Improving estimates of gross primary productivity by assimilating solar-induced fluorescence satellite retrievals in a terrestrial biosphere model using a process-based SIF model

C. Bacour, F. Maignan, N. Macbean, A. Porcar-castell, J. Flexas, C. Frankenberg, Philippe Peylin, F. Chevallier, N. Vuichard, V. Bastrikov

JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2019JG005040

Improving Estimates of Gross Primary Productivity by Assimilating Solar-Induced Fluorescence Satellite Retrievals in a Terrestrial Biosphere Model Using a Process-Based SIF Model

C. Bacour[1], F. Maignan[2], N. MacBean[3], A. Porcar-Castell[4], J. Flexas[5], C. Frankenberg[6,7], P. Peylin[2], F. Chevallier[2], N. Vuichard[2], and V. Bastrikov[2]

[1]NOVELTIS, Labège, France, [2]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif sur Yvette, France, [3]Department of Geography, Indiana University, Bloomington, IN, USA, [4]Optics of Photosynthesis Laboratory, Institute for Atmospheric and Earth System Research/Forest Sciences, University of Helsinki, Helsinki, Finland, [5]Research Group on Plant Biology under Mediterranean Conditions, Universitat de les Illes Balears - Instituto de Investigaciones Agroambientales y de Economía del Agua, Palma, Spain, [6]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA, [7]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Key Points:
- We developed a process-based SIF observation operator in a terrestrial biosphere model
- We assimilated monthly OCO-2 SIF products to optimize model photosynthesis and phenology-related parameters
- The optimized GPP is considerably reduced with spatio-temporal patterns in closer agreement with independent products

Correspondence to:
C. Bacour,
cedric.bacour@noveltis.fr

Citation:
Bacour, C., Maignan, F., MacBean, N., Porcar-Castell, A., Flexas, J., Frankenberg, C., et al. (2019). Improving estimates of gross primary productivity by assimilating solar-induced fluorescence satellite retrievals in a terrestrial biosphere model using a process-based SIF model. *Journal of Geophysical Research: Biogeosciences*, *124*, 3281–3306. https://doi.org/10.1029/2019JG005040

Received 24 JAN 2019
Accepted 11 SEP 2019
Accepted article online 10 OCT 2019
Published online 7 NOV 2019

Author Contributions:
Conceptualization: C. Bacour
Data curation: F. Maignan, A. Porcar-Castell, J. Flexas
Formal analysis: C. Bacour, F. Maignan, N. MacBean, A. Porcar-Castell, J. Flexas, C. Frankenberg, P. Peylin
Funding acquisition: F. Maignan, F. Chevallier
Investigation: C. Bacour, F. Maignan
Methodology: C. Bacour, F. Maignan, N. MacBean, P. Peylin
Resources: F. Maignan, P. Peylin
Software: C. Bacour, F. Maignan, N. Vuichard, V. Bastrikov
Supervision: F. Maignan
(continued)

Abstract Over the last few years, solar-induced chlorophyll fluorescence (SIF) observations from space have emerged as a promising resource for evaluating the spatio-temporal distribution of gross primary productivity (GPP) simulated by global terrestrial biosphere models. SIF can be used to improve GPP simulations by optimizing critical model parameters through statistical Bayesian data assimilation techniques. A prerequisite is the availability of a functional link between GPP and SIF in terrestrial biosphere models. Here we present the development of a mechanistic SIF observation operator in the ORCHIDEE (Organizing Carbon and Hydrology In Dynamic Ecosystems) terrestrial biosphere model. It simulates the regulation of photosystem II fluorescence quantum yield at the leaf level thanks to a novel parameterization of non-photochemical quenching as a function of temperature, photosynthetically active radiation, and normalized quantum yield of photochemistry. It emulates the radiative transfer of chlorophyll fluorescence to the top of the canopy using a parametric simplification of the SCOPE (Soil Canopy Observation Photosynthesis Energy) model. We assimilate two years of monthly OCO-2 (Orbiting Carbon Observatory-2) SIF product at 0.5° (2015–2016) to optimize ORCHIDEE photosynthesis and phenological parameters over an ensemble of grid points for all plant functional types. The impact on the simulated GPP is considerable with a large decrease of the global scale budget by 28 GtC/year over the period 1990–2009. The optimized GPP budget (134/136 GtC/year over 1990–2009/2001–2009) remarkably agrees with independent GPP estimates, FLUXSAT (137 GtC/year over 2001–2009) in particular and FLUXCOM (121 GtC/year over 1990–2009). Our results also suggest a biome dependency of the SIF-GPP relationship that needs to be improved for some plant functional types.

1. Introduction

The terrestrial biosphere currently plays a pivotal role for climate by offsetting about one fourth of carbon dioxide (CO_2) emissions released by anthropogenic activities into the atmosphere (Le Quéré et al., 2018). Terrestrial ecosystems assimilate CO_2 in leaf chloroplasts by photosynthesis, but most of the assimilated carbon is released back into the atmosphere (mainly as CO_2) through ecosystem respiration. Although the net carbon budget of terrestrial ecosystems is the main quantity of interest for climate studies, quantifying the gross fluxes (photosynthesis and respiration) is crucial in order to better understand the drivers of the net fluxes. Current knowledge of the spatial and temporal distribution of net and gross carbon fluxes over the globe is mainly driven by atmospheric and ecosystem measurements, but our ability to anticipate their evolution under a changing climate largely relies on global terrestrial biosphere models (TBMs) (Sitch et al., 2015). TBMs numerically represent the different processes controlling the exchanges of trace gases (including CO_2), water, and energy, between the atmosphere and the biosphere in a simplified manner. The large uncertainties that remain in our understanding of the carbon sequestration in land ecosystems (Anav

Visualization: C. Bacour
Writing - original draft: C. Bacour
Writing – review & editing: C. Bacour, F. Maignan, N. MacBean, A. Porcar-Castell, J. Flexas, C. Frankenberg, P. Peylin, F. Chevallier, N. Vuichard

et al., 2015; Friedlingstein et al., 2014) lead to a large spread in TBM simulations, especially in terms of spatio-temporal patterns of the terrestrial gross primary productivity (GPP). In practice, the spread results from structural differences in the representation of key processes, that is, their mathematical formulation (Jung et al., 2007; Kuppel et al., 2013), and from the different fixed values assigned to the associated parameters. For instance, the impact on simulated GPP resulting from a plausible change in parameter values ranges from about 10 to 15 GtC/year in the global GPP budget for the Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) model (Peylin et al., 2016) to 3–4 times these values for the JSBACH model (Schürmann et al., 2016). The question of optimizing parameter values in TBMs from the existing limited number of observations, obtained at heterogeneous spatial scales (from the individual plant organs to the plant community) and under specific environmental conditions, has therefore grown in importance over the recent past and now involves sophisticated data assimilation approaches as well as simple heuristic ones.

Data assimilation (DA) techniques enable the numerical minimization of the mismatch between model predictions and corresponding independent observations, within prescribed confidence intervals, by adjusting the model parameters (Kaminski et al., 2002; Raupach et al., 2005). In addition to atmospheric CO_2 mole fraction measurements (Koffi et al., 2012; Peylin et al., 2016; Rayner et al., 2005; Schürmann et al., 2016), local scale eddy-covariance flux measurements (Fox et al., 2009; Groenendijk et al., 2011; Knorr & Kattge, 2005; Sacks et al., 2007; Wang et al., 2001; Williams et al., 2005) or remote sensing proxies of vegetation greenness like vegetation indices (MacBean et al., 2015; Migliavacca et al., 2009) and the fraction of absorbed photosynthetically active radiation (FAPAR) (Bacour et al., 2015; Kaminski et al., 2012; Knorr et al., 2010; Stöckli et al., 2008; Zobitz et al., 2014), are traditionally assimilated to constrain the value of the main model parameters controlling the carbon cycle. Eddy covariance flux and atmospheric CO_2 measurements provide ecosystem- to large-scale information on net ecosystem exchange. Vegetation indices (VI) or FAPAR space-borne products, on the other hand, provide information on the photosynthetic potential of the ecosystems—i.e. the "greenness" of the vegetation; however, these measures do not always give accurate information about the actual photosynthetic activity (Grace et al., 2007; Zhang et al., 2016) or the short-term variations of GPP (Yang et al., 2015), given that their estimates also depend on nongreen canopy elements (i.e., not only by chlorophyll) and are not directly related to the photosynthetic activity itself (Smith et al., 2018; Walther et al., 2016). Satellite retrievals of solar-induced fluorescence (SIF) over land surfaces have been available for nearly a decade and their use for model parameter optimization is starting. These observations exhibit a strong correlation with GPP at broad spatial and temporal scales (Frankenberg et al., 2011; Guanter et al., 2012; Zhang et al., 2016) and at site scale for different vegetation types (Joiner et al., 2013; Madani et al., 2017; Sanders et al., 2016; Yang et al., 2015). SIF tracks GPP seasonality better than traditional vegetation indices (Joiner et al., 2014; Smith et al., 2018; Walther et al., 2016) and provide complementary information on the drivers of plant phenology and physiology (Jeong et al., 2017). There is theoretical evidence to support the covariation of SIF and GPP. Chlorophyll fluorescence originates in the chloroplasts inside the leaf. It is strongly linked to CO_2 fixation (Meroni et al., 2009; Porcar-Castell et al., 2014) and is affected by regulation mechanisms modulating photosynthesis due to environmental conditions, for example, water availability (Flexas et al., 2002; Sun et al., 2015; Wang et al., 2016). The top of canopy SIF emission can be expressed as the product of the absorbed photosynthetically active radiation (APAR) by the fluorescence yield of the canopy Φ_F (also referred as the SIF light-use efficiency (Guanter et al., 2014) that accounts for these mechanisms occurring inside the leaves) modulated by the probability Ω_c that a SIF photon emitted within the canopy escapes to space (Frankenberg & Berry, 2018): $\mathrm{SIF} = \mathrm{APAR} \times \Phi_F \times \Omega_c$. This theoretical expression is analogous to the classical expression of GPP as a function of photosynthetic light use efficiency (LUE): $\mathrm{GPP} = \mathrm{APAR} \times \mathrm{LUE}$ (Guanter et al., 2014; Joiner et al., 2014).

Given the recent availability of SIF observations, modeling groups are thus adapting TBMs to take advantage of this GPP proxy in order to evaluate simulations of gross productivity (Lee et al., 2015; Thum et al., 2017). The next logical step is to assimilate the SIF data to optimize their parameters (Koffi et al., 2015; MacBean et al., 2018; Norton et al., 2018). Both model evaluation and optimization (via DA techniques) require the development of an observation operator that functionally relates the model variables (e.g., photosynthesis and fluorescence) and related parameters to the observed SIF.

Two different classes of method can be distinguished. The first simple approach is based on a statistical linear relationship between GPP and SIF (MacBean et al., 2018), usually defined at the level of plant functional type (PFT), the typical unit of vegetation description in TBMs. However, such simple relationship limits the

potential of SIF data to constrain photosynthetic parameters, as part of the SIF information is hidden by the statistical relationship. The second more process-based approach relies on the incorporation of (or some part of) the Soil Canopy Observation Photosynthesis Energy (SCOPE) model (van der Tol et al., 2009, 2014) in TBMs. SCOPE combines a fully prognostic leaf biochemical model and a one-dimensional (1-D) canopy radiative transfer model. SCOPE therefore enables simulations of chlorophyll fluorescence emitted by the canopy depending on the leaf biochemistry, canopy structure, meteorological conditions (including incoming radiation), and observation geometry. However, the use of SCOPE in TBMs is computationally expensive. Its implementation within the Biosphere Energy Transfer Hydrology (BETHY) model (Norton et al., 2018) has increased CPU computation expenses by more than 3 orders of magnitude in data assimilation runs (Norton, personal communication)—the major computational cost being attributed to the radiative transfer scheme. As a consequence, alternatives or simplified versions of such 1-D radiative transfer modeling scheme need to be incorporated into complex TBMs for efficient and computationally affordable assimilation purposes.

In this context, the objectives of our paper are the following:

- To describe a new simplified process-based model of canopy fluorescence implemented in the ORCHIDEE (Krinner et al., 2005) TBM. The approach relies on a new parametric model to estimate the total non-photochemical quenching (NPQ) and a simplification of the 1-D radiative transfer model in SCOPE to reduce the computational burden;
- To use the SIF observation operator to assimilate SIF retrievals from the Orbiting Carbon Observatory-2 (OCO-2) instrument (Frankenberg et al., 2014, Sun et al., 2018) in ORCHIDEE in order to constraint its main photosynthetic and phenology parameters;
- To assess the overall potential of the SIF data to optimize the photosynthesis of ORCHIDEE, comparing the prior and posterior (after the assimilation of SIF data) spatial and temporal variation of GPP against independent information.

The modeling approach of SIF in ORCHIDEE is described in section 2. The simulation setup, the data used for assimilation and evaluation of the optimized model, and the assimilation scheme are described in section 3. The results relative to assimilation are presented in section 4. In section 5, we then discuss the constraints brought by SIF on photosynthesis and phenology in TBMs, the limits of the approach, and propose strategies aiming at making the best use of SIF observations for reducing uncertainties in TBM parameterization.

2. Modeling SIF in ORCHIDEE

2.1. Overview of the ORCHIDEE Terrestrial Biosphere Model

ORCHIDEE is a mechanistic TBM that simulates the exchanges of carbon, water, and energy between biosphere and atmosphere (Krinner et al., 2005). It is the land surface component of the Earth System Model of Institut Pierre-Simon Laplace IPSL-CM (Dufresne et al., 2013) and therefore has been contributing to the recent exercises of the Coupled Model Intercomparison Project (CMPI) established by the World Climate Research Programme (https://www.wcrp-climate.org/wgcm-cmip). Photosynthesis and all components of the surface energy and water budgets are calculated at a half-hourly resolution while the dynamics of the carbon storage (including carbon allocation in plant reservoirs, soil carbon dynamics, and litter decomposition) are resolved on a daily basis. Photosynthesis depends on light availability and CO_2 concentration, soil moisture, and temperature and is parameterized based on Farquhar et al. (1980) and Collatz et al. (1992) for C3 and C4 plants, respectively. The formulations of Yin and Struik (2009) are used to describe the main photosynthesis processes, including the stomatal conductance and the rates of CO_2 assimilation under Rubisco and electron transport limitations and the intercellular CO_2 partial pressure (Vuichard et al., 2018). As in most TBMs, the spatial distribution of vegetation is represented using fractions of PFTs for each grid point (Cramer, 1997; Prentice et al., 1992; Wullschleger et al., 2014). Here we use a new land cover map for the year 2015 with 14 PFTs (among which are two representative C3 and C4 crop PFTs) in addition to bare soil. It is derived from the ESA CCI Land Cover (LC) products (see https://orchidas.lsce.ipsl.fr/dev/lccci/ for more details) and a LC-to-PFT cross-walking approach described in Poulter et al. (2015) (see Table 3 for the description of the PFTs).

We use the soil map from Zobler (1986) to prescribe the texture class in each grid cell. The impacts of land use change, forest management, harvesting, and fires are not included in this study. Except for the phenology (detailed in Botta et al. (2000) and MacBean et al. (2015)), processes are described with the same governing equations for all PFTs, but usually with different parameter values.

