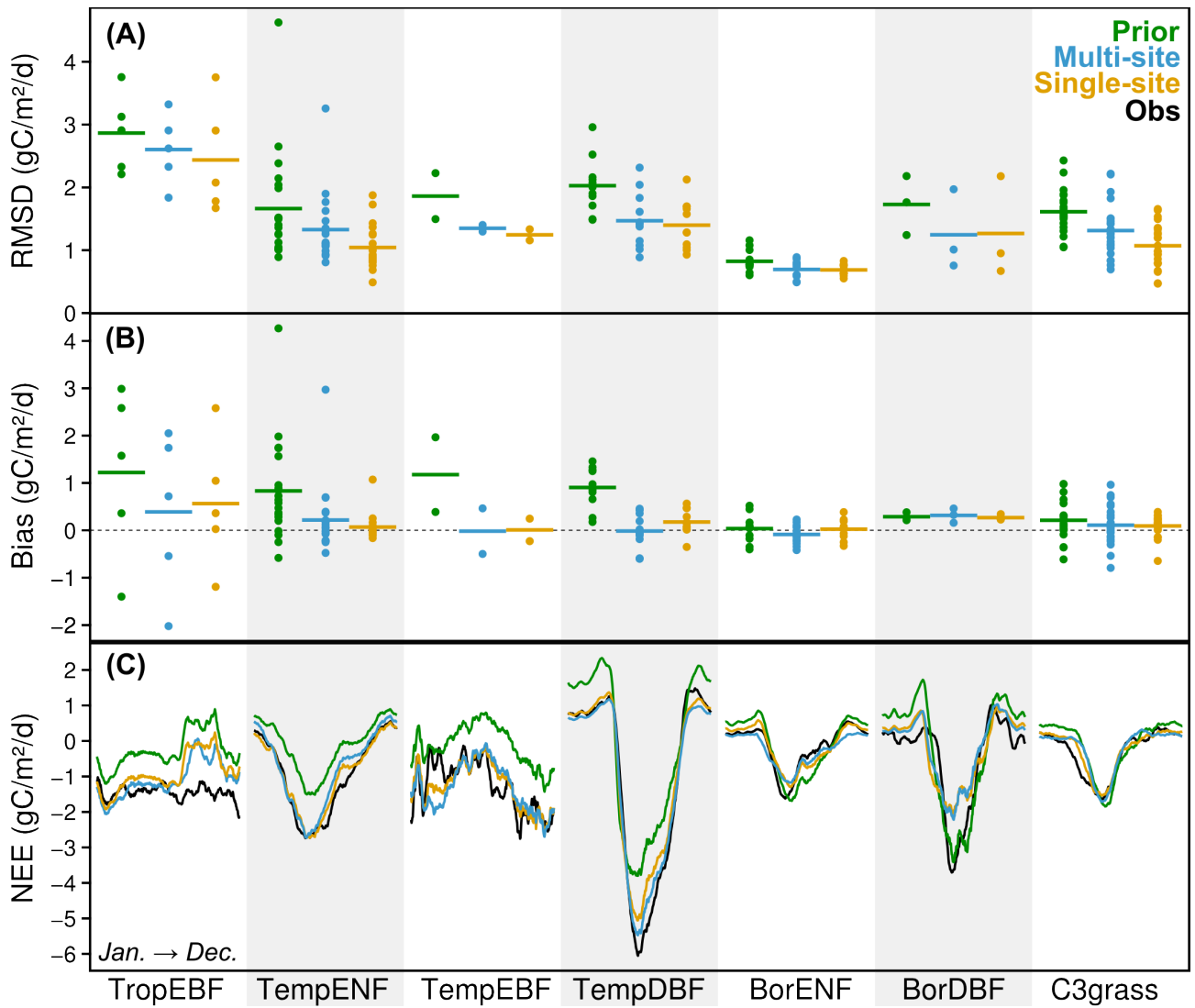


Figure 1. Model-data (A) RMSD and (B) bias for the daily NEE time series at each site (filled circles), grouped and averaged by PFT (horizontal bars), in three cases: prior model (green), multi-site optimization (blue) and single-site optimization (orange). (C) PFT-averaged mean seasonal cycle of NEE, for the training observations (black) and the three aforementioned cases, smoothed with a 15-day-moving-average window.

