Co-variations between plant functional traits emerge from constraining parameterization of a terrestrial biosphere model.

Running title: Functional traits variability inferred from data assimilation

Marc Peaucelle[1,2], Cédric Bacour[3], Philippe Ciais[1], Nicolas Vuichard[1], Sylvain Kuppel[4], Josep Peñuelas[2,5], Luca Belelli Marchesini[6,7], Peter D. Blanken[8], Nina Buchmann[9], Jiquan Chen[10], Nicolas Delpierre[11], Ankur Desai[12], Eric Dufrene[11], Damiano Gianelle[6], Cristina Gimeno-Colera[13], Carsten Gruening[14], Carole Helfter[15], Lukas Hörtnagl[9], Andreas Ibrom[16], Richard Joffre[17], Tomomichi Kato[18,19], Thomas E. Kolb[20], Beverly Law[21], Anders Lindroth[22], Ivan Mammarella[23], Lutz Merbold[24], Stefano Minerbi[25], Leonardo Montagnani[25,26], Ladislav Šigut[27], Mark Sutton[15], Andrej Varlagin[28], Timo Vesala[29,30], Georg Wohlfahrt[31], Sebastian Wolf[32], Dan Yakir[33] and Nicolas Viovy[1].

1Laboratoire des Sciences du Climat et de l'Environnement, CEA/CNRS/UVSQ, Gif-sur-Yvette, France
2CREAF, Cerdanyola del Vallès, Barcelona 08193, Catalonia, Spain
3NOVELTIS, Labège, France
4Northern Rivers Institute, School of Geosciences, University of Aberdeen
5CSIC, Global Ecology Unit CREAF -CSIC-UAB, Bellaterra, Barcelona 11 08193, Catalonia, Spain
6Department of Sustainable Agro-Ecosystems and Bioresources, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy
7Department of Landscape Design and Sustainable Ecosystems, Agrarian-Technological Institute, RUDN University, Moscow, Russia
8Department of Geography, University of Colorado, Boulder, Colorado
9Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland
10LEES Lab, Department of Geography and Spatial Sciences, Michigan State University, East Lansing, Michigan
11Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, Orsay, France
12Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison, Madison, Wisconsin
13Fundación CEAM, Parque Tecnológico, Paterna, Spain
14European Commission, Joint Research Centre, Ispra, Italy
15Centre for Ecology and Hydrology, Penicuik, UK
16Department of Environmental Engineering, Technical University of Denmark, Lyngby, Denmark
17CEFE, CNRS – Université de Montpellier – Université Paul-Valéry Montpellier – EPHE – IRD, Montpellier, France
18Research Faculty of Agriculture, Hokkaido University, Sapporo, Japan
19Global Station for Food, Land and Water Resources (GSF), GI-CoRE, Hokkaido University, Sapporo, Japan
20School of Forestry, Northern Arizona University, Flagstaff, Arizona
21Forest Ecosystems and Society Dept, College of Forestry, Oregon State University, Corvallis, Oregon
22Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
23Institute for Atmosphere and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland
24Mazingira Centre, International Livestock Research Institute (ILRI), Nairobi, Kenya
25Forest Services, Autonomous Province of Bolzano, Bolzano, Italy
26Faculty of Science and Technology, Free University of Bolzano, Piazza Università, Bolzano, Italy
27Department of Matter and Energy Fluxes, Global Change Research Institute CAS, Brno, Czech Republic

28A.N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia
29Institute for Atmosphere and Earth System Research/Forest Sciences, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland
30Viikki Plant Science, University of Helsinki, Helsinki, Finland
31Institute for Ecology, University of Innsbruck, Innsbruck, Austria
32Institute of Terrestrial Ecosystems, ETH Zurich, Zurich, Switzerland
33Earth and Planetary Sciences, Weizmann Institute of Science, Rehovot, Israel

Marc Peaucelle: marc.peaucelle@lsce.ipsl.fr (corresponding author)