2.2. General SIF Modeling Approach

The modeling of SIF in ORCHIDEE relies on the formalism developed in SCOPE (v1.61) at the leaf level and transposed at the canopy scale, in order to calculate (1) the relative contribution of photosystems I and II (denoted PSI and PSII, respectively) to leaf level fluorescence and (2) the total fluorescence flux emitted at the top of the canopy (TOC). At TOC, the total fluorescence spectrum *SIF* (in $W.m^{-2}.\mu m^{-1}.sr^{-1}$) at wavelength λ is expressed as a linear combination of the fluorescence spectra of PSI and PSII, with the PSII fluorescence flux modulated by a regulation factor ε (Verrelst et al., 2015; Zhang et al., 2016):

$$SIF(\lambda) = SIF_{PSI}(\lambda) + \varepsilon SIF_{PSII}(\lambda) \quad (1)$$

Parameter ε is a regulatory scaling factor that represents the changes in the fluorescence emission of PSII in response to the physiological regulation of PSII related to meteorological forcing and plant stress (see equation (2)). As a first approximation, it is calculated for the TOC leaves. Also, given that the spatio-temporal dynamics of PSI fluorescence remain unknown, we assumed that the fluorescence emitted by PSI does not vary with environmental conditions or across species and that its fluorescence yield remains constant (Pedrós et al., 2010; Porcar-Castell et al., 2014). TOC PSI and PSII *SIF* fluxes are calculated with a parametric representation of the 1-D radiative transfer model in SCOPE, which accounts for the effect of the canopy structure on re-distribution (absorption/diffusion) of fluorescence fluxes through the canopy.

2.3. Leaf Scale

The competing processes of photosynthesis, fluorescence and thermal energy dissipation, occur in the chloroplast; however, we modeled them at the leaf level, which is the smallest physiological unit considered in ORCHIDEE. In a first step, we define the regulatory scaling factor ε that accounts for the changes in fluorescence yield in PSII at the steady state ($\Phi F'_{PSII}$;see equation (3)), relative to a minimal fluorescence level during summer ($\Phi F_{0\ PSII}$) obtained in the absence of physiological stress (NPQ approaching zero) and with a maximum fraction of open and functional PSII reaction centers (Baker, 2008).

$$\varepsilon = \Phi F'_{PSII} / \Phi F_{0\ PSII} \quad (2)$$

To compute the quantum yields of the processes related to PSII, we assumed that there is a large connectivity between photosynthetic units, correctly represented by the "lake" model described for instance in Lavergne and Trissl (1995), Kramer et al. (2004), Porcar-Castell (2011), or Amarnath et al. (2016). In this model, each yield equals the ratio of the rate constant of each process divided by the sum of the rate constants over all the competing processes. In the lake model formalism, $\Phi F'_{PSII}$ is expressed as a function of the photochemistry yield for PSII ($\Phi P'_{PSII}$) (Porcar-Castell et al., 2014; Lazár, 2015; Koffi et al., 2015; Lee et al., 2015):

$$\Phi F'_{PSII} = \frac{k_F}{k_D + k_F + k_{NPQ}} \left(1 - \Phi P'_{PSII}\right) \quad (3)$$

and depends on the rate constants of fluorescence (k_F) and of thermal energy dissipation. Thermal energy dissipation is divided between an invariant part (k_D, basal or constitutive thermal energy dissipation) and a variable part, k_{NPQ}, related to the total NPQ process (i.e., reversible and sustained NPQ). The fluorescence rate constant k_F does not vary (Porcar-Castell, 2011), while the rate constant related to NPQ is regulated by environmental conditions. For simplicity, all rate constants are expressed in units relative to the sum of k_D and k_F, with k_D set to 0.95 and k_F to 0.05, and the rate constant of photochemistry k_P was set to 4 as in SCOPE (van der Tol et al., 2014). This corresponds to a value of the reference fluorescence yield $\Phi F_{0\ PSII}$ (without NPQ) of 0.01, and a corresponding maximum photochemical yield $\Phi P_{PSII\ MAX}$ of 0.8. In equation (3), only k_{NPQ} and $\Phi P'_{PSII}$ are unknown. We calculate these variables using the approaches presented in sections 2.3.1 and 2.3.2.

Table 1
Values of the Parameters of the Parametric Models for α_{leaf}, k_{NPQ}, and Top of Canopy fluorescence Flux at 760 nm

Parameters	Value
Modeling of α_{leaf}(equation (5))	
a_0	0.9
b_0	0.611
c_0	12.132
d_0	−0.086
Modeling of k_{NPQ} (equation (6))	
a_1	16.042
b_1	5.74
c_1	2.167
d_1	−0.014
e_1	−0.00437
f_1	0.00576
Modeling of the TOC fluorescence flux SIF (760 nm)	
a_2	0.00561 and 0.00917, respectively, for PSI and PSII
b_2	0.0434
c_2	0.802
d_2	32.472
e_2	0.1404
f_2	0.688
g_2	−0.1197

2.3.1. Modeling of the PSII Photochemical Yield

The yield of photochemistry for PSII ($\Phi P'_{PSII}$) is a function of the most limiting factors between electron transport rate (Maxwell & Johnson, 2000; Porcar-Castell et al., 2014; Ruuska et al., 2000) and the carboxylation rate (see equation (4)). These two factors are already calculated by the photosynthesis module of ORCHIDEE, based on the Farquhar et al. (1980) and Collatz et al. (1992) models for C3 and C4 plants, respectively:

$$\Phi P'_{PSII} = MIN\left(ETR^{c}_{PSII}, ETR\right)/(APAR.c) \quad (4)$$

where ETR^{c}_{PSII}, the electron transport rate (in μmol e^{-} $.m^{-2}.s^{-1}$) limited by the carboxylation rate, derives from equation (2) (C3 plants) and equation 23 (C4 species) in Yin and Struik (2009), while *ETR* is calculated according to equation 4 of the same reference. The factor c (=0.5 mol e^{-} .(mol photon absorbed)$^{-1}$) represents the partition of energy between the two photosystems, which is assumed identical (as in Maxwell & Johnson, 2000; Hendrickson et al., 2004; Galmés et al., 2007). The absorbed photosynthetically active radiation (APAR) is expressed as the incoming photon flux density (μmol photon $.m^{-2}.s^{-1}$) multiplied by the leaf absorptance (α_{leaf}).

In the standard ORCHIDEE version used for this study, the conversion efficiency of absorbed light into ETR at limiting light relies on a foliar absorptance of 0.8 (Medlyn et al., 2002). Baker (2008) and further Porcar-Castell et al. (2014) report a value of 0.84 for α_{leaf}. Instead of using a constant value, we have introduced a parameterization of α_{leaf} as a function of the leaf chlorophyll *a* and *b* content (C_{ab}). This approach also links the effects of the photosynthetic pigments on the leaf photochemistry and on the radiative transfer within the canopy: C_{ab} is indeed one of the input parameters of SCOPE that has the largest impact in the PAR domain (400–700 nm; cf. section 2.4), because it also controls the amount of fluoresced light re-absorbed by leaves in the red domain. The parameterization relies on leaf absorptance simulations performed with the PROSPECT model of leaf optical properties (Feret et al., 2008; Jacquemoud & Baret, 1990), where the parameters of leaf structure, carotenoid content and chlorophyll content, were varied at the same time, within respective prescribed ranges of variation. The average leaf absorptance over the PAR domain is expressed as

$$\alpha_{leaf} = a_0\left(1 - b_0.\exp\left(-\frac{Cab}{c_0} + d_0\right)\right) \quad (5)$$

The parameter values were optimized against PROSPECT simulations (Table 1); with those values, leaf absorptance for a leaf chlorophyll content of 40 $\mu g.cm^{-2}$ is 0.88. The C_{ab} parameter was introduced in ORCHIDEE as PFT dependent, and a prior value of 40 $\mu g.cm^{-2}$ was set for all vegetation types (Table 3). In addition, we also account for temporal variations of C_{ab} with leaf age in the same way the control of leaf age on the maximum carboxylation capacity is already implemented in ORCHIDEE (see Krinner et al., 2005; Santaren et al., 2007), a covariation of these traits having been documented by Niinemets et al. (2004), Muraoka and Koizumi (2005), and Šesták and Čatský (1962).

2.3.2. Modeling of the Non-photochemical Quenching Effect, k_{NPQ}

Because there is no simple model in the literature that calculates k_{NPQ} (for total NPQ) as a function of the ORCHIDEE state variables related to the photosystem functioning, we followed an approach similar to the one used in SCOPE (van der Tol et al., 2014, see also Lee et al., 2015; Koffi et al., 2015) that consists in determining a parametric model of k_{NPQ} based on leaf fluorescence measurements. Compared with the above-mentioned studies however, we did not use the same measurement data sets and we explicitly accounted for the effect of incoming radiation and temperature on total NPQ. The model for k_{NPQ} was calibrated using the measurements of (i) Porcar-Castell (2011) for *Pinus sylvestris* that are adapted to boreal climate and (ii) Flexas et al. (2002) for several broadleaf (*Celtis australis* L., *Pistacia terebinthus* L., *Quercus ilex* L.) and crop (*Solanum melongena* L.) species that are more adapted to a Mediterranean climate. Since we did

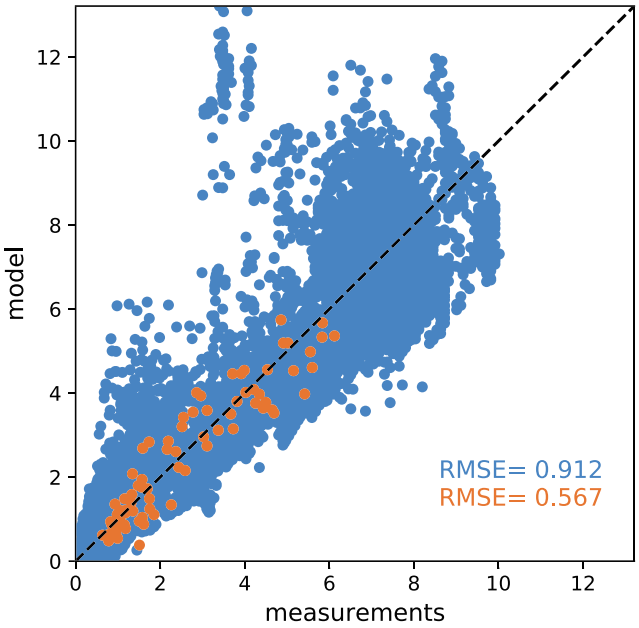

Figure 1. Modeled versus measured k_{NPQ} values (in s^{-1}) for the data sets of Porcar-Castell (2011)—in blue and Flexas et al. (2002)—in orange, with the corresponding RMSEs.

not have any reasonable proxies to estimate the contribution of PSI to the PAM fluorescence data, we decided not to attempt any correction for PSI fluorescence in the calibration PAM data set and avoid any additional source of uncertainty:

$$k_{NPQ} = a.x^c.\frac{1+b}{b+x^c}.\frac{\exp^{(d.T+e)}}{PAR^f} \quad (6)$$

The values of the parameters are provided in Table 1. T is the air temperature, bounded to −2.5 °C. As in van der Tol et al. (2014), the parameter $\boldsymbol{x}$ is defined as

$$x = 1-\Phi P'_{PSII}/\Phi P_{PSII\ MAX} \quad (7)$$

Despite the differences in terms of plant species and climate conditions between the two data sets used for tuning the k_{NPQ} model (equation (6)), the model is able to reproduce both observation data sets reasonably well, as illustrated in Figure 1. The better performance seen for the Flexas data set is explained by the smaller number of measurements (78), also performed under more controlled conditions (acquisitions performed at midday in summer), than the 24,500 *Pinus sylvestris* measurements that correspond to about one year of continuous monitoring. We evaluated the models of van der Tol et al. (2014) and Lee (Lee et al., 2015) on the tuning data set, and they showed lesser performance: the root-mean-square error (RMSE) were, respectively, of 2.05 and 2.25 s^{-1} over the Porcar-Castell (2011) data, and 0.75 and 1.45 s^{-1} over the Flexas et al. (2002) data.

2.4. Canopy Level

For the calculation of the TOC fluorescence fluxes of PSI and PSII (equation (1)), we set up a parametric model that we tuned from an ensemble of simulations of the SCOPE (v1.61) model (equation (8)). The ensemble was designed by simultaneously varying the most influential model parameters on canopy fluorescence found in the sensitivity study conducted by Verrelst et al. (2015). In the selection of key parameters, we also considered the possibility to relate them to ORCHIDEE parameters or state variables. In the PAR domain, the main contributors to the total canopy fluorescence are the leaf area index (LAI; which is prognostically calculated by ORCHIDEE), the chlorophyll a + b content, the leaf angle distribution, the dry matter content, and the incoming radiation. The leaf angle distribution implemented in SCOPE depends on two parameters. In order to reduce the number of degrees of freedom of the parametric model, we replaced it by the ellipsoidal distribution (Campbell, 1990) that depends on the average leaf angle (*ALA*) parameter only. Dry matter content is inversely proportional to specific leaf area (*SLA*) which is an input parameter of ORCHIDEE. Given we aim at assimilating OCO-2 observations acquired at nadir, we fixed the view zenith angle at 0° while variations of the sun zenith angle are accounted for.

We conducted a preliminary sensitivity analysis of the SCOPE model to identify its main drivers (within the confines of parameters and variables used or simulated by ORCHIDEE) and mimic their mean effects through mathematical functions. The simulations were determined thanks to a statistical approach used in Design Of Experiments for Simulation (DOES) (Bacour et al., 2002). A Hyper Graeco Latin Geometric sampling scheme made of 2,401 simulations was used (as in Jacquemoud et al., 2009). Table 2 lists SCOPE input parameters and indicates those varying with each simulation, including their variation interval, and those kept fixed, with their associated value. Along the simulations, the varying parameters took seven discrete values evenly spread over their variation range.

The determination of the mean effect of each parameter from the DOES approach helped to identify the mathematical functions that best reproduce them, and therefore to further build the parametric model. Several formulations were tested and the one that best fits the original SCOPE fluorescence is expressed as

Table 2
SCOPE Input Parameters and Values/Range of Variation Used in the DOES Approach (See Text)

Parameter	Description	Fixed value or [discrete values]
Leaf		
N	Leaf structure parameter	1.4
Cab	Chlorophyll content ($\mu g.cm^{-2}$)	[20, 30, 40, 50, 60, 70, 80]
Cca	Carotenoid content ($\mu g.cm^{-2}$)	20
Cdm	Dry matter content ($g.cm^{-2}$)	[0.001, 0.0092, 0.0173, 0.0255, 0.0337, 0.0418, 0.05]
Cs	Senescence factor	0
Cw	Water content ($g.cm^{-2}$)	0.009
Canopy		
LAI	Leaf area index ($m^2.m^{-2}$)	[0.1, 1.3, 2.6, 3.8, 5, 6.3, 7.5]
ALA	Average leaf angle (°)	[28, 33.7, 39.3, 45, 50.7, 56.3, 62]
lw	Leaf width (m)	0.1
hc	Canopy height (m)	1
Geometry/forcing		
SZA	Sun zenith angle (°)	[5, 15, 25, 35, 45, 55, 65]
VZA	View zenith angle (°)	0
RAA	Relative azimuth angle (°)	180
Rin	Incoming shortwave radiation ($W.m^{-2}$)	[1, 234, 467, 700, 934, 1167, 1400]

Note. Parameters in bold are those that vary and the mean effect of which is further assessed.

$$SIF(\lambda) = \frac{a_2.Rin}{1+\exp(-b_2.C_{ab})}.\frac{1}{1+LAI^{c_2}}.\frac{1}{1+d_2.Cdm}\left[1-\exp\left(-e_2.\frac{\cos(ALA)^{f_2}}{\cos(SZA)^{g_2}}.LAI\right)\right] \quad (8)$$

for a given wavelength λ and with parameters a_2 to g_2 estimated by a nonlinear least squares optimization procedure, using SCOPE radiative transfer simulations as pseudo-observations (after deactivating the biochemical module). For an application to OCO-2 space-borne observations, the model was adjusted to a wavelength λ of 760 nm and a constant view geometry set at nadir was considered. Note that at this wavelength, the contribution of *Cab* to the total canopy fluorescence is small (Verrelst et al., 2015). The values of the parameters are listed in Table 1. The ability of the parametric model to reproduce the original SCOPE simulations was assessed over an ensemble of 300 simulations where the dependent parameters were changed randomly within their variation range. The high values of the correlation coefficients (0.996 for both PSI and PSII) together with low RMSE values (0.015 $W.m^{-2}.\mu m^{-1}.sr^{-1}$ for PSI and 0.025 $W.m^{-2}.\mu m^{-1}.sr^{-1}$ for PSII) highlights the ability of the parametric model to accurately emulate the radiative transfer module of SCOPE. The parameters a_2 to g_2 will not be further optimized against OCO-2 SIF observations, which is equivalent to assume that the SCOPE model (and its emulated version) is perfect.