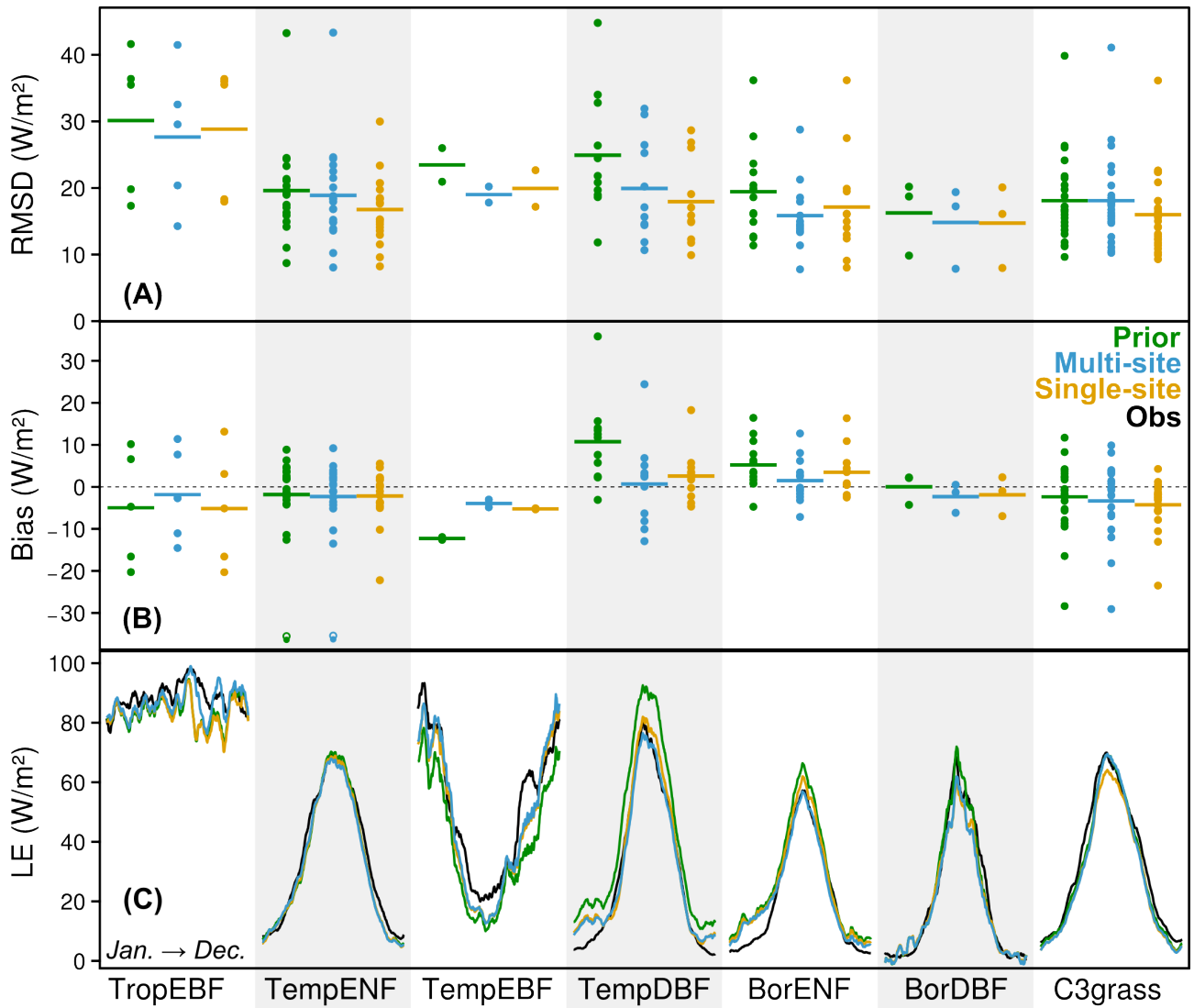


Figure 2. Model-data (A) RMSD and (B) bias for the daily LE time series at each site (filled circles), grouped and averaged by PFT (horizontal bars), in three cases: prior model (green), multi-site optimization (blue) and single-site optimization (orange). (C) PFT-averaged mean seasonal cycle of LE, for the training observations (black) and the three aforementioned cases, smoothed with a 15-day-moving-average window.