Acknowledgments

This work was performed using HPC resources from GENCI-TGCC (Grant 2017-A0030106328). The authors would like to acknowledge the financial support from the European Research Council Synergy grant ERC-SyG-2013-610028 IMBALANCE-P. The study was supported by the National Centre of Excellence (272041), ICOS-Finland (281255) and Academy professor project (284701) funded by the Academy of Finland. N.B. acknowledges various funding sources for the Swiss FluxNet, particularly from the SNF (grants: 20FI21_148992, 20FI20_173691). L.M. was supported by the Swiss National Science Foundation under the 40FA40_154245/1 grant agreement. This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux and AsiaFlux offices. L.B.M. acknowledges the support of the RUDN University program 5-100. L L.Š. was supported by the Ministry of Education, Youth and Sports of CR within the National Sustainability Program I (NPU I), grant number LO1415.

Co-variations between plant functional traits emerge from constraining parameterization of a terrestrial biosphere model.

Running title: Functional traits variability inferred from data assimilation

Key words: Plant functional traits, ORCHIDEE, terrestrial model, optimization, data assimilation, plant acclimation.

Abstract

Aim

Mechanisms of plant trait adaptation and acclimation are still poorly understood and consequently lack a consistent representation in terrestrial biosphere models (TBMs). Despite the increasing availability of geo-referenced trait observations, current databases are still insufficient to cover all vegetation types and environmental conditions. In parallel, the growing number of continuous eddy-covariance observations of energy and CO_2 fluxes has enabled modelers to optimize TBMs with these data. Past attempts to optimize TBMs parameters mostly focused on model performance overlooking the ecological properties of ecosystems. The aim of this study is to assess the ecological consistency of optimized trait-related parameters while improving the model performances for gross primary productivity (GPP) at sites.

Location

World

Time period

1992-2012

Major taxa studied

Trees and C3 grasses

Methods

We optimized parameters of the ORCHIDEE model against 371 site-years of GPP estimates from the FLUXNET network and we looked at global co-variation among parameters and with climate.

Results

The optimized parameter values are shown to be consistent with leaf-scale traits, in particular well-known trade-offs observed at the leaf level, echoing the leaf economic spectrum theory. Results show a marked sensitivity of trait-related parameters to local bio-climatic variables and reproduce observed relationships between traits and climate.

Main conclusions

Our approach validates some biological processes implemented in the model and enables us to study ecological properties of vegetation at the canopy level, as well as some traits that are difficult to observe experimentally. This study stresses the need for 1) implementing explicit trade-offs and acclimation processes in TBMs, 2) improving the representation of processes to avoid model-specific parameterization as well as 3) performing systematic traits measurements at FLUXNET sites in order to gather information on plant ecophysiology and plant diversity, together with micro-meteorological conditions.

Introduction

Terrestrial biosphere models (TBMs) describe the different processes controlling exchanges of energy and trace gases between the atmosphere and the biosphere. Process-based TBMs are useful tools for understanding the dynamics of ecosystems under changing environment, for present-day to future conditions.

In most TBMs, the worldwide vegetation is divided into plant functional types (PFTs) based on general characteristics of the photosynthetic pathways, phenology, structure and physiology. Different PFTs usually share the same equations but use different parameter values to describe generic processes (photosynthesis, respiration), while biome-specific equations may be used for phenology and allocation. Therefore, for a given PFT, only the differences in climate and soil properties can determine spatial and temporal gradients in ecosystem state variables.

The prescribed values of PFT-specific parameters are derived from discrete observations obtained at varying spatial scales (organs, individuals, ecosystems; Reich *et al.*, 2007; Kattge *et al.*, 2009) and in specific environmental conditions, despite the modulation of real world plant traits by climate (Wright *et al.*, 2005; Ordoñez *et al.*, 2009; van Ommen Kloeke *et al.*, 2012; Maire *et al.*, 2015) and soil properties (Fisher *et al.*, 2012). In addition, some TBM parameters relate to traits that are difficult to measure experimentally (e.g. root turnovers or carbon allocation), or are model-specific. These parameters can hardly be directly optimized from observations and their adjustment to varying environmental conditions can only be determined by labor intensive multi-factorial ecosystem manipulation experiments (Luo *et al.*, 2017). This rigid determination of parameter values, combined with the use of single PFT to cover a range of different species (Peaucelle *et al.*, 2016), hinders a realistic representation of the past, present and future ecosystem dynamics both at the local or regional scale, and their