We recall that, while the parametric model somewhat represents the fluorescence regime (emission and re-absorption) within the canopy, we consider that the regulation of PSII fluorescence is essentially determined by the top (illuminated) layer of the canopy. This important hypothesis is discussed in section 5.

3. Data and Data Assimilation Methods

3.1. ORCHIDEE Configuration

ORCHIDEE is forced by 6-hourly meteorological fields from CRU-NCEP v7 (a product combining the climate reanalysis from the National Centers for Environmental Prediction and observations from the Climatic Research Unit) at 0.5° resolution (Viovy, 2018), which are interpolated to a half-hour time step within ORCHIDEE (linearly for all fields except for the shortwave downwelling radiation which is interpolated using the cosine of the solar zenith angle). A prior spin-up simulation was performed at the global scale by recycling forcing data over the 1901–1910 period in order to bring plant carbon reservoirs to equilibrium. This was followed by a transient simulation up to the observation period accounting for changing climatic conditions and the yearly increase of the global atmospheric CO_2 concentrations. Grid-point-scale simulations (section 3.4.1) rely on the initial conditions extracted from this global run.

In order to explicitly account for the impact of the time of OCO-2 overpass in the estimation of SIF, the simulated TOC SIF corresponds to the value at 1:30 pm local time. ORCHIDEE TOC SIF simulations were performed every day (hence accounting for the temporal evolution of the illumination angle) and were then averaged at a monthly resolution (cf. section 3.2.1).

The overarching hypothesis of our approach is that instantaneous SIF estimates (here derived from OCO-2 observations at ~1:30 local time) are able to constrain daily GPP (here simulated by the ORCHIDEE model). It is supported by the recent work of Zhang, Xiao, et al., 2018).

3.2. Data Sets

3.2.1. Assimilated Data: OCO-2 SIF Products

The OCO-2 instrument was launched in July 2014 to monitor atmospheric CO_2. Thanks to its high spectral resolution in the O_2 A-band, the land surface SIF is estimated and distributed as a separate product (Frankenberg et al., 2014). We used the OCO-2 B7300 SIF data available at ftp://fluo.gps.caltech.edu/data/

OCO2/sif_lite_B7300/ which has shown very good consistency with airborne SIF measurements (Sun et al., 2017). In this study, we used the monthly synthesis products at a 0.5° × 0.5° spatial resolution for years 2015–2016. The high spatial resolution of OCO-2 observations (footprint of 1.3 × 2.25 km^2) combined with their low revisit frequency (16 days) make such spatio-temporal aggregates best suited for comparison with a TBM at the global scale. The aggregated products also have reduced random errors. We only considered nadir observations to remove directional effects not accounted for in this study. OCO-2 has a sun-synchronous orbit with an ascending node at 1:30 pm local time; as we explicitly accounted for the time of the satellite overpass in ORCHIDEE simulations, we assimilated the raw OCO-2 SIF product and not the daily normalized product.

3.2.2. GPP Evaluation

3.2.2.1. Procedure

We compared ORCHIDEE GPP simulations with respect to two independent GPP estimates derived from data-driven approaches. We used different years for the evaluation of the simulated GPP than for the assimilation of SIF data (2015–2016); this can be seen as an opportunity to assess the temporal validity of the optimized model. To do so, we used the parameter values optimized over the assimilation period to perform global-scale hindcast simulations of GPP.

3.2.2.2. Data Sets

In the first place, we choose the recent FLUXCOM GPP products described in Tramontana et al. (2016), based on an upscaling of *in situ* estimates of GPP at FLUXNET sites (http://fluxnet.ornl.gov) with several global MODIS (Moderate-Resolution Imaging Spectroradiometer) satellite products and meteorological forcing. In this study, we used the median between the three machine learning estimates as a reference GPP and based on the flux partitioning approach of Reichstein et al. (2005). The spatial and temporal resolutions of the FLUXCOM GPP are 0.5° and monthly, the same as the ORCHIDEE simulations. FLUXCOM should not be considered as a validation product (as in Norton et al., 2018) given the sparsity of the training flux sites and other shortcoming issues (for instance, it does not represent the fertilization effect of increasing CO_2 atmospheric concentrations), but it is a commonly used product usually taken as a benchmark. In Figure 3 of Tramontana et al. (2016), a scatterplot of their eight-day product against independent observations shows a RMSE of 1.5 $gC.m^{-2}.d^{-1}$. We also use the FLUXSAT GPP estimates described in Joiner et al. (2018). Similarly to FLUXCOM, FLUXSAT also derives from a data-driven approach relying on FLUXNET measurements and MODIS reflectances in seven spectral bands and calibrated against FLUXNET measurements. One noteworthy difference is that FLUXSAT does not use any meteorological forcing. A dual-calibration procedure is applied by discriminating low versus high productive FLUXNET sites, the identification of which is based on space-borne SIF products derived from Global Ozone Monitoring Experiment 2 (GOME-2) observations. FLUXSAT GPP calibration remains independent from SIF data and the GPP estimates are not impacted by GOME-2 SIF artifacts (instrument degradation; Zhang, Joiner, et al., 2018).

The FLUXCOM product is used as the main benchmark for the analyses presented in the following, given that it has been used in numerous evaluation studies regarding land surface model simulations (Norton et al., 2018; Sun et al., 2019, among others) and space-borne SIF estimates (Walther et al., 2016; Sun et al., 2018; Zhang, Xiao, et al., 2018). We used FLUXCOM data available over the 1990–2013 period, with specific time windows depending on the evaluation purpose: focus on year 2008 for assessing the modeled GPP spatial patterns at the global scale relative to SIF data, 2008–2013 for the temporal evolution of the modeled GPP for an ensemble of OCO-2 pixels, and 1990–2009 for characterizing the trend of the yearly global GPP budget. FLUXSAT is available from 2000 onward. We used the data over 2001–2009 to assess the yearly variations of the GPP budget at the global scale.

3.3. Overview of the Data Assimilation Framework

The optimization of ORCHIDEE with regard to OCO-2 SIF products was performed with the ORCHIDEE Data Assimilation System (ORCHIDAS) (https://orchidas.lsce.ipsl.fr/). ORCHIDAS has been described in detail in previous studies focusing on the assimilation of *in situ* net surface CO_2 and energy flux observations, satellite-derived vegetation indices or FAPAR products and atmospheric CO_2 concentration data (Bacour et al., 2015; Kuppel et al., 2012; MacBean et al., 2015; Peylin et al., 2016), and on the assimilation of GOME-2 SIF products using a simple linear SIF-GPP relationship (MacBean et al., 2018).

ORCHIDAS relies on a Bayesian framework to optimize a set of selected ORCHIDEE parameters (called *control variables* and gathered in a *control vector* $\mathbf{x}$) from a background (*a priori* knowledge) estimate ($\mathbf{x_b}$), using OCO-2 SIF observations ($\mathbf{y}$), while explicitly accounting for the uncertainties in parameters and observations (assuming Gaussian probability distributions as in most data assimilation studies). The optimal set of parameters is found by minimizing a function $J(\mathbf{x})$ that gauges the misfit between model outputs ($H(\mathbf{x})$) and observations, together with the departure of $\mathbf{x}$ from $\mathbf{x_b}$, relative to the error covariance matrices of the observations ($\mathbf{R}$, including model and observation errors) and of the parameters ($\mathbf{B}$; Tarantola, 2005):

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{y}-H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x})) + (\mathbf{x}-\mathbf{x_b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b}) \quad (9)$$

How the $\mathbf{B}$ and $\mathbf{R}$ matrices are determined is described, respectively, in sections 3.4.1 and 3.4.2.

After optimization, several metrics can be calculated to assess the knowledge improvement brought by the assimilation. The posterior error covariance matrix of the optimized parameters ($\mathbf{A}$) is computed under the hypothesis that the model is linear in the vicinity of the solution. It is a function of $\mathbf{H}_\infty$, the Jacobian matrix of the model at the solution:

$$\boldsymbol{A} = \left[\mathbf{H}_\infty{}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_\infty + \mathbf{B}^{-1}\right]^{-1} \quad (10)$$

The uncertainty reduction on model parameters is quantified by $1-\sigma_{post}/\sigma_{prior}$, where σ_{post} and σ_{prior} are the error standard deviations derived from the posterior ($\mathbf{A}$) and prior ($\mathbf{B}$) covariance matrices, respectively.

Finally, to quantify the model improvement, we calculated the reduction of the root-mean-square difference (RMSD) between model and data, prior and posterior to optimization, expressed in %, as $100 \times (1 - \mathrm{RMSD}_{post}/\mathrm{RMSD}_{prior})$.

3.4. Data Assimilation Setup

3.4.1. Optimized Parameters and Associated Uncertainties

We optimized several ORCHIDEE parameters involved in carbon assimilation and hence directly impacting GPP dynamics through photosynthesis and phenology processes. The parameter values vary with PFT. Most of these parameters have already been identified in previous studies (Kuppel et al., 2014; MacBean et al., 2018; Peylin et al., 2016) that focused on an older version of the model. The process equations they are involved in are described in these papers as well as in Krinner et al. (2005), Santaren et al. (2007), and MacBean et al. (2016). New ones were introduced in the control vector $\mathbf{x}$ for this version of ORCHIDEE. The list of optimized PFT-dependent parameters is provided in Table 3, with their prior value and defined interval of variation (defined from a literature survey and from expert knowledge). The "photosynthesis" parameters in Table 3 have a direct impact on the ε regulatory scaling factor of PSII fluorescence emission (equation (1)). Note that ETR is controlled by maximum rate of e^- transport rate under saturated light (J_{max25}) which is inferred from the maximum rate of Rubisco activity-limited carboxylation at 25 °C (V_{cmax25} parameter). For C3 plants, the ratio J_{max25}/V_{cmax25} is a linear function of the plant growth temperature, following Kattge and Knorr (2007); the offset is 2.59 and the slope is −0.035 (see their Table 3). For C4 plants, the model does not consider temperature acclimation, with a fixed ratio of 1.715. Two parameters specific to the SIF model were also adjusted, C_{ab} and ALA. These two parameters are currently only impacting the computation of the SIF fluxes and not the photosynthesis calculation. Like many ORCHIDEE parameters, C_{ab} and ALA are PFT dependent. However, without any prior information on their distribution with vegetation type, the same *a priori* values and variation ranges were used across PFTs. The a priori value for ALA corresponds to a spherical distribution of leaves. Only the diagonal elements of the uncertainty matrix of the parameters $\mathbf{B}$ were accounted for, with the standard deviation of the error associated to each parameter being set to 40% of its variation range (Kuppel et al., 2012).

3.4.2. Selection of OCO-2 Grid Points and Observation Uncertainties

The spatial resolution of the OCO-2 product used here (0.5°) somewhat precludes the assimilation of global observation maps because of excessive computational times. Instead, we assimilated OCO-2 time series for an ensemble of selected grid points, an approach already employed in MacBean et al. (2015, 2018). For each PFT, 15 grid points were randomly chosen among those with the highest fractions of the targeted PFT for the optimization of the model, and 10 other ones for evaluation purposes. Grid points with a

Journal of Geophysical Research: Biogeosciences 10.1029/2019JG005040

Table 3
Definition of the Optimized Parameters, With Their Prior Value and Range of Variation (in Brackets) for the Different PFTs

Name	Description	TrEBF	TrDBF	TeENF	TeEBF	TeDBF	BoENF	BoDBF
Photosynthesis								
A1	Empirical factor involved in the calculation of effect of leaf-to-air vapor difference to the coupling between CO_2 assimilation and stomatal conductance	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]
B1	Empirical factor involved in the calculation of effect of leaf-to-air vapor difference: Ball-Bery offset	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]
V_{cmax25}	Maximum rate of Rubisco activity-limited carboxylation at 25 °C ($\mu mol\ [CO_2] \cdot m^{-2} \cdot s^{-1}$)	45 [27, 63]	45 [27, 63]	35 [21, 49]	40 [24, 56]	50 [20, 80]	45 [27, 63]	35 [20, 60]
Hum_{cste}	Root profile (m)	0.8 [0.6, 1]	0.8 [0.6, 1]	1 [0.75, 1.25]	0.8 [0.6, 1]	0.8 [0.6, 1]	1 [0.75, 1.25]	1 [0.75, 1.25]
SLA	Specific leaf area ($m^2.g^{-1}$)	0.0153 [0.011, 0.019]	0.026 [0.019, 0.033]	0.0093 [0.007, 0.012]	0.02 [0.015, 0.025]	0.026 [0.018, 0.04]	0.0093 [0.007, 0.012]	0.026 [0.019, 0.033]
Phenology								
LAI_{max}	Maximum LAI value ($m^2.m^{-2}$)	7 [4.2, 9.8]	5 [3, 7]	5 [3, 7]	4 [2.4, 5.6]	5 [2.5, 7.5]	3.5 [2.1, 4.9]	4 [2,.6]
T_{sen}	Temperature threshold for senescence (°C)	–	–	–	–	16 [8, 24]	–	14 [7, 21]
$L_{age\text{-}crit}$	Average critical leaves age (days)	730 [547, 913]	180 [135, 225]	910 [682, 1138]	730 [547, 913]	160 [120, 200]	910 [682, 1138]	220 [165, 275]
L_{fall}	Length of leaves shedding (days)	–	10 [6, 14]	–	–	30 [18, 42]	–	5 [3, 7]
$K_{LAI\text{-}happy}$	LAI threshold to stop using carbohydrate reserves during growth (–)	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]
$T_{leaf\text{-}init}$	Time to attain the initial foliage using the carbohydrate reserve (days)	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]
GDD_{crit}	Critical growing-degree days threshold (°C)	–	–	–	–	–	–	–
NGD_{crit}	Critical number of growing days threshold (°C)	–	–	–	–	–	–	–
Hum_{sen}	Critical relative moisture availability for senescence (unitless)	–	0.3 [0.15, 0.45]	–	–	–	–	–
Hum_{nosen}	Relative moisture availability above which there is no humidity-related senescence (unitless)	–	0.8 [0.6, 1]	–	–	–	–	–
$L_{ageMin\text{-}Sen}$	Minimum leaf age to allow senescence (days)	–	90 [67.5, 112.5]	–	–	90 [67.5, 112.5]	–	60 [45, 75]
$TGDD_{crit}$	Temperature threshold (days)	–	–	–	–	–	–	–
SIF model								
C_{ab}	Chlorophyll $a + b$ content ($\mu g.cm^{-2}$)	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]
ALA	Average leaf angle (°)	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]

Note. The PFTs are tropical evergreen broadleaf (TrEBF), tropical broadleaf deciduous (TrDBF), temperate evergreen needleleaf (TeENF), temperate evergreen broadleaf (TeEBF), temperate deciduous broadleaf (TeDBF), boreal evergreen needleleaf (BoENF), boreal deciduous broadleaf (BoBDF), boreal needleleaf deciduous (BoNDF), forests, and C3 grasslands (temperate—TeC3GRA, tropical—TrC3GRA, and boreal—BoC3GRA) and C4 grasslands (C4GRA), as well as C3 and C4 crops (C3CRO and C4CRO).