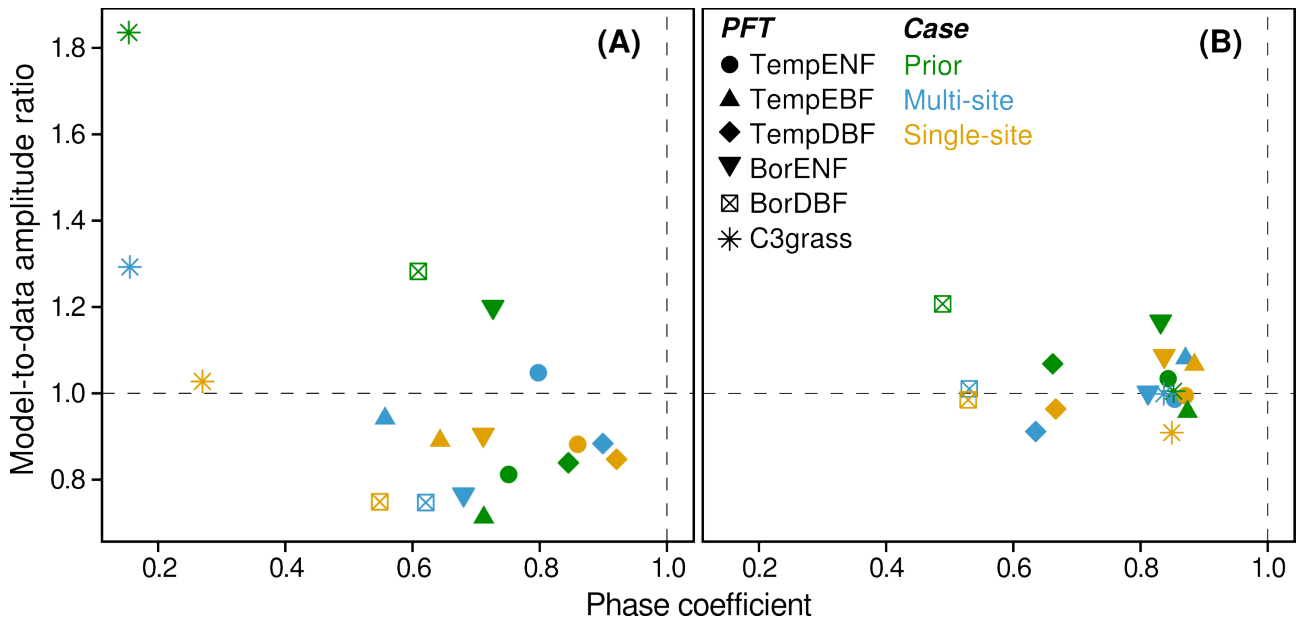


Figure 3. PFT-averaged model phase coefficient versus model-to-data amplitude ratio, for the detrended smooth seasonal cycles of (A) NEE and (B) LE fluxes. Simulations using prior parameters (green) are compared to multi-site (blue) and single-site (orange) optimizations, with the measured reference indicated by the intersection of the dashed lines.