response to global drivers such as climate, elevated CO_2 and nutrient fertilization (Hartig *et al.*, 2012, Atkin *et al.*, 2015; Kroner & Way, 2016; Reich *et al.*, 2016).

To overcome the rigidity of the PFT representation, various approaches have been proposed to provide continuous distributions of plant functional traits related to model parameters. These approaches range from extrapolating trait observations across spatial gradients (Verheijen *et al.*, 2015), to estimating optimal trait values according to ecological theories and plant-centered approaches (Reu *et al.*, 2011; Pavlick *et al.*, 2013; Prentice *et al.*, 2014). The drawback of these different approaches is that they require both spatial and temporal observations for model calibration and/or validation. Despite the increasing number of geo-referenced trait observations (Kattge *et al.*, 2011), current databases are insufficient to cover all vegetation types and environmental conditions for projections at the ecosystem level (Musavi *et al.*, 2015, 2016). Moreover, trait observations should be co-located with process and meteorology data to understand linkages between traits and ecosystem function (Law *et al.*, 2008), which is rare in existing databases although increasingly addressed for some biomes (Bjorkman *et al.*, 2018). Long-term monitoring of functional traits is needed to assess the adjustments to climate. As such information is still lacking, approaches have been developed that confound the spatial and temporal dimensions of trait variability.

Another modeling strategy consists in optimizing TBMs against observed variables sensitive to ecosystem-level parameters in order to overcome these limitations. This approach assumes that the model structure is unbiased, so that realistic parameters values can be estimated when model simulations best match observations. Because biometric variables are sparse and often depend on processes not represented in models (Thum *et al.*, 2017), energy and trace gas flux measurements are more appealing to optimize TBM parameters. Eddy-covariance data provide near-continuous observations of CO_2, latent heat and sensible heat fluxes, and are therefore well suited for better constraining photosynthesis, respiration, transpiration and

carbon phenology model parameters. Eddy-covariance measurements have been extensively used to improve specific performances of TBMs, i.e. their ability to reproduce specific observed ecosystem behaviors (Carvalhais *et al.*, 2010; Kuppel *et al.*, 2012; Santaren *et al.*, 2014; Schürmann *et al.*, 2016). However, such model calibrations are disconnected, by construction, from ecological theory or trait-based relationships, and do not exploit the full potential of continuous flux observations across the globe, which provide both spatial and temporal information.

In this study, we aim at assessing the consistency of model trait parameters optimized against eddy-covariance flux tower measurements of growth primary productivity (GPP) using the state-of-the-art ORCHIDEE land surface model (Krinner *et al.*, 2005). In addition to classical optimization analyses (i.e. looking for the optimal parameter sets resulting in the highest model improvement), we focus here on the variability of optimized parameter values and on inter-traits correlations or trait-environment correlations. By doing so, we address the following research questions: 1) Are the parameters retrieved by optimizing the model against flux tower records consistent with known relationships between traits (i.e. trade-offs) ? or 2) between traits and environmental variables? and 3) What new relationship can be identified with this approach?

Methods

The ORCHIDEE model

The land surface model ORCHIDEE (v1.9.6, without nitrogen cycle) computes biosphere-atmosphere exchanges, consistently with water and carbon storage using ordinary differential equations (Krinner *et al.*, 2005) (Figure1). Given meteorological forcing, plant and soil conditions, the model simulates photosynthesis, all components of the surface energy budget and hydrological processes at a half-hourly time step, while the dynamics of carbon storage are calculated daily. In ORCHIDEE, the land surface is discretized into 12 plant functional types (PFT) and bare soil (Table S1.1, Appendix S1). All PFTs share the same equations, but use different parameter values, except for phenology (budburst/senescence), which is PFT-specific (Botta *et al.*, 2000).