Journal of Geophysical Research: Biogeosciences

10.1029/2019JG005040

Table 3
(continued)

Name	BoNDF	TeC3GRA	TrC3GRA	BoC3GRA	C4GRA	C3CRO	C4CRO
Photosynthesis							
A1	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.85 [0.64, 1.06]	0.72 [0.54, 0.9]	0.85 [0.64, 1.06]	0.72 [0.54, 0.9]
B1	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.14 [0.1, 0.18]	0.2 [0.15, 0.25]	0.14 [0.1, 0.18]	0.2 [0.15, 0.25]
V_{cmax25}	35 [20, 60]	50 [30, 70]	50 [30, 70]	40 [24, 56]	50 [30, 70]	60 [36, 84]	60 [36, 84]
Hum_{cste}	0.8 [0.6, 1]	4 [3, 5]	4 [3, 5]	4 [3, 5]	4 [3, 5]	4 [3, 5]	4 [3, 5]
SLA	0.019 [0.012, 0.03]	0.026 [0.019, 0.033]	0.026 [0.019, 0.033]	0.026 [0.019, 0.033]	0.026 [0.019, 0.033]	0.026 [0.019, 0.033]	0.026 [0.019, 0.033]
Phenology							
LAI_{max}	3 [1.5, 5.5]	2.5 [1.5, 3.5]	2.5 [1.5, 3.5]	2 [1.2, 2.8]	2 [1.2, 2.8]	5 [3, 7]	5 [3, 7]
T_{sen}	10 [5, 15]	5 [2.5, 7.5]	5 [2.5, 7.5]	5 [2.5, 7.5]	5 [2.5, 7.5]	5 [2.5, 7.5]	10 [5, 15]
$L_{age\text{-}crit}$	120 [90, 150]	80 [60, 100]	80 [60, 100]	80 [60, 100]	120 [90, 150]	90 [67.5, 112.5]	90 [67.5, 112.5]
L_{fall}	10 [6, 14]	10 [6, 14]	10 [6, 14]	10 [6, 14]	10 [6, 14]	10 [6, 14]	10 [6, 14]
$K_{LAI\text{-}happy}$	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]	0.5 [0.37, 0.63]
$T_{leaf\text{-}init}$	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]	10 [7.5, 12.5]
GDD_{crit}	–	320 [240, 400]	320 [240, 400]	320 [240, 400]	400 [300, 500]	320 [240, 400]	700 [525, 875]
NGD_{crit}	17 [10.2, 23.8]	–	–	–	–	–	–
Hum_{sen}	–	0.2 [0.1, 0.3]	0.2 [0.1, 0.3]	0.2 [0.1, 0.3]	0.2 [0.1, 0.3]	0.3 [0.15, 0.45]	0.2 [0.1, 0.3]
Hum_{nosen}	–	0.6 [0.45, 0.75]	0.6 [0.45, 0.75]	0.6 [0.45, 0.75]	0.3 [0.22, 0.38]	0.3 [0.22, 0.38]	0.3 [0.22, 0.38]
$L_{ageMin\text{-}Sen}$	60 [45, 75]	30 [22.5, 37.5]	30 [22.5, 37.5]	30 [22.5, 37.5]	30 [22.5, 37.5]	30 [22.5, 37.5]	30 [22.5, 37.5]
$TGDD_{crit}$	–	–	–	–	20 [15, 25]	–	–
SIF model							
C_{ab}	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]	40 [30, 50]
ALA	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]	57.4 [43, 72]

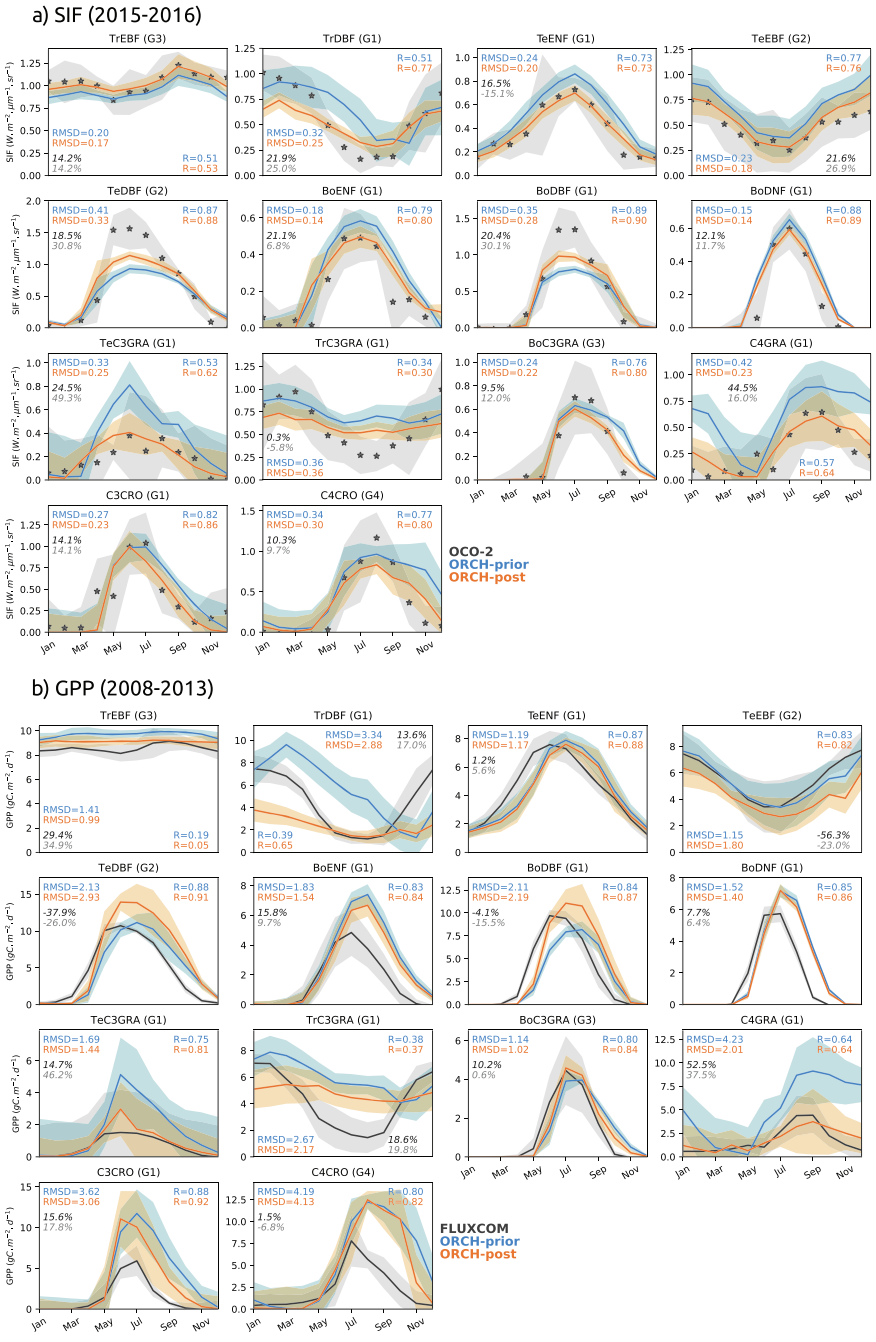

Figure 2. Comparison of the monthly time series of (a) SIF (averaged over 2015 and 2016) and (b) GPP (mean seasonal cycles over 2008–2013), estimates (in black) with ORCHIDEE simulations (prior in blue and posterior in orange) for the 14 vegetated PFTs over the grid points used for optimization. Plain lines correspond to the median SIF values over the 15 optimization grid points selected for each PFT, and the shaded areas is the corresponding standard deviation. For each grid point, the posterior simulations are performed with the estimated values for all PFTs (and not only the dominant one). The RMSD of fit and correlation coefficient over all grid points before and after assimilation is provided for each vegetation type. The reduction of the RMSD is provided (%) for the 15 optimization grid points (black italic) and for 10 validation grid points (grey italic). The group (letter G next to PFT name) refers to model changes relative to SIF and GPP, and is described in section 4.2.1.

fraction of the target PFT lower than the 85th percentile were discarded in order to select highly homogeneous grid points. The higher biome homogeneities were found for TrEBF, TeEBF, TeC3GRA, BoC3GRA, and C4GRA, vegetation types with 85th percentile values of the vegetation fraction above

70% for the selected grid points. For the other PFTs, the 85th percentile threshold corresponds to vegetation fractions usually higher than 50%, except for C4CRO (between 40% and 56%) and TrC3GRA (46–73%). In addition, grid cells exhibiting a low temporal correlation between the ORCHIDEE prior SIF simulation and OCO-2 SIF observations were discarded from the selection, given that they likely reflect model-data inconsistencies that the assimilation would not reconcile. To do this, we removed grid points with a correlation below a prescribed threshold (55th percentile of the correlation distribution estimated PFT per PFT). Finally, only grid cells with more than six valid OCO-2 monthly means per year were retained.

As in previous ORCHIDAS studies (MacBean et al., 2015; Santaren et al., 2014), the observation uncertainty (**R** matrix) was set diagonal and the prior error was defined for each PFT as the RMSD between the prior model simulations and OCO-2 SIF data over all 15 grid points selected for assimilation. The values of the prior RMSD for each PFT, averaged over the 15 grid points, are provided in Figure 2a. The observation uncertainty accounts for uncertainties in the SIF retrieval processing chain (from the radiance measurements to the monthly gridded products) and uncertainties originating from the ORCHIDEE-SIF model (including parameterization uncertainty as well as uncertainties in the meteorological forcing and in the prescribed land cover).

3.4.3. Data Assimilation Procedure

The assimilation was performed in two steps in order to reduce the computational burden. At the first step, it was performed PFT per PFT (over the 15 selected grid points) for the 14 vegetated PFTs, and only the ORCHIDEE parameters of the main PFT at the 15 grid points were adjusted by the optimization (similar to MacBean et al., 2018). In a second step, all pixels/PFTs were combined, in order to better account for the vegetation mixing within grid cells, the optimization process starting from the parameter values pre-optimized at the first step. The iterative minimization of $J(\mathbf{x})$ (equation (9)) was achieved by the gradient-based L-BFGS-B algorithm (Byrd et al., 1995) which accounts for parameter bounds. It requires the estimation of the gradient of the misfit function at each iteration, here calculated by a finite difference approach. The minimum of $J(\mathbf{x})$ detected by the assimilation algorithm may be sensitive to the starting point of $\mathbf{x}$ due to possible multiple local minima (arising from model nonlinearity) (Bastrikov et al., 2018). For this reason, we carried out 20 optimizations at step 1 starting from 20 different parameter first guesses, chosen randomly within their range of variation. We selected the solution of step 1 as the one out of the 20 that resulted in the lowest posterior value of $J(\mathbf{x})$; for parameters that do not depend on PFT, we retained the median value over all PFTs as the estimated value. The values estimated at step 1 were then used as first guesses for the second step of the assimilation procedure. Note that the **R** and **B** matrices remained unchanged from the first to the second step.

4. Results

4.1. Improvement of the Modeled SIF After Optimization

Figure 2a illustrates the mean improvement of the model with respect to SIF observations for the grid points used for optimization and validation. The prior model already shows relatively good agreements with the mean OCO-2 SIF time series, but disagreements in magnitude and seasonality are nonetheless visible, the importance of which varies with the PFT. An overestimation of the prior SIF magnitude is seen for most PFTs, except for TeDBF, BoDBF, TrEBF, BoC3GRA, C3CRO, and C4CRO, where it is underestimated. For deciduous PFTs, the prior model also generally simulates a longer growing season. The assimilation generally corrects these general biased features in the model. For the optimization grid points, the median value of the reduction in RMSD between observed and simulated SIF is 18% across all PFTs, ranging from only 0.3% for TrC3GRA to about 45% for C4GRA.

The RMSD reduction for the 10 evaluation grid points is usually lower, with a median value of 14% over all PFTs. Noticeably, it is however higher than the one computed over the optimization grid points for six PFTs (TrDBF, TeEBF, TeDBF, BoDNF, TeC3GRA, BoC3GRA) while higher posterior RMSD are obtained for TeENF and TrC3GRA which indicates a degradation of the model performance.

For the temperate and boreal broadleaf deciduous forests (TeDBF, BoDBF), and, to a lesser extent, boreal C3 grass (BoC3GRA) and crops (C3CRO and C4CRO), the optimized (posterior) model still underestimates the

peak of SIF during the growing season; for C3CRO the only improvement is related to phenology. For the several deciduous PFTs, the temporal improvement of SIF is achieved by a shortening of the growing season length mostly via an earlier senescence.

4.2. Impact on GPP Spatio-temporal Patterns

4.2.1. GPP Temporal Dynamics at the Optimization Grid Points

The impact on the GPP simulation resulting from the assimilation of OCO-2 SIF retrievals is presented in Figure 2b for the mean cycle over the 2008–2013 period. For 11 PFTs out of 14, we note an increase of the consistency between the model and FLUXCOM-GPP after optimization with a median improvement of about 6% over all PFTs. The highest improvements are seen for C4GRA (by 53%), TrEBF (29%), TrC3GRA (18%), BoENF (16%), C3CRO (16%), and TeC3GRA (15%). At the opposite, an increase in RMSD is obtained for four PFTs: for temperate (by 38% for TeDBF, up to 56% for TeEBF) and boreal (by 4% for BoDBF) forests, as well as C4CRO but only for the validation pixels (about 7%).

For all PFTs but three, the variation in SIF magnitude through the optimization (increase or decrease) is consistent with that of GPP. For TrEBF, BoC3GRA, and C4CRO, however, opposite variations in modeled SIF and GPP are obtained after assimilation. These opposite variations are partly attributed to two parameters, C_{ab} and ALA, that depend only on the SIF model. The relative differences between ORCHIDEE SIF, OCO-2 SIF, and FLUXCOM-GPP magnitude suggest some biome dependence in terms of PSII fluorescence yield and/or of radiative transfer regimes.

A first group of PFTs (G1 in Figure 2) encompasses those showing similar relative changes (posterior–prior) in SIF and GPP magnitudes simulated by ORCHIDEE and for which the assimilation results in a joint improvement with respect to both SIF and GPP benchmark. It encompasses nine PFTs: TrDBF, TeENF, boreal forests (BoENF, BoDBF, BoDNF), TeC3GRA and TrC3GRA, C4GRA, and C3CRO. For TrDBF and TrC3GRA however, the posterior amplitude of the SIF and GPP temporal profiles remains lower than that recorded by OCO2-SIF or estimated by FLUXCOM-GPP.

The second group of PFTs (G2: TeEBF and TeDBF) corresponds to those exhibiting a coherent change in simulated magnitude for SIF and GPP after assimilation, but for which the optimized GPP further departs from FLUXCOM. For these PFTs, the prior model was already in good agreement with FLUXCOM estimates. The change in modeled SIF magnitude required to fit OCO-2 data results in a strong underestimation/overestimation with respect to FLUXCOM-GPP for TeEBF/TeDBF. Note that for TeDBF, the increase of SIF amplitude during the growing season in the optimized model still underestimates OCO-2 SIF amplitude. Importantly however, the posterior GPP overestimates FLUXCOM and is degraded compared to the prior model-data fit. This apparent inconsistency between the model, OCO-2 SIF and FLUXCOM, still need to be resolved, particularly given that these PFTs are well represented by eddy covariance tower sites in the FLUXCOM product.

The third and fourth groups (G3: TrEBF and BoC3GRA; G4: C4CRO) correspond to PFTs for which the model optimization results in opposite variation of SIF and GPP. While G3 shows an improved agreement between the optimized ORCHIDEE model and FLUXCOM, G4 shows an increased mismatch but only for the validation pixels.

With regard to the seasonality, we mostly observe an enhanced agreement between the optimized model and both OCO-2 SIF observations and FLUXCOM-GPP estimates (higher correlation coefficients for 11 PFTs out of 14). For several PFTs (TeENF, TeDBF, BoENF, BoDBF, BoDNF), the optimized model however simulates a shifted carbon uptake period as compared to FLUXCOM (later start and later senescence), although it captures reasonably well what is actually monitored by OCO-2 SIF.

The fact that we improve the consistency between ORCHIDEE-GPP with FLUXCOM for tropical tree and most herbaceous PFTS and degrade it for some temperate and boreal tree PFTs, while always improving the modeled SIF with respect to OCO-2 data, may indicate that we estimated sub-optimal parameter values. This may also point to a deficiency in our model to represent a biome dependency with respect to the PSII fluorescence yield at the leaf scale or to the radiative transfer.