1

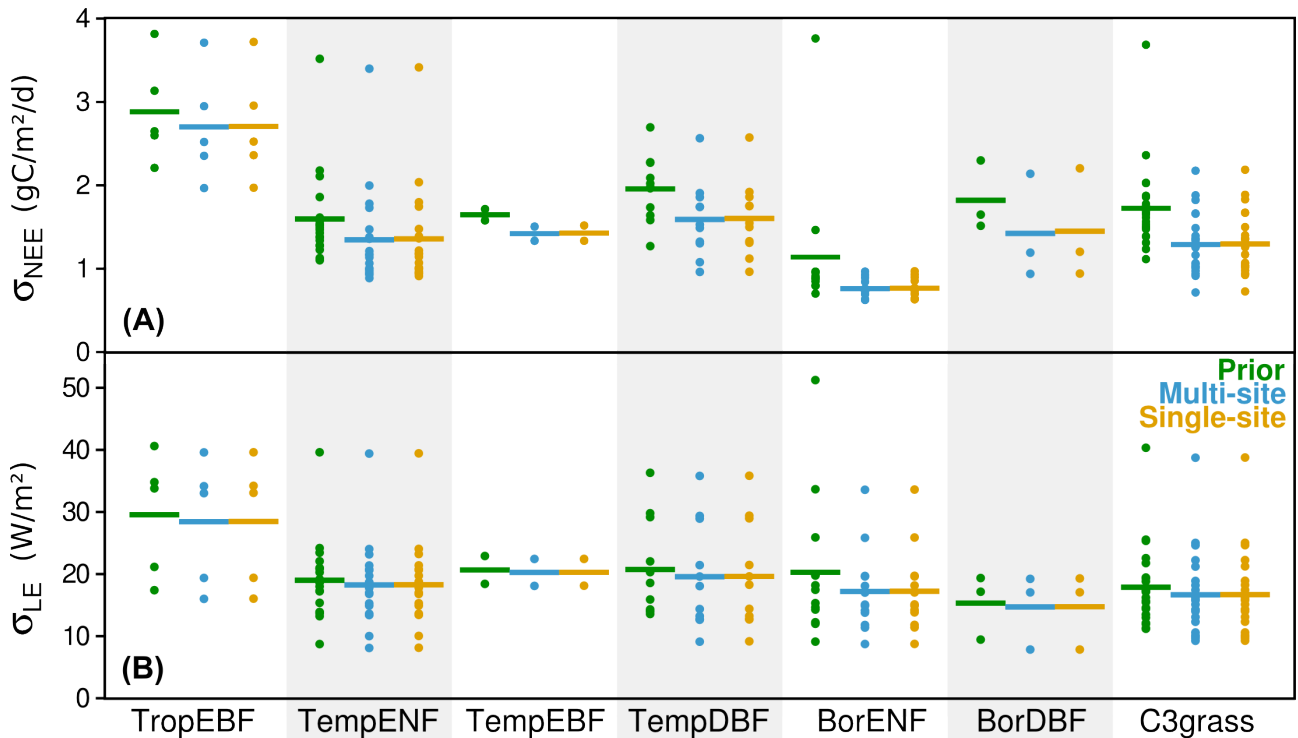


Figure 4. Uncertainty of simulated daily (A) NEE and (B) LE fluxes. For each PFT, the horizontal lines give the average of the individual site values (filled circles), in three cases: prior model (green), multi-site optimization (blue) and single-site optimization (orange).