Eddy-covariance GPP

We used half-hourly flux observations from eddy-covariance sites within the FLUXNET network (https://fluxnet.fluxdata.org). The sites were selected on the basis of spatial homogeneity and the dominance of a vegetation type that could easily be matched to one of the PFTs in ORCHIDEE, excluding crops and C_4 grasses. The vegetation type information at each site was obtained from http://fluxnet.ornl.gov. The list of analyzed FLUXNET sites (98 sites, 371 site-year) and the corresponding PFTs is given in Appendix S2. The following analyses rely on GPP derived from net ecosystem exchange (NEE; reference with variable USTAR threshold) after accounting for ecosystem respiration calculated using the method of Reichstein *et al.*, (2005) provided in the FLUXNET dataset. Years with less than 80% of available half hourly observations were discarded.

212 *Meteorological data*

213 Because ORCHIDEE needs continuous half-hourly meteorological forcing, we gap-filled time
214 series of weather variables using the interpolation algorithm developed by Vuichard & Papale
215 (2015). Linear interpolation was applied between available observations when the gap-
216 duration in the meteorological data was less than six hours. Otherwise, the variables were
217 interpolated and bias corrected using the ERA-interim reanalysis (~80km, Dee *et al.*, 2011).
218 Snow and rain were identified according to air temperature (threshold for snow being 0°C).

219

220 *Data assimilation procedure*

221 The parameters of ORCHIDEE were optimized with the ORCHIDAS package developed by:
222 Kuppel *et al.*, (2012); Bacour *et al.*, (2015); MacBean *et al.*, (2015) and Peylin *et al.*, (2016);
223 (https://orchidas.lsce.ipsl.fr/; Figure 1). Gaussian distributions of parameter and observation
224 errors being assumed, a gradient-based approach was used to minimize the Bayesian cost
225 function *J* (Tarantola, 2005):

$$\mathrm{J(x)} = \frac{1}{2}\left[\left(\mathbf{y} - H(\mathrm{x})\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y} - H(\mathrm{x})\right) + (\mathbf{x} - \mathbf{x_b})^{\mathrm{T}}\mathbf{B^{-1}}(\mathbf{x} - \mathbf{x_b})\right] \quad (1)$$

226 This function quantifies the difference between observations (y) and simulations (H(x)) (here
227 GPP), and between *a priori* (x_b) and optimized parameters (x). The *B* and *R* matrices are the
228 prior error covariance matrices for parameters and observations, respectively (including in the
229 latter case eddy-covariance measurement and model errors).

230 Both *R* and *B* were taken as diagonal, as discussed in Kuppel *et al.* (2012). The *J(x)* function
231 was iteratively minimized with the L-BFGS-B algorithm (Byrd *et al.*, 1995), which notably
232 allows bounding the range of variation of the parameters to optimize. After model calibration
233 (i.e. minimizing *J*), the posterior error covariance matrix (*A*), providing the full statistical
234 distribution of the optimized parameters was estimated by:

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

where H is the Jacobian of model at the minimum of J (Tarantola, 2005). The covariances of errors between parameters contained in the non-diagonal terms of A inform about the ability of observations given the structure of H to solve for parameters individually, or in combination. High error covariance between two parameters relates to the equifinality problem, whereby different values of these parameters result in model outputs equally matching the observations (relative to R).