4.2.2. GPP Spatial Dynamic at Global Scale

The maps in Figure 3 show the changes in ORCHIDEE annual mean SIF and GPP for all continental grid points after assimilation (over 2015–2016 for SIF, and 2008 for GPP), compared to OCO-2 SIF observations

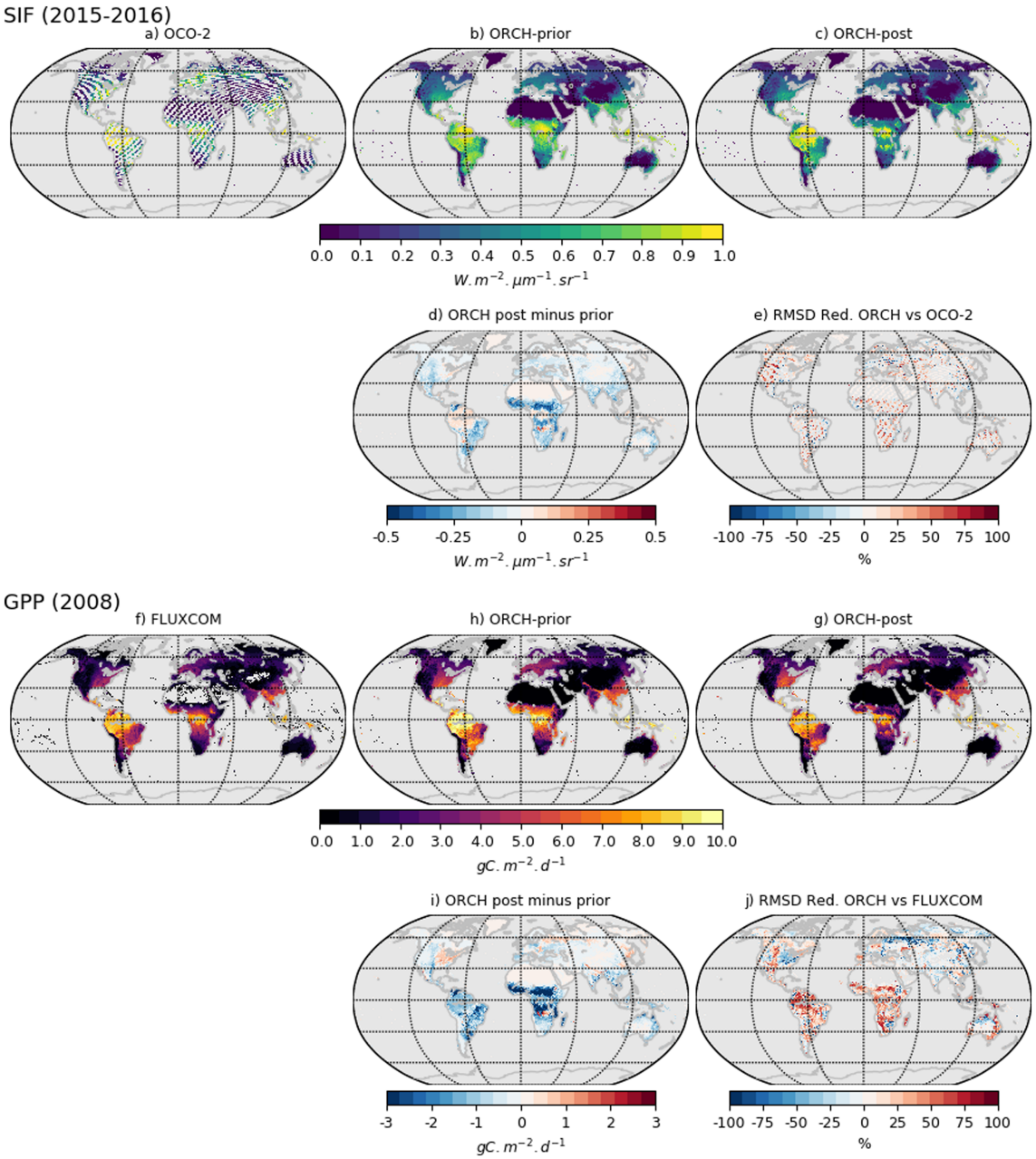

Figure 3. Global maps of (top) SIF (2015–2016) and (bottom) GPP (year 2008) annual means at 0.5°, with ORCHIDEE simulated variables before and after assimilation being compared to (a) OCO-2 SIF product and (f) FLUXCOM-GPP: ORCHIDEE prior simulations of SIF and GPP are shown in (b) and (h), respectively; the simulations obtained with the optimized model are seen in (c) for SIF and (g) for GPP. The differences between ORCHIDEE simulations obtained with the optimized and the prior model are seen in (d) for SIF and (i) for GPP. The reduction in RMSD between ORCHIDEE outputs and OCO-2 following the data assimilation is shown in (e), and in (j) for FLUXCOM-GPP.

and FLUXCOM-GPP estimates. The *a priori* ORCHIDEE model already reproduces OCO-2 SIF observations reasonably well in terms of spatial patterns and annual magnitude, with however a tendency to underestimate SIF annual-mean values over the tropical rain forests (in particular in the Amazon and Congo basins) and to overestimate SIF in the surrounding regions (Figures 3a and 3b). ORCHIDEE SIF also underestimates SIF product over the temperate Northern Hemisphere (eastern part of the United States and western Europe). Of particular note, OCO-2 exhibits about as high SIF values over the temperate Northern Hemisphere as over the tropics (values above 1.2 $W.m^{-2}.\mu m^{-1}.sr^{-1}$), while

FLUXCOM shows a ratio of GPP annual magnitude of about two between the tropics and temperate Northern Hemisphere. Otherwise, the spatial gradients of FLUXCOM-GPP (Figure 3f) and OCO-2-SIF appear consistent (despite the different periods considered; see also Sun et al., 2018), with a spatial correlation coefficient between the annual means of 0.8. The ORCHIDEE simulations show a higher spatial correlation between prior SIF and prior GPP magnitudes (0.95). The spatial correlation between prior ORCHIDEE-SIF and OCO-2 (0.78) is also slightly lower than the spatial correlation between prior ORCHIDEE-GPP and FLUXCOM-GPP (0.84).

The assimilation of OCO-2 fluorescence estimates decreased the modeled yearly SIF for most regions (blue pixels in Figure 3d), which generally goes along with an increased consistency with respect to OCO-2 SIF. The model improvement occurs at most grid points (red pixels in Figure 3e), but a degradation can be seen for some, in particular in Europe and Asia. In parallel, we observe a redistribution of the modeled GPP at large scale (Figure 3i), with typically an increase in the U.S. corn belt mostly as well as over boreal regions but to a lower extent, and a decrease elsewhere (even more pronounced in South Africa and South America). Where simulated SIF and GPP are both decreased, their joint reduction seems proportional over most regions (varying between 10% and 40%), but a higher decrease in GPP (above 65%) than in SIF (about 50%) can be seen over the tropics (South American and Africa). The large regions in Africa and South America surrounding the tropical rain forest (composed mainly of TrDBF, TrC3GRA, C4GRA, and C3 and C4 crops) are those where the decrease both in SIF and GPP is the most important. Over the tropical rain forests (Amazonia and Africa), noticeably, the model optimization mostly leads to an increased annual SIF and an opposite reduction of yearly GPP (which goes along with an enhanced agreement of ORCHIDEE with FLUXCOM-GPP, as already noticed above with Figure 2).

The reduction in RMSD with respect to FLUXCOM-GPP shows more contrasted results than for SIF with a larger number of pixels with increased RMSD (blue pixels in Figure 3j), in particular in the U.S. corn belt and Siberia where GPP is increased. Across all grid points, the RMSD between ORCHIDEE-SIF and OCO-2 SIF reduced from 0.26 to 0.23 W m^{-2} $\mu m^{-1}.sr^{-1}$ over 2015–2016 (Figure 3e), and from 2.09 to 1.97 gC $m^{-2}.d^{-1}$ between ORCHIDEE-GPP and FLUXCOM-GPP for year 2008 (Figure 3j). This corresponds to a large reduction of the global-scale mean annual GPP.

Over the 1990–2009 period, the optimization of ORCHIDEE resulted in a mean GPP reduction of about 28 $GtC.yr^{-1}$ at global scale (from 162 to 134 GtC/year), which is more in agreement with that of FLUXCOM (mean value of 121 GtC/year). Over the 2001–2009 period, the global-scale optimized GPP is in an even closer agreement with FLUXSAT than with FLUXCOM: the ORCHIDEE GPP budget (136 GtC/year) is comparable to that of FLUXSAT (137 GtC/year), whereas FLUXCOM estimates is lower (121 GtC/year). The positive trend of the model was decreased from 0.56 to 0.43 $GtC.yr^{-2}$ for both 1990–2009 and 2001–2009 periods. It is hence in closer agreement with mean linear trends estimated in the literature from several models (for instance 0.41 GtC/yr^2 both over the period 1990–2009 in Anav et al. (2015) and over the period 2000–2010 in Chen et al. (2017), 0.35 $GtC.yr^{-2}$ over the 1981–2010 period in Ito et al. (2017), 0.39 $GtC.yr^{-2}$ over 2000–2013 in Zhang et al. (2017)), but departs from the higher positive trend of FLUXSAT (1.35 GtC/yr^2 over 2001–2009). That positive trend in ORCHIDEE is due to the fertilization effect of increasing CO_2 atmospheric concentrations to which the model is sensitive (Sun et al., 2019), although moderated by a downregulation effect that mimics the role of nitrogen limitation on photosynthesis in the version used, and to the increasing surface temperatures (Anav et al., 2015). In parallel, the interannual variation (IAV) was increased from 1.67 to 1.93 $GtC.yr^{-1}$ over the 1990–2009 period, which remains higher than FLUXCOM IAV (0.31); over 2001–2009 however, ORCHIDEE-GPP IAV was decreased from 1.43 to 1.19 GtC/year which is closer to that estimated in FLUXSAT (1.25 GtC/year) but departs from the mean IAV of an ensemble of land surface models during the 2000–2010 period in Chen et al. (2017) (1.48).

4.2.3. GPP at the Regional and Ecozone Scales

The latitudinal and ecozonal changes in GPP monthly averages following the assimilation of OCO-2-SIF (2015–2016) are detailed over seven land regions in Figure 4 and illustrated for year 2008. We focused on three latitudinal bands (temperate north (30°–60°), tropics north (0°–30°), and tropics south (0°–30°)) as well as four bioclimatic zones inferred from the grouping of several Koppen Geiger classes derived from Peel et al. (2007). The GPP annual budget simulated by ORCHIDEE was reduced after assimilation for all regions except boreal biomes where a slight increase is observed. The change in GPP varies from only +0.9% for

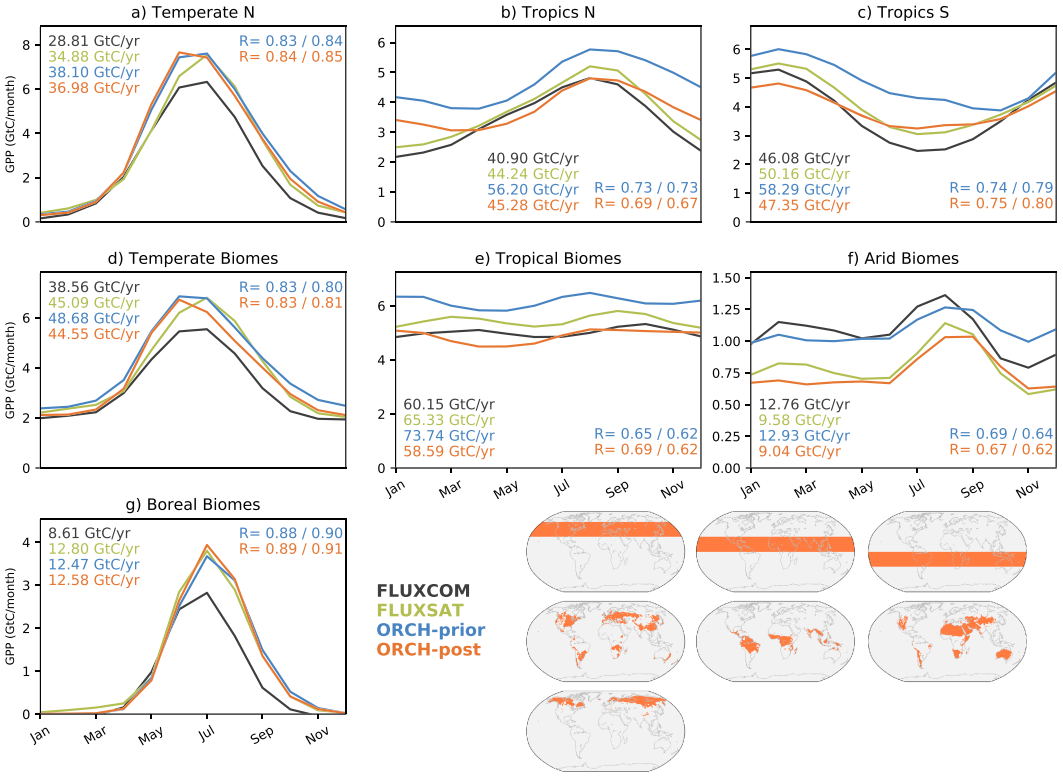

Figure 4. Monthly variations of GPP (GtC/month) over three latitudinal zones bands ((a) temperate north, (b) tropics north, (c) tropics south) and four bioclimatic regions ((d) temperate, (e) tropical, (f) arid, and (g) boreal biomes) for year 2008 derived from FLUXCOM (black) and FLUXSAT (green), and compared to ORCHIDEE simulations before (blue) and after (orange) grid-point-scale assimilation of OCO-2 SIF estimates. The annual budget associated to each GPP estimate is provided for each region as well as the correlation coefficient (R) between ORCHIDEE simulations and FLUXCOM/FLUXSAT, respectively. The map inserts illustrate each analyzed region. Three latitudinal bands are considered: temperate north (30°–60°), tropics north (0°–30°), and tropics south (0°–30°). The bioclimatic zones are grouped in the following way: "tropical biomes" correspond to class A in the Koppen Geiger scheme (Peel et al., 2007); "arid biomes" correspond to class B, "temperate biomes" correspond to class C plus the a and b subgroups of the D (cold) class (i.e., with hot and warm summer, respectively), and "boreal biomes" correspond to the c and d subcategories of the D class (i.e., cold and very cold winters).

the boreal ecozone to −8% for temperate biomes, −20% for tropical ecosystems, and up to −31% for arid biomes. For the tropics north and south regions, the decrease in GPP occur all year long. Although this result seems to contradict the previous findings on the increase of the modeled GPP for TeBDF or BoDNF, at the grid-point scale, this effect is largely compensated at the larger scales by the decreasing of GPP for the other PFTs present in each grid cell (notably TeC3GRA). The optimization results in a closer agreement (as compared to the prior simulations) of the GPP time series simulated by ORCHIDEE with respect to both FLUXCOM and FLUXSAT for all regions, except for arid biomes in the case of FLUXCOM. FLUXSAT and the optimized ORCHIDEE model exhibit closer GPP budgets for year 2008 than with FLUXCOM over all regions/biomes, and noteworthy over arid ecosystems (about 9 GtC for year 2008). However, the assimilation also results in a slight decrease of the correlation between ORCHIDEE time series and both FLUXCOM and FLUXSAT, except for boreal biomes.

4.3. SIF Constraint on Parameter Values and Uncertainties

The large changes in the optimized parameter values relative to the prior values for photosynthesis (mainly V_{cmax25}, $A1$, SLA—see Table 3 for their definition) and phenology (LAI_{MAX}, T_{sen}, $K_{LAI\text{-}happy}$, $T_{leaf\text{-}init}$, $L_{age\text{-}crit}$ for all PFTs but C3 and C4 grasses and crops) parameters highlight how SIF observations constrain different processes (Figure 5).