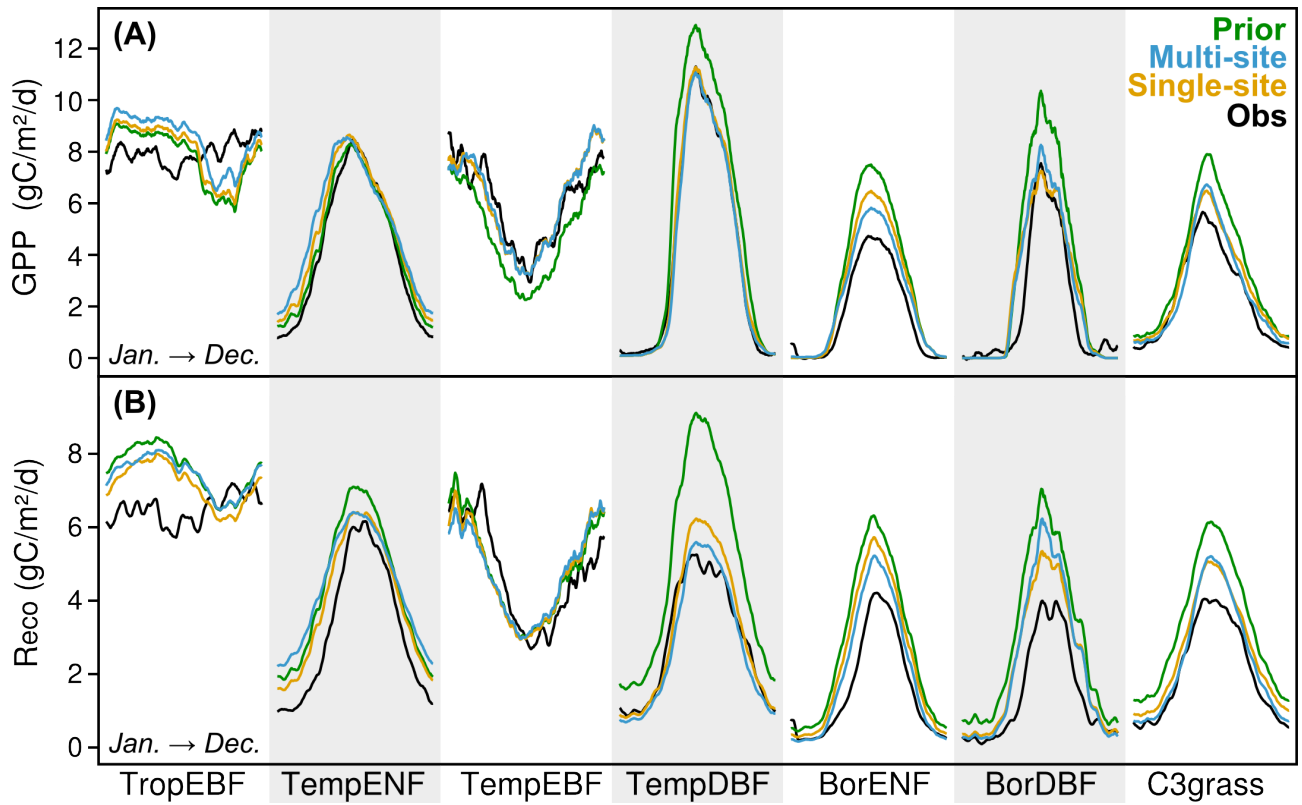


Figure 5. PFT-averaged mean seasonal cycles of **(A)** the photosynthetic carbon flux and **(B)** the respiration flux, smoothed with a 15-day-moving-average window. The simulations using prior (green), single-site (orange) and multi-site (blue) parameterizations are compared to the evaluative observation-derived flux estimates (black).