Optimized parameters

We restricted our exercise to the parameters involved in the assimilation of CO_2 following previous sensitivity analyses from Kuppel (2012). We analyzed 14 parameters controlling long-term and inter-annual GPP variability (Table 1). The key equations involving each optimized parameter as well as their effect on the simulated GPP are described in Table S1.2 (Appendix S1). The parameters were related to photosynthetic capacity, phenology, carbon allocation and the water budget. Photosynthetic capacity parameters were the maximal rate of carboxylation limited by CO_2 (*Vcmax*), the ratio between the maximal rate of carboxylation limited by light and *Vcmax* (*Vj/Vc*), the optimal temperature of photosynthesis (*Topt*) and the slope of the Ball-Berry model for stomatal conductance (*gslope*). Parameters driving phenology were the specific leaf area (*SLA*), leaf longevity (*Lage*), summer maximal leaf area index (*LAImax*) and the temperature for leaf senescence (*Csenes*). Allocation parameters were the minimal fraction of *LAImax* for the use of carbohydrate reserves (*Klai*) and the period after budburst during which the use of carbohydrates is allowed (*tauleaf*) for the formation of new leaves. Finally, two parameters involved in the water status of the plant were the exponential factor describing the root profile and length (*Kroot*) and the minimal threshold at which photosynthesis becomes limited by minimum water potential (*Wlim*). In addition, two

scaling factor *Kbm* (initial biomass of leaves for evergreen species) and *bbdate* (spring burdburst date) were added in the optimization to allow adjusting the seasonal timing of GPP.

The range in variation of the three parameters corresponding to observable traits (*SLA*, *Vcmax* and *Lage)* was set from the TRY database for each PFT (Niinemets *et al.*, 1999; Deng *et al.*, 2004; Meir *et al.*, 2007; Kattge *et al.*, 2009, 2011; Domingues *et al.*, 2010; Cernusak *et al.*, 2011; Azevedo & Marenco, 2012; van de Weg *et al.*, 2012; Nascimento & Marenco, 2013). Species from the TRY database were assigned to corresponding PFTs based on available metadata about plant structure, leaf phenology and climate information extracted from species' latitude and longitude coordinates. We chose as a reference range the 2.5 - 97.5 percentile of the trait distributions from TRY. The variation ranges for the other parameters were fixed based on expert judgment (Kuppel *et al.*, 2014).

Simulations and assimilation set-up

At each flux tower site, we assumed that the eddy-covariance flux footprint was entirely composed by a single PFT (Appendix S2). The model was forced by local meteorological observations (see *Meteorological data* section) and soil texture from the harmonized worldwide soil database (Nachtergaele *et al.*, 2012) to define the residual and saturation water contents, and the saturated hydraulic conductivity in the soil model (Ducoudré *et al.*, 1993; Krinner *et al.*, 2005) based on Van Genuchten (1980). Initial soil carbon pools in equilibrium with local climate were obtained with an analytical spin-up procedure (Lardy *et al.*, 2011; Xia *et al.*, 2012). Initial biomass was simulated until reaching equilibrium (generally after a ~300 year-long simulations using the studied year meteorological data and constant CO_2 set to level of the year), thus different from the real stand age observed at each site.

We optimized GPP averaged over 15 days using moving windows to avoid noise from high frequency variations in the parameter optimization that could induce convergence issues (Bacour *et al.*, 2015). As far as test data from eddy-covariance measurements are concerned, high frequency variations in fluxes include also variation in the boundary layer that are unrelated to the fluxes at the surface (Ibrom *et al.*, 2006). Santaren *et al.* (2007) estimated that for parameters related to photosynthesis and phenology, optimization based on half-hourly observations did not improve the results. For each site, the optimizations were conducted year-by-year to account for trait variability over time (Wu *et al.*, 2013).

Following MacBean *et al.* (2015), each calibration (site-year) used ten replicates representing different starting parameter sets with values randomly picked within their allowed variation range (Table S1.3). Only the best calibration out of these ten replicates was retained for analyses. This procedure increases the chances of finding the global minimum of *J* as Santaren *et al.* (2014) showed that the gradient-based algorithm was sensitive to initial conditions with a non-linear and complex model such as ORCHIDEE.

Analyses

We only retained calibrations for which the optimized model reproduced GPP observations with high precision. The rationale for this was that optimized parameters from model runs which agreed poorly with GPP observations provided little or no useable information. The filtering was performed using a two-step procedure.