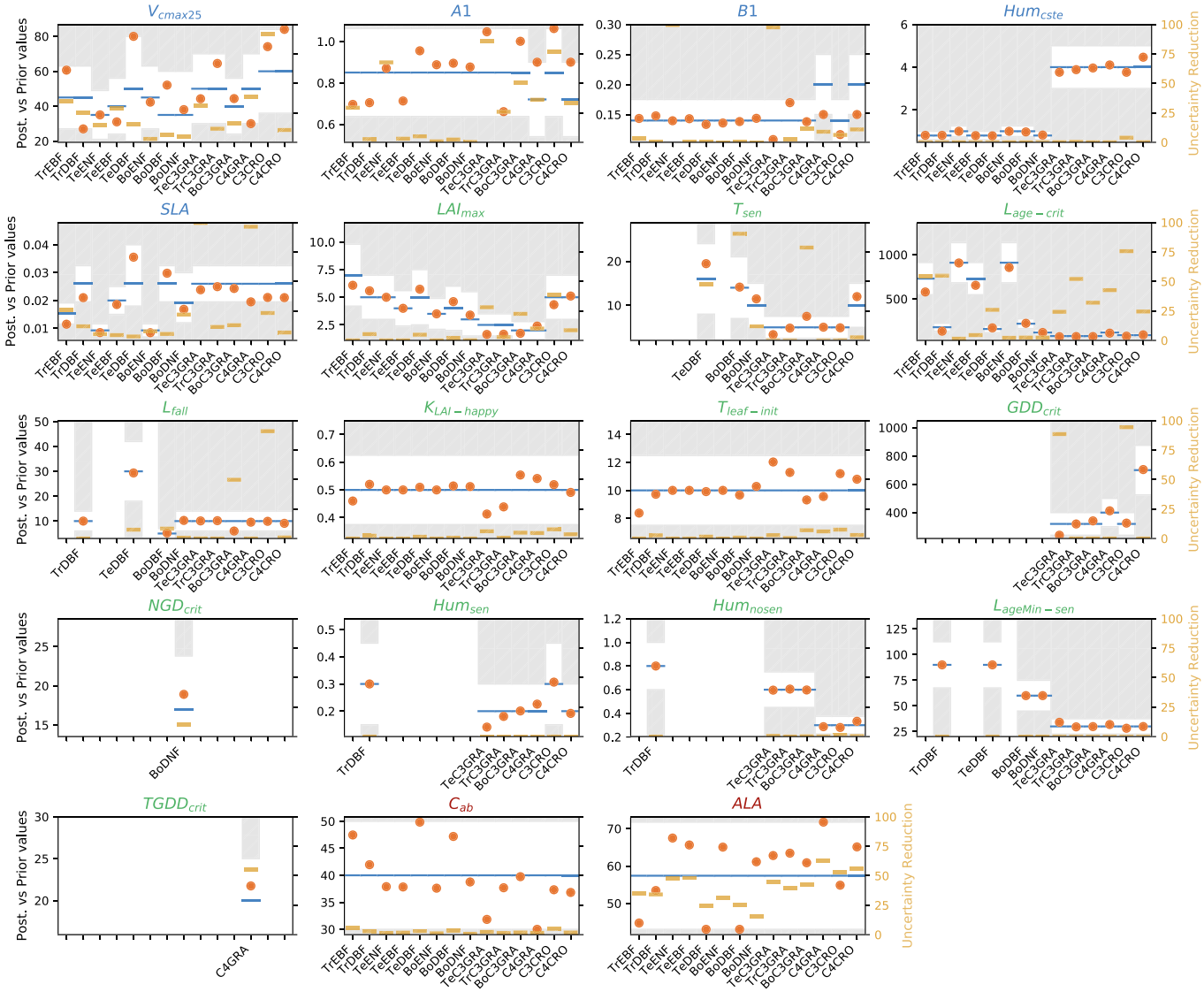

Figure 5. Prior (horizontal blue line) and posterior (orange dot) parameter values for each PFT (left *y* axis). The horizontal orange line is the uncertainty reduction (right *y* axis) expressed in percentage. Grey areas correspond to values outside the parameter variation range.

The increase of the modeled SIF magnitude required to match the OCO-2 retrievals for TrEBF, TeDBF, and BoDBF PFTs (Figure 2a) is tied to a noticeable increase of V_{cmax25}, which boosts the photosynthetic capacity. While a parallel increase in leaf area is obtained for TeDBF and BoDBF (through the increase of both *SLA* and LAI_{max}), a decrease is observed for TrEBF (*SLA* and LAI_{max} are decreased). For that matter, *SLA* and maximum LAI decreased for all PFTs except TeDBF and BoDBF (plus TrDBF, TeEBF, and C4 for maximum LAI but to a lesser extent). Noticeably, the largest joint increase of V_{cmax25} and *SLA* are obtained for two aforementioned PFTs, TeDBF and BoDBF, while *ALA* reaches its lower bound. The highest number of "edge-hitting" parameters—that is, parameters whose posterior values reach one of the limits of their predefined bounds—is also obtained for these PFTs. The impacted parameters are V_{cmax25}, C_{ab}, and *ALA*. It consequently limits the potential model improvement in terms of increase of carbon uptake and explains the residual underestimation of the modeled SIF amplitude regarding to OCO-2 time series for these PFTs. The general increment in T_{sen} (for all PFTs but TeC3GRA) and the general decrease of both $L_{age\text{-}crit}$ and L_{fall} contribute to advance the start of senescence and increase the leaf turnover. The changes in the parameters controlling the start of the growing season and the increase in leaf biomass ($K_{LAI\text{-}happy}$ and $T_{leaf\text{-}}$

$_{init}$) exhibit a general increase over the PFTs, except noticeably for temperate and tropical C3 grasses. Overall, the result is a shortening of the growing season for deciduous (temperate and boreal) and grass and crop PFTs.

The SIF-related parameters C_{ab} and ALA show strong and often opposite changes, with a large variability between PFTs. We recall that these two parameters only contribute to the SIF observation operator without any feedback on the photosynthesis and phenology processes. There are several cases where opposite changes in parameters result in opposing changes in the modeled SIF magnitude: for instance, the increase/decrease of V_{cmax25}/SLA which results in a SIF decrease for BoDNF, TrC3GRA, BoC3GRA, and C3 and C4 crops, while it corresponds to a SIF increase for TrEBF. These opposite changes highlight nonlinear interactions between optimized parameters and confounding effects (e.g., MacBean et al., 2018). This makes it difficult to interpret which parameters contribute the most to the model improvement relative to SIF data, and hence the change in GPP. However, the relative comparison of the uncertainty reductions between parameters and PFTs somewhat inform on those who are the most constrained by SIF data. The highest uncertainty reductions (above 50%) are achieved for the V_{cmax25}, $A1$, SLA, T_{sen}, $L_{age\text{-}crit}$, and ALA parameters, but not always for the same PFT. The lowest uncertainty reductions (below 10%) are obtained also for $B1$, as well as for the phenology-related parameters $K_{LAI\text{-}happy}$, $Tl_{eaf\text{-}init}$, GDD_{crit}, Hum_{sen}, Hum_{nosen}, and $L_{ageMin\text{-}sen}$, and for the C_{ab} parameter of the SIF model. The reasons for these low uncertainty reductions may however be different. They may express a low sensitivity of the model to some of the parameters (GDD_{crit}, Hum_{nosen}, LAI_{MAX}, for some PFTs for instance, or C_{ab} that has little influence on the TOC fluorescence in the near infrared).

5. Discussion

5.1. SIF Constraints on GPP-Related Parameters and Processes

Our assimilation framework of SIF data has achieved a strong reduction in the global mean annual GPP simulated with the ORCHIDEE model. At global scale, the reduction of GPP amounts to 28 GtC.yr^{-1} on average over the 1990–2009 period (from 162 to 134 GtC/year). For the sole year 2008, the reduction is of 28.8 GtC, which is comparable to the correction achieved by MacBean et al. (2018) by assimilating GOME-2 SIF data over 2007–2011 with ORCHIDAS also, based however on a simple statistical observation operator that does not use any mechanistic SIF model, and a different setup (with an older version of ORCHIDEE).

Although the analysis of the estimated parameter values after assimilation does not enable to really identify the main parameters causing the change in GPP for all PFTs, the reduction in GPP budget are attributed to a decrease of GPP magnitude (due to the lowering of photosynthetic capacities and leaf area) and/or of the seasonal cycle length depending on PFT. The decrease in GPP trend is attributed to the parameters controlling the GPP magnitude, and which optimization resulted in a reduced sensitivity of GPP to the increase of atmospheric CO_2 concentration. The difficulty to identify the driving parameters is revealed by error correlations between optimized parameters and relatively low and equal sensitivity of SIF to some parameters. These aspects partly explain the rather low uncertainty reduction on parameters, which is generally below 50%. Our results differ from those of MacBean et al. (2018) who found higher constraint levels on ORCHIDEE parameters brought by SIF observations, which was also partly due to their larger prior parameter bounds.

5.2. Leaf-Scale Modeling of SIF

The development and optimization of our parametric model of NPQ relies on leaf-scale fluorescence measurements over a limited range of vegetation species. In addition, we have made the intrinsic hypothesis that the processes driving non-photochemical quenching will operate similarly across all vegetation types and climates. Extensive measurement campaigns would be required to validate this assumption and, if need be, to improve our modeling framework. The parametric model also does not enable us to represent various factors that could impact NPQ and the link between SIF and GPP: (1) stresses other than light-limiting photosynthesis (e.g., water stress; van der Tol et al., 2014; which is however for the moment directly applied on V_{cmax} and on the stomatal conductance), (2) the possible sustained decrease of photochemical quenching due to photoinhibition, (3) the partitioning of NPQ between sustainable and reversible mechanisms that operates at different time scales and with different kinetics (Porcar-Castell et al., 2014), or (4) the possible variation in the partitioning of energy between PSI and PSII (parameter c, in equation (4)) and its potential effect

on PSI contribution, NPQ, and coupling between SIF and GPP. Obviously, the scientific community would strongly benefit from a full characterization of these factors across ecosystems and in response to different sources of stress, and the subsequent development of a physical model of NPQ that would incorporate these processes.

The values of the rate constants (k_D and k_F) and so of the maximum yield of photochemistry ($\Phi P_{PSII\ MAX}$), here derived from SCOPE, are empirical and are also subject to uncertainty: a value of 0.8 for $\Phi P_{PSII\ MAX}$ was here consistently used in our model, whereas this parameter is known to vary slightly between species with mean values around 0.83 (Baker, 2008; Björkman & Demmig, 1987). The parameters could be even biome specific (Frankenberg & Berry, 2018). The sensitivity of the ORCHIDEE SIF model to those parameterizations will be further investigated as they impact the upscaling from SIF to GPP. These parameters were not optimized in this study, to reduce the complexity of the optimization, but they should be considered in the future.

5.3. Canopy-Scale Modeling of SIF

The radiative transfer to simulate top of canopy SIF, based on SCOPE, was simplified in order to speed up the computation time. Although it is simpler than a recent machine learning approach applied to the SCOPE model (Verrelst et al., 2017), the fairly good performance of the optimized parametric SIF model to reproduce SCOPE SIF simulations validates the use of this approach.

The SIF of plant canopies simulated by ORCHIDEE assumes that it is essentially determined by the illuminated leaves of the top layer. This simplification is questionable given that the distribution between illuminated and shaded leaves may influence the total fluorescence flux. It is however justified by the fact that the major part of the canopy fluorescence originates from the top layers. For example, Thum et al. (2017) showed that more than 85% of the SIF is produced in the uppermost layer of their model. Yang and van der Tol (2018) also showed that SIF observed at the top of canopy at 760 nm largely originates from first-order scattering and thus from sunlit leaves, mainly situated at the top layers.

SIF simulations with ORCHIDEE were performed at the time of the satellite overpass. Possible disagreements between meteorological forcing and actual meteorological conditions at the synoptic scale (e.g., instantaneous cloud coverage impacting the PAR) were smoothed in the monthly averaged output.

The major limitation of the ORCHIDEE SIF model, as highlighted earlier, is its lack of biome dependency (a feature that is also not intrinsically represented in SCOPE). Actually, differences in leaf optical properties and canopy structure between vegetation species influence the radiation regime within the canopy and hence the escape probability of emitted SIF photons (Damm et al., 2015; Guanter et al., 2014; Sun et al., 2018; Zhang et al., 2016). We have initially chosen to have a representation of the SIF radiative transfer that is generic across all PFTs by not constraining the prior distribution and uncertainty of the two SIF-related parameters C_{ab} and *ALA,* hence relying on the data assimilation system to handle possible PFT dependencies. Note that this approach is somewhat similar to MacBean et al. (2018) who did not consider any a priori PFT dependency of the two empirical parameters of the SIF-GPP relationship, or to Norton et al. (2018) who assumed the structural parameters of SCOPE to be global. The optimized values of C_{ab} and *ALA* show large differences between PFTs. However, correlations between parameters during assimilation do not allow us to constrain sufficiently the distribution of the parameter values given that opposing changes in SIF and GPP are observed for some PFTs. This implementation can be improved given new evidence of different biome dependencies in the SIF-GPP relationship (Guanter et al., 2012; Sun et al., 2018; Zhang et al., 2016). That biome dependency may be attributable to both different physiological regulation mechanisms and differences in the radiative transfer within the canopy due to specific biome characteristics in terms of leaf optical properties and canopy structure that influence the amount of re-absorption and scattering of the chlorophyll fluorescence. The influence of structural effects is for instance supported by the recent work of Zhang, Zhang, et al., 2018), who observed directional variations in SIF monitored by OCO-2 in target mode, the behavior of which depended on the vegetation type. Therefore, with the current framework, it seems necessary to prescribe biome-specific prior parameter values and bounds given that they indirectly impact GPP due to parameter interactions within the assimilation procedure. Additionally, introducing a leaf-clumping parameter that characterizes the grouping of leaves within the canopy (Nilson, 1971) could bolster the representation of PFT-specific structural effects on top of canopy SIF, given that ecosystems exhibit distinct clumping index values (Chen et al.,

2005). Note that, in SCOPE, the clumping index equals 1, which corresponds to a random grouping of leaves. The implementation of a leaf clumping parameter is feasible with the current version of ORCHIDEE relying on a big-leaf approach, but would necessitate cautious constraining of the covariations of that parameter and those of C_{ab} and ALA.

5.4. SIF Model Coupling With ORCHIDEE

In the current implementation, C_{ab} and ALA do not impact the photosynthesis module in ORCHIDEE through the calculation of the absorbed PAR or the extinction coefficient of incoming radiation within the canopy. They therefore help to absorb possible bias in the observation data (Bacour et al., 2019; Köhler et al., 2015; Zhang, Joiner, et al., 2018) and act similarly to the offset parameter of the SIF-GPP linear model in MacBean et al. (2018). The next logical step is to have a consistent accounting of these parameters between the fluorescence and photosynthesis modules.

In the future, the radiative modeling of SIF in ORCHIDEE will benefit from and evolve with the model development. In particular, the ORCHIDEE-CANopy version (Naudts et al., 2015) that represents the vertical distribution of leaves use a more physical description of the radiative transfer within the canopy (two stream model; Pinty et al., 2006). Similarly, we should also include the consideration of the quality of the incoming photosynthetically active radiation (direct versus diffuse light) that largely influences the SIF emission. Also, the upcoming version of ORCHIDEE will include the nitrogen cycle (Vuichard et al., 2018) and therefore will better represent nutrient limitations on CO_2 fertilization (He et al., 2017) . This should enable to better represent the covariations of V_{cmax} and leaf chlorophyll content, and hence the impact of C_{ab} on the SIF signal. Indeed, leaf chlorophyll content is related to leaf nitrogen content (Gholizadeh et al., 2017; Yoder & Pettigrew-Crosby, 1995) with a dependence that may however vary depending on the season, species, and allocation of nitrogen for different plant processes (Croft et al., 2017; Ghimire et al., 2017; Kalacska et al., 2015). In parallel, a more direct relationship between C_{ab} and V_{cmax} shall be implemented to represent the dynamic of leaf chlorophyll content (Croft et al., 2017; Houborg et al., 2013; Migliavacca et al., 2017).