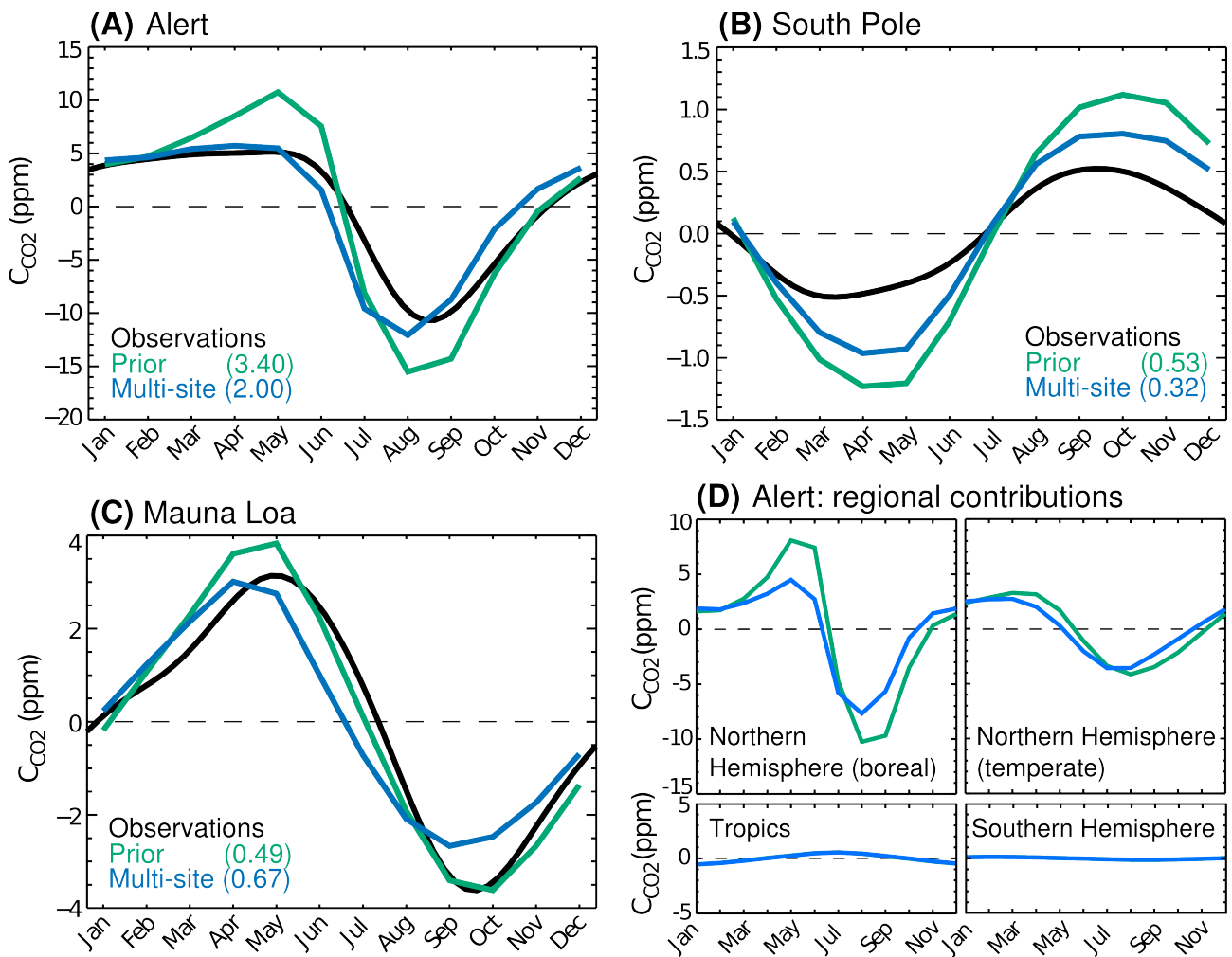
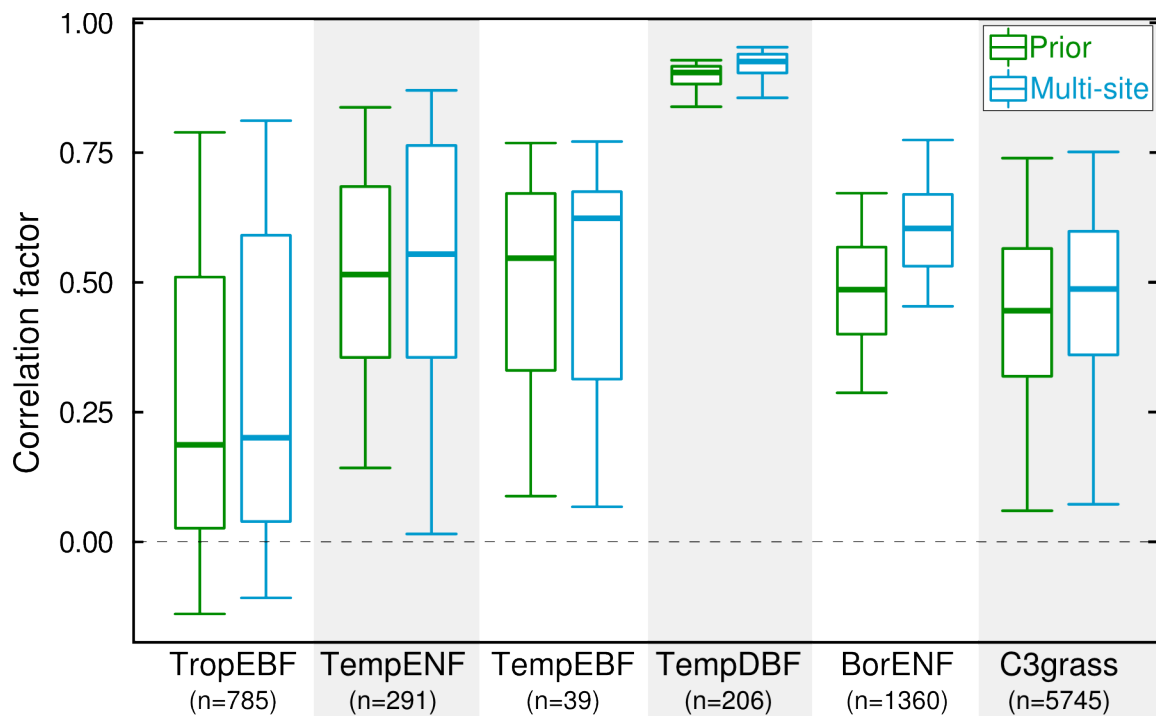


Figure 6. Detrended mean seasonal cycle of the atmospheric CO<sub>2</sub> concentrations at (A) Alert, (B) South Pole and (C) Mauna Loa locations during the 1989-2009 period: the optimization-independent concentrations records (black) are compared to simulations where the biospheric contribution is calculated using the ORCHIDEE model with default (green) and multi-site (blue) parameterization, with the model-data RMSD given between brackets. (D) Regional contributions to the mean seasonal cycle simulated at Alert.



1 Figure 7. Correlation factor between weekly time series of modelled FAPAR and independent  
 2 measurements of NDVI, for the 2000-2010 period. The results are grouped using the dominant PFT  
 3 at each pixel, for global simulations with default (green) and multi-site parameterization (blue). The  
 4 central horizontal bar indicates the median value, the top and bottom of the boxes correspond to the  
 5 first and last quartile, and the 5- and 95-percentile are given by the 'error bars'.