6. Summary and Conclusion

We have presented the development of a process-based model for solar-induced fluorescence implemented in the ORCHIDEE TBM. It relies (1) on the simulation of the regulation of the PSII fluorescence quantum yield at the leaf level, following the lake model formalism and using a novel parameterization of NPQ as a function of temperature, PAR, and the normalized quantum yield of photochemistry, and (2) a parametric simplification of the SCOPE model to emulate the radiative transfer of chlorophyll fluorescence to the top of the canopy. The prior ORCHIDEE-SIF model enabled to simulate levels and temporal variations of SIF already in agreement with those recorded by OCO-2 for most PFTs.

We applied a Bayesian assimilation framework to assimilate OCO-2-SIF products and hence optimize ORCHIDEE over an ensemble of grid points for all vegetation PFTs. It resulted in a substantial improvement of the simulated amplitudes and temporal seasonality for all PFTs, and of the simulated spatio-temporal patterns of yearly SIF with respect to OCO-2 observations.

Through the optimization of photosynthesis- and phenology-related ORCHIDEE parameters the impact on the simulation of GPP was considerable with a decrease of the global budget by 28/29 GtC.yr^{-1} in average over the 1990–2009/2001–2009 periods (from 162/165 to 134/136 GtC/year). The global GPP budget of the optimized model was then in closer agreement with that of FLUXCOM estimates (121 GtC/year over 1990–2009) and even more with that of FLUXSAT (137 GtC/year over 2001–2009). The reasons for this large change vary between PFTs. For deciduous trees and C3 and C4 grasses and crops, it is attributed to a decrease of the growing season length; for tropical forests, temperate and boreal evergreen trees, and C3 and C4 grasses and crops, it can also be related to a reduction of the leaf area. At global scale, this translated into a decrease of the gross primary productivity over the tropics principally, Europe, and over a large area covering the northwestern part of North America; at the opposite, an increase in GPP was obtained in the U.S. corn belt and over boreal Asia.

Despite the strong parallel improvements of the optimized model both regarding OCO-2 SIF data and independent FLUXCOM-GPP estimates, the assimilation however failed to reconcile these two data sets and the ORCHIDEE model for temperate and boreal deciduous forests. Indeed, the increase of the simulated

SIF amplitude required to match OCO-2 time series for those PFTs resulted in a parallel increasing of the modeled GPP which in turn degraded the model agreement with FLUXCOM-GPP estimates.

Uncertainties in the FLUXCOM-GPP and OCO-2 SIF estimates put aside, this may point out some model deficiencies that can be explained by an additional biome dependency of the SIF-GPP relationship (Guanter et al., 2012; Sun et al., 2018; Wood et al., 2017; Zhang et al., 2016) not represented in ORCHIDEE (nor in the SCOPE model).

The validation of SIF models and of their representation/parameterization of the processes regulating plant photosynthesis and fluorescence emission is critical for using SIF estimates in data assimilation systems. The biome dependency in particular needs to be resolved to enhance the observational constraint brought by space-borne SIF products on vegetation productivity and reinforce the predictive capabilities of TBMs relative to the fate of the carbon cycle.

Acknowledgments
The authors are very grateful to the whole ORCHIDEE development team (and to Josefine Ghattas in particular) and to LSCE's computer team for the computing support and resources. This study was supported by CNES-TOSCA under the FluOR and ECOFLUO projects, the Copernicus Atmosphere Monitoring Service (CAMS41 project) implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Commission, and the European Union's Horizon 2020 research and innovation programme under Grant Agreement number 776810. A. Porcar-Castell was supported by the Academy of Finland (288039). The authors thank Joris Timmermans (University College London, UK) and Yongguang Zhang (Nanjing University, China) for their valuable feedback on the use of SCOPE. The Hyper Graeco Latin Geometric simulation table was determined by Yves Tourbier (Renault SAS, France). The OCO-2 B7300 SIF data are available at ftp://fluo.gps.caltech.edu/data/OCO2/sif_lite_B7300/. The FLUXCOM products can be obtained from the Data Portal of the Max Planck Institute for Biochemistry (https://www.bgc-jena.mpg.de/geodb/projects/Data.php). The FLUXSAT GPP product is available at https://avdc.gsfc.nasa.gov. The SCOPE model is available from the Github repository (https://github.com/Christiaanvandertol/SCOPE). The code of ORCHIDEE is managed using the Subversion (svn) version control system, and in this study, we used revision 4810 of the svn trunk; details on how downloading and using the model are provided at http://forge.ipsl.jussieu.fr/orchidee.

References

Amarnath, K., Bennett, D. I., Schneider, A. R., & Fleming, G. R. (2016). Multiscale model of light harvesting by photosystem II in plants. *Proceedings of the National Academy of Sciences*, *113*(5), 1156–1161.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., et al. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. *Reviews of Geophysics*, *53*, 785–818. https://doi.org/10.1002/2015RG000483

Bacour, C., Jacquemoud, S., Tourbier, Y., Dechambre, M., & Frangi, J.-P. (2002). Design and analysis of numerical experiments to compare four canopy reflectance models. *Remote Sensing of Environment*, *79*(1), 72–83.

Bacour, C., Maignan, F., Peylin, P., MacBean, N., Bastrikov, V., Joiner, J., et al. (2019). Differences between OCO2 and GOME2 SIF products from a model-data fusion perspective. *Journal of Geophysical Research: Biogeosciences*, in revision. https://doi.org/10.1029/2018JG004938

Bacour, C., Peylin, P., MacBean, N., Rayner, P. J., Delage, F., Chevallier, F., et al. (2015). Joint assimilation of eddy covariance flux measurements and FAPAR products over temperate forests within a process-oriented biosphere model. *Journal of Geophysical Research: Biogeosciences*, *120*, 1839–1857. https://doi.org/10.1002/2015JG002966

Baker, N. R. (2008). Chlorophyll fluorescence: A probe of photosynthesis in vivo. *Annual Review of Plant Biology*, *59*, 89–113. https://doi.org/10.1146/annurev.arplant.59.032607.092759

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

Björkman, O., & Demmig, B. (1987). Photon yield of O_2 evolution and chlorophyll fluorescence characteristics at 77 K among vascular plants of diverse origins. *Planta*, *170*(4), 489–504. https://doi.org/10.1007/BF00402983

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

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

Campbell, G. S. (1990). Derivation of an angle density function for canopies with ellipsoidal leaf angle distributions. *Agricultural and Forest Meteorology*, *49*(3), 173–176.

Chen, J. M., Menges, C. H., & Leblanc, S. G. (2005). Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sensing of Environment*, *97*(4), 447–457.

Chen, M., Rafique, R., Asrar, G. R., Bond-Lamberty, B., Ciais, P., Zhao, F., et al. (2017). Regional contribution to variability and trends of global gross primary productivity. *Environmental Research Letters*, *12*(10). https://doi.org/10.1088/1748-9326/aa8978

Collatz, G. J., Ribas-Carbo, M., & Berry, J. A. (1992). Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Functional Plant Biology*, *19*(5), 519–538.

Cramer, W. (1997). *Using plant functional types in a global vegetation model. Plant Functional Types: Their Relevance to Ecosystem Properties and Global Change*. Cambridge: Cambridge University Press.

Croft, H., Chen, J. M., Luo, X., Bartlett, P., Chen, B., & Staebler, R. M. (2017). Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology*, *23*(9), 3513–3524. https://doi.org/10.1111/gcb.13599

Damm, A., Guanter, L., Paul-Limoges, E., van der Tol, C., Hueni, A., Buchmann, N., et al. (2015). Far-red Sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modelling approaches. *Remote Sensing of Environment*, *166*, 91–105. https://doi.org/10.1016/j.rse.2015.06.004

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., et al. (2013). Climate change projections using the IPSL-CM5 Earth system model: From CMIP3 to CMIP5. *Climate Dynamics*, *40*(9–10), 2123–2165. https://doi.org/10.1007/s00382-012-1636-1

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta*, *149*(1), 78–90. https://doi.org/10.1007/BF00386231

Feret, J.-B., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P. R., et al. (2008). PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, *112*(6), 3030–3043. https://doi.org/10.1016/j.rse.2008.02.012

Flexas, J., Escalona, J. M., Evain, S., Gulías, J., Moya, I., Osmond, C. B., & Medrano, H. (2002). Steady-state chlorophyll fluorescence (Fs) measurements as a tool to follow variations of net CO_2 assimilation and stomatal conductance during water-stress in C3 plants. *Physiologia Plantarum*, *114*(2), 231–240. https://doi.org/10.1034/j.1399-3054.2002.1140209.x

Fox, A., Williams, M., Richardson, A. D., Cameron, D., Gove, J. H., Quaife, T., et al. (2009). The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. *Agricultural and Forest Meteorology*, *149*(10), 1597–1615. https://doi.org/10.1016/j.agrformet.2009.05.002

Frankenberg, C., & Berry, J. (2018). Solar induced chlorophyll fluorescence: Origins, relation to photosynthesis and retrieval. In *Reference Module in Earth Systems and Environmental Sciences: Comprehensive Remote Sensing*, (pp. 143–162). Oxford.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., et al. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, *38*, L17706. https://doi.org/10.1029/2011GL048738

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

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, *27*(2), 511–526. https://doi.org/10.1175/JCLI-D-12-00579.1

Galmés, J., Medrano, H., & Flexas, J. (2007). Photosynthetic limitations in response to water stress and recovery in Mediterranean plants with different growth forms. *New Phytologist*, *175*(1), 81–93. https://doi.org/10.1111/j.1469-8137.2007.02087.x

Ghimire, B., Riley, W. J., Koven, C. D., Kattge, J., Rogers, A., Reich, P. B., & Wright, I. J. (2017). A global trait-based approach to estimate leaf nitrogen functional allocation from observations. *Ecological Applications*, *27*(5), 1421–1434. https://doi.org/10.1002/eap.1542

Gholizadeh, A., Saberioon, M., Borůvka, L., Wayayok, A., & Soom, M. A. M. (2017). Leaf chlorophyll and nitrogen dynamics and their relationship to lowland rice yield for site-specific paddy management. *Information Processing in Agriculture*, *4*(4), 259–268.

Grace, J., Nichol, C., Disney, M., Lewis, P., Quaife, T., & Bowyer, P. (2007). Can we measure terrestrial photosynthesis from space directly, using spectral reflectance and fluorescence? *Global Change Biology*, *13*(7), 1484–1497.

Groenendijk, M., Dolman, A. J., van der Molen, M. K., Leuning, R., Arneth, A., Delpierre, N., et al. (2011). Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data. *Agricultural and Forest Meteorology*, *151*(1), 22–38. https://doi.org/10.1016/j.agrformet.2010.08.013

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

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., et al. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences*, *111*(14), E1327–E1333. https://doi.org/10.1073/pnas.1320008111

He, L., Chen, J. M., Croft, H., Gonsamo, A., Luo, X., Liu, J., et al. (2017). Nitrogen availability dampens the positive impacts of CO_2 fertilization on terrestrial ecosystem carbon and water cycles. *Geophysical Research Letters*, *44*, 11–590. https://doi.org/10.1002/2017GL075981

Hendrickson, L., Furbank, R. T., & Chow, W. S. (2004). A simple alternative approach to assessing the fate of absorbed light energy using chlorophyll fluorescence. *Photosynthesis Research*, *82*(1), 73–81. https://doi.org/10.1023/B:PRES.0000040446.87305.f4

Houborg, R., Cescatti, A., Migliavacca, M., & Kustas, W. P. (2013). Satellite retrievals of leaf chlorophyll and photosynthetic capacity for improved modelling of GPP. *Agricultural and Forest Meteorology*, *177*, 10–23.

Ito, A., Nishina, K., Reyer, C. P. O., François, L., Henrot, A. J., Munhoven, G., et al. (2017). Photosynthetic productivity and its efficiencies in ISIMIP2a biome models: Benchmarking for impact assessment studies. *Environmental Research Letters*, *12*(8). https://doi.org/10.1088/1748-9326/aa7a19

Jacquemoud, S., & Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, *34*(2), 75–91.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., et al. (2009). PROSPECT+ SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, *113*, S56–S66. https://doi.org/10.1016/j.rse.2008.01.026

Jeong, S.-J., Schimel, D., Frankenberg, C., Drewry, D. T., Fisher, J. B., Verma, M., et al. (2017). Application of satellite solar-induced chlorophyll fluorescence to understanding large-scale variations in vegetation phenology and function over northern high latitude forests. *Remote Sensing of Environment*, *190*, 178–187. https://doi.org/10.1016/j.rse.2016.11.021

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., et al. (2013). Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: Methodology, simulations, and application to GOME-2. *Atmospheric Measurement Techniques*, *6*(10), 2803–2823. https://doi.org/10.5194/amt-6-2803-2013

Joiner, J., Yoshida, Y., Vasilkov, A. P., Schaefer, K., Jung, M., Guanter, L., et al. (2014). The seasonal cycle of satellite chlorophyll fluorescence observations and its relationship to vegetation phenology and ecosystem atmosphere carbon exchange. *Remote Sensing of Environment*, *152*, 375–391. https://doi.org/10.1016/j.rse.2014.06.022

Joiner, J., Yoshida, Y., Zhang, Y., Duveiller, G., Jung, M., Lyapustin, A., et al. (2018). Estimation of terrestrial global gross primary production (GPP) with satellite data-driven models and eddy covariance flux data. *Remote Sensing*, *10*(9), 1346. https://doi.org/10.3390/rs10091346

Jung, M., Vetter, M., Herold, M., Churkina, G., Reichstein, M., Zaehle, S., et al. (2007). Uncertainties of modelling gross primary productivity over Europe: A systematic study on the effects of using different drivers and terrestrial biosphere models. *Global Biogeochemical Cycles*, *21*, GB4021. https://doi.org/10.1029/2006GB002915

Kalacska, M., Lalonde, M., & Moore, T. R. (2015). Estimation of foliar chlorophyll and nitrogen content in an ombrotrophic bog from hyperspectral data: Scaling from leaf to image. *Remote Sensing of Environment*, *169*, 270–279.

Kaminski, T., Knorr, W., Rayner, P. J., & Heimann, M. (2002). Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle. *Global Biogeochemical Cycles*, *16*(4), 1066. https://doi.org/10.1029/2001GB001463

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., & Mathieu, P.-P. (2012). Consistent assimilation of MERIS FAPAR and atmospheric CO_2 into a terrestrial vegetation model and interactive mission benefit analysis. *Biogeosciences*, *9*(8), 3173–3184.

Kattge, J., & Knorr, W. (2007). Temperature acclimation in a biochemical model of photosynthesis: A reanalysis of data from 36 species. *Plant, Cell & Environment*, *30*(9), 1176–1190. https://doi.org/10.1111/j.1365-3040.2007.01690.x

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., & Mathieu, P.-P. (2010). Carbon cycle data assimilation with a generic phenology model. *Journal of Geophysical Research*, *115*, G04017. https://doi.org/10.1029/2009JG001119

Knorr, W., & Kattge, J. (2005). Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. *Global Change Biology*, *11*(8), 1333–1351.

Koffi, E. N., Rayner, P. J., Norton, A. J., Frankenberg, C., & Scholze, M. (2015). Investigating the usefulness of satellite-derived fluorescence data in inferring gross primary productivity within the carbon cycle data assimilation system. *Biogeosciences*, *12*(13), 4067–4084.

Koffi, E. N., Rayner, P. J., Scholze, M., & Beer, C. (2012). Atmospheric constraints on gross primary productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation system. *Global Biogeochemical Cycles*, *26*, GB1024. https://doi.org/10.1029/2010GB003900

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

Kramer, D. M., Johnson, G., Kiirats, O., & Edwards, G. E. (2004). New fluorescence parameters for the determination of QA redox state and excitation energy fluxes. *Photosynthesis Research*, *79*(2), 209. https://doi.org/10.1023/B:PRES.0000015391.99477.0d

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., et al. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*, GB1015. https://doi.org/10.1029/2003GB002199

Kuppel, S., Chevallier, F., & Peylin, P. (2013). Quantifying the model structural error in carbon cycle data assimilation systems. *Geoscientific Model Development*, *6*, 45–55.

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

Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., & Cescatti, A. (2014). Model–data fusion across ecosystems: From multisite optimizations to global simulations. *Geoscientific Model Development*, *7*(6), 2581–2597.

Lavergne, J., & Trissl, H.-W. (1995). Theory of fluorescence induction in photosystem II: Derivation of analytical expressions in a model including exciton-radical-pair equilibrium and restricted energy transfer between photosynthetic units. *Biophysical Journal*, *68*(6), 2474–2492. https://doi.org/10.1016/S0006-3495(95)80429-7

Lazár, D. (2015). Parameters of photosynthetic energy partitioning. *Journal of Plant Physiology*, *175*, 131–147. https://doi.org/10.1016/j.jplph.2014.10.021

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., et al. (2018). Global carbon budget 2017. *Earth System Science Data*, *10*, 405–448.

Lee, J.-E., Berry, J. A., van der Tol, C., Yang, X., Guanter, L., Damm, A., et al. (2015). Simulations of chlorophyll fluorescence incorporated into the Community Land Model version 4. *Global Change Biology*, *21*(9), 3469–3477. https://doi.org/10.1111/gcb.12948

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., et al. (2018). Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. *Scientific Reports*, *8*(1), 1973. https://doi.org/10.1038/s41598-018-20024-w

MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., & Ciais, P. (2015). Using satellite data to improve the leaf phenology of a global terrestrial biosphere model. *Biogeosciences*, *12*(23), 7185–7208.

MacBean, N., Peylin, P., Chevallier, F., Scholze, M., & Schürmann, G. (2016). Consistent assimilation of multiple data streams in a carbon cycle data assimilation system. *Geoscientific Model Development*, *9*(10), 3569–3588.

Madani, N., Kimball, J. S., Jones, L. A., Parazoo, N. C., & Guan, K. (2017). Global analysis of bioclimatic controls on ecosystem productivity using satellite observations of solar-induced chlorophyll fluorescence. *Remote Sensing*, *9*(6), 530.

Maxwell, K., & Johnson, G. N. (2000). Chlorophyll fluorescence—A practical guide. *Journal of Experimental Botany*, *51*(345), 659–668. https://doi.org/10.1093/jxb/51.345.659

Medlyn, B. E., Dreyer, E., Ellsworth, D., Forstreuter, M., Harley, P. C., Kirschbaum, M. U. F., et al. (2002). Temperature response of parameters of a biochemically based model of photosynthesis. II. A review of experimental data. *Plant, Cell & Environment*, *25*(9), 1167–1179. https://doi.org/10.1046/j.1365-3040.2002.00891.x

Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R., & Moreno, J. (2009). Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sensing of Environment*, *113*(10), 2037–2051.

Migliavacca, M., Meroni, M., Busetto, L., Colombo, R., Zenone, T., Matteucci, G., et al. (2009). Modelling gross primary production of agro-forestry ecosystems by assimilation of satellite-derived information in a process-based model. *Sensors*, *9*(2), 922–942. https://doi.org/10.3390/s90200922

Migliavacca, M., Perez-Priego, O., Rossini, M., el-Madany, T. S., Moreno, G., van der Tol, C., et al. (2017). Plant functional traits and canopy structure control the relationship between photosynthetic CO_2 uptake and far-red Sun-induced fluorescence in a Mediterranean grassland under different nutrient availability. *New Phytologist*, *214*(3), 1078–1091. https://doi.org/10.1111/nph.14437

Muraoka, H., & Koizumi, H. (2005). Photosynthetic and structural characteristics of canopy and shrub trees in a cool-temperate deciduous broadleaved forest: Implication to the ecosystem carbon gain. *Agricultural and Forest Meteorology*, *134*(1–4), 39–59.

Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., et al. (2015). A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes. *Geoscientific Model Development*, *8*(7), 2035–2065. https://doi.org/10.5194/gmd-8-2035-2015

Niinemets, Ü., Kull, O., & Tenhunen, J. D. (2004). Within-canopy variation in the rate of development of photosynthetic capacity is proportional to integrated quantum flux density in temperate deciduous trees. *Plant, Cell & Environment*, *27*(3), 293–313.

Nilson, T. (1971). A theoretical analysis of the frequency of gaps in plant stands. *Agricultural Meteorology*, *8*, 25–38.

Norton, A. J., Rayner, P. J., Koffi, E. N., Scholze, M., Silver, J. D., & Wang, Y.-P. (2018). Estimating global gross primary productivity using chlorophyll fluorescence and a data assimilation system with the BETHY-SCOPE model. *Biogeosciences Discussions*, *2018*, 1–40.

Pedrós, R., Goulas, Y., Jacquemoud, S., Louis, J., & Moya, I. (2010). FluorMODleaf: A new leaf fluorescence emission model based on the PROSPECT model. *Remote Sensing of Environment*, *114*(1), 155–167.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences Discussions*, *4*(2), 439–473.

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

Pinty, B., Lavergne, T., Dickinson, R. E., Widlowski, J.-L., Gobron, N., & Verstraete, M. M. (2006). Simplifying the interaction of land surfaces with radiation for relating remote sensing products to climate models. *Journal of Geophysical Research*, *111*, D02116. https://doi.org/10.1029/2005JD005952

Porcar-Castell, A. (2011). A high-resolution portrait of the annual dynamics of photochemical and non-photochemical quenching in needles of *Pinus sylvestris*. *Physiologia Plantarum*, *143*(2), 139–153. https://doi.org/10.1111/j.1399-3054.2011.01488.x

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., van der Tol, C., Flexas, J., Pfündel, E. E., et al. (2014). Linking chlorophyll *a* fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges. *Journal of Experimental Botany*, *65*(15), 4065–4095. https://doi.org/10.1093/jxb/eru191

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., et al. (2015). Plant functional type classification for earth system models: Results from the European Space Agency's Land Cover Climate Change Initiative. *Geoscientific Model Development*, *8*(7), 2315–2328. https://doi.org/10.5194/gmd-8-2315-2015

Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., & Solomon, A. M. (1992). Special paper: A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, 117–134.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., et al. (2005). Model–data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications. *Global Change Biology*, *11*(3), 378–397. https://doi.org/10.1111/j.1365-2486.2005.00917.x

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., & Widmann, H. (2005). Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Global Biogeochemical Cycles*, *19*, GB2026. https://doi.org/10.1029/2004GB002254

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., et al. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. *Global Change Biology*, *11*(9), 1424–1439. https://doi.org/10.1111/j.1365-2486.2005.001002.x

Ruuska, S. A., Badger, M. R., Andrews, T. J., & Von Caemmerer, S. (2000). Photosynthetic electron sinks in transgenic tobacco with reduced amounts of Rubisco: Little evidence for significant Mehler reaction. *Journal of Experimental Botany*, *51*(suppl_1), 357–368.

Sacks, W. J., Schimel, D. S., & Monson, R. K. (2007). Coupling between carbon cycling and climate in a high-elevation, subalpine forest: A model-data fusion analysis. *Oecologia*, *151*(1), 54–68. https://doi.org/10.1007/s00442-006-0565-2

Sanders, A. F., Verstraeten, W. W., Kooreman, M. L., Van Leth, T. C., Beringer, J., & Joiner, J. (2016). Spaceborne Sun-induced vegetation fluorescence time series from 2007 to 2015 evaluated with Australian flux tower measurements. *Remote Sensing*, *8*(11), 895.

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

Santaren, D., Peylin, P., Viovy, N., & Ciais, P. (2007). Optimizing a process-based ecosystem model with eddy-covariance flux measurements: A pine forest in southern France. *Global Biogeochemical Cycles*, *21*, GB2013. https://doi.org/10.1029/2006GB002834

Schürmann, G. J., Kaminski, T., Köstler, C., Carvalhais, N., Voßbeck, M., Kattge, J., et al. (2016). Constraining a land-surface model with multiple observations by application of the MPI-Carbon Cycle Data Assimilation System v1.0. *Geoscientific Model Development*, *9*(9), 2999–3026. https://doi.org/10.5194/gmd-9-2999-2016

Šesták, Z., & Čatský, J. (1962). Intensity of photosynthesis and chlorophyll content as related to leaf age in *Nicotiana sanderae* hort. *Biologia Plantarum*, *4*(2), 131.

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

Smith, W. K., Biederman, J. A., Scott, R. L., Moore, D. J. P., He, M., Kimball, J. S., et al. (2018). Chlorophyll fluorescence better captures seasonal and interannual gross primary productivity dynamics across dryland ecosystems of southwestern North America. *Geophysical Research Letters*, *45*, 748–757. https://doi.org/10.1002/2017GL075922

Stöckli, R., Rutishauser, T., Dragoni, D., O'Keefe, J., Thornton, P. E., Jolly, M., et al. (2008). Remote sensing data assimilation for a prognostic phenology model. *Journal of Geophysical Research*, *113*, G04021. https://doi.org/10.1029/2008JG000781

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Köhler, P., & Magney, T. (2018). Overview of solar-induced chlorophyll fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP. *Remote Sensing of Environment*, *209*, 808–823.

Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D. S., Jung, M., Guanter, L., et al. (2017). OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. *Science*, *358*(6360), eaam5747. https://doi.org/10.1126/science.aam5747

Sun, Y., Fu, R., Dickinson, R., Joiner, J., Frankenberg, C., Gu, L., et al. (2015). Drought onset mechanisms revealed by satellite solar-induced chlorophyll fluorescence: Insights from two contrasting extreme events. *Journal of Geophysical Research: Biogeosciences*, *120*, 2427–2440. https://doi.org/10.1002/2015JG003150

Sun, Z., Wang, X., Zhang, X., Tani, H., Guo, E., Yin, S., & Zhang, T. (2019). Evaluating and comparing remote sensing terrestrial GPP models for their response to climate variability and CO_2 trends. *Science of the Total Environment*. https://doi.org/10.1007/978-3-030-16729-5

Tarantola, A. (2005). Inverse problem theory and methods for model parameter estimation (Vol. 89). siam.

Thum, T., Zaehle, S., Köhler, P., Aalto, T., Aurela, M., Guanter, L., et al. (2017). Modelling Sun-induced fluorescence and photosynthesis with a land surface model at local and regional scales in northern Europe. *Biogeosciences*, *14*(7), 1969–1987. https://doi.org/10.5194/bg-14-1969-2017

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., et al. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, *13*(14), 4291–4313. https://doi.org/10.5194/bg-13-4291-2016

van der Tol, C., Berry, J. A., Campbell, P. K. E., & Rascher, U. (2014). Models of fluorescence and photosynthesis for interpreting measurements of solar-induced chlorophyll fluorescence. *Journal of Geophysical Research: Biogeosciences*, *119*, 2312–2327. https://doi.org/10.1002/2014JG002713

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

Verrelst, J., Rivera Caicedo, J. P., Muñoz-Marí, J., Camps-Valls, G., & Moreno, J. (2017). SCOPE-based emulators for fast generation of synthetic canopy reflectance and Sun-induced fluorescence spectra. *Remote Sensing*, (9), 927. https://doi.org/10.3390/rs9090927

Verrelst, J., Rivera, J. P., van der Tol, C., Magnani, F., Mohammed, G., & Moreno, J. (2015). Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving Sun-induced fluorescence? *Remote Sensing of Environment*, *166*, 8–21.

Viovy, N. (2018). CRUNCEP version 7—Atmospheric Forcing Data for the Community Land Model [Data set]. Retrieved from https://rda.ucar.edu/datasets/ds314.3/

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., et al. (2018). Accounting for Carbon and Nitrogen interactions in the Global Terrestrial Ecosystem Model ORCHIDEE (trunk version, rev 4999): Multi-scale evaluation of gross primary production. *Geoscientific Model Development Discussions*, 1–41. https://doi.org/10.5194/gmd-2018-261

Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Köhler, P., et al. (2016). Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. *Global Change Biology*, *22*(9), 2979–2996. https://doi.org/10.1111/gcb.13200

Wang, S., Huang, C., Zhang, L., Lin, Y., Cen, Y., & Wu, T. (2016). Monitoring and assessing the 2012 drought in the Great Plains: Analyzing satellite-retrieved solar-induced chlorophyll fluorescence, drought indices, and gross primary production. *Remote Sensing*, *8*(2), 61.

Wang, Y.-P., Leuning, R., Cleugh, H. A., & Coppin, P. A. (2001). Parameter estimation in surface exchange models using nonlinear inversion: How many parameters can we estimate and which measurements are most useful? *Global Change Biology*, *7*(5), 495–510.

Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., & Kurpius, M. R. (2005). An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology*, *11*(1), 89–105.

Wood, J. D., Griffis, T. J., Baker, J. M., Frankenberg, C., Verma, M., & Yuen, K. (2017). Multiscale analyses of solar-induced florescence and gross primary production. *Geophysical Research Letters*, *44*, 533–541. https://doi.org/10.1002/2016GL070775

Wullschleger, S. D., Epstein, H. E., Box, E. O., Euskirchen, E. S., Goswami, S., Iversen, C. M., et al. (2014). Plant functional types in Earth system models: Past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. *Annals of Botany*, *114*(1), 1–16. https://doi.org/10.1093/aob/mcu077

Yang, P., & Van Der Tol, C. (2018). Linking canopy scattering of far-red Sun-induced chlorophyll fluorescence with reflectance. *Remote Sensing of Environment*, *209*, 456–467.

Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., et al. (2015). Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. *Geophysical Research Letters*, *42*, 2977–2987. https://doi.org/10.1002/2015GL063201

Yin, X., & Struik, P. C. (2009). C3 and C4 photosynthesis models: An overview from the perspective of crop modelling. *NJAS-Wageningen Journal of Life Sciences*, *57*(1), 27–38.

Yoder, B. J., & Pettigrew-Crosby, R. E. (1995). Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400–2500 nm) at leaf and canopy scales. *Remote Sensing of Environment*, *53*(3), 199–211.

Zhang, Y., Guanter, L., Berry, J. A., van der Tol, C., Yang, X., Tang, J., & Zhang, F. (2016). Model-based analysis of the relationship between Sun-induced chlorophyll fluorescence and gross primary production for remote sensing applications. *Remote Sensing of Environment*, *187*, 145–155.

Zhang, Y., Joiner, J., Gentine, P., & Zhou, S. (2018). Reduced solar-induced chlorophyll fluorescence from GOME-2 during Amazon drought caused by dataset artifacts. *Global Change Biology*, *24*(6), 2229–2230. https://doi.org/10.1111/gcb.14134

Zhang, Y., Xiao, X., Wu, X., Zhou, S., Zhang, G., Qin, Y., & Dong, J. (2017). A global moderate resolution dataset of gross primary production of vegetation for 2000–2016. *Scientific Data*, *4*, 170,165.

Zhang, Y., Xiao, X., Zhang, Y., Wolf, S., Zhou, S., Joiner, J., et al. (2018). On the relationship between sub-daily instantaneous and daily total gross primary production: Implications for interpreting satellite-based SIF retrievals. *Remote Sensing of Environment*, *205*, 276–289. https://doi.org/10.1016/j.rse.2017.12.009

Zhang, Z., Zhang, Y., Joiner, J., & Migliavacca, M. (2018). Angle matters: Bidirectional effects impact the slope of relationship between gross primary productivity and Sun-induced chlorophyll fluorescence from Orbiting Carbon Observatory-2 across biomes. *Global Change Biology*, *24*(11), 5017–5020. https://doi.org/10.1111/gcb.14427

Zobitz, J. M., Moore, D. J., Quaife, T., Braswell, B. H., Bergeson, A., Anthony, J. A., & Monson, R. K. (2014). Joint data assimilation of satellite reflectance and net ecosystem exchange data constrains ecosystem carbon fluxes at a high-elevation subalpine forest. *Agricultural and Forest Meteorology*, *195*, 73–88.

Zobler, L. (1986). A world soil file global climate modelling. *NASA Technical Memorandum*, *32*.