First, the criterion for 'improved GPP simulation' was the relative site-year posterior RMSE ($RMSE_{re}$) between observed and optimized GPP:

$$RMSE_{re} = \frac{RMSE}{mean(GPP_{obs})} \quad (4)$$

Whenever the value of $RMSE_{re}$ was higher than the all-$RMSE_{re}$ median plus one interquartile range (IQR), the site-year was removed from the analysis. We also discarded sites with 'inconsistent parameters values', i.e. with too large differences between the ten replicates at the same site reflecting convergence issues (equifinality) of the algorithm.

Secondly, for sites with at least two $RMSE_{re}$ below 10 % among the ten replicates, we estimated the coefficient of variation (CV) of parameters across the replicates. We retained only years for which the median CV was below the median of all CV plus one IQR of their distribution. This filtering provided optimized parameters from 371 site-years (over 516 initially considered) for 98 sites (over 116; Appendix S2) spanning seven PFTs located in boreal, temperate and tropical areas (Table S3.4; Appendix S3).

For each parameter, we calculated the uncertainty reduction (*UR*) as:

$$UR = 1 - \frac{\sigma_{post}}{\sigma_{prior}} \quad (5)$$

With σ_{post} and σ_{prior} being the *posterior* and *prior* parameter uncertainties (square root of the diagonal elements of *A* and *B*). We then separated in the analysis the well- from the poorly-constrained parameters. Well-constrained parameters are defined as those with 1) *UR* higher than the median of *UR* distributions for all parameters and 2) a low correlation of error with other parameters (from the A matrix, Eq. 2). Note that a strong error correlation making two parameters poorly constrained individually is still an interesting result as it indicates a range of possible tradeoffs between these two parameters.

The optimized parameter values were regressed against the local background bio-climatic variables (Table 2) for each site, and against the soil relative water content (volume of water by volume of soil) simulated by ORCHIDEE. Bio-climatic variables were averaged over the whole year and over the length of the growing season (GSL). For temperate sites, the growing

328 season was defined as the period with daily temperature above 5°C and relative soil water
329 content above 0.2 (Violle *et al.*, 2015). In some tropical regions, the growing season length is
330 potentially limited by water availability (wet/dry seasons), we thus kept the same definition as
331 for temperate ecosystems. For boreal sites, we adapted the definition of the growing season
332 such as weekly temperature must be above 0°C. Analyses were performed with the R.3.2
333 software (R Core Team, 2016) and standardized major axis (SMA) analyses were performed
334 with the 'lmodel2' package (Legendre, 2014). Because we sought to compare simulated
335 correlations with common ecological properties observed at the global scale, we analyzed
336 different groups of PFTs: all PFTs together; deciduous versus evergreens; needleleaves versus
337 broadleaves; and C_3 grasses (Table S1.1). Regressions were performed both with and without
338 a logarithmic transformation of the data.

Results and comparison to existing literature

Optimization performances

A full description of the optimization performances and parameter uncertainty reduction can be found in Appendix S3.

In all cases, the optimized GPP time series better agrees with observations than the prior ones, with the RMSE being reduced by 76.6 ± 13.0 % (Table S3.4; Appendix S3). The median posterior $RMSE_{re}$ is 0.19 and the IQR is 0.11. The median CV over all parameters is 0.24 (IQR=0.13). After optimization, the parameter uncertainty (Eq. 5) is reduced by 30 % on average (Table S3.5; Appendix S3).

The posterior error correlation matrix *A* (Eq.2) reveals a positive correlation between *Vcmax* and several other parameters including (Figure 2): *Topt* (r=0.57±0.05); *gslope* (r=-0.37±0.04); *Kroot* (r=0.24±0.07) and *Vj/Vc* (r=-0.31±0.04). There also exists a negative correlation between *Kroot* and *gslope* (r=-0.38±0.08), between *Kroot* and *Wlim* (r=-0.30±0.09) and between *LAImax* and *Klai* (r=-0.37±0.16) (Figure 2).

Jointly analyzing information from the uncertainty reduction (Appendix S3) and the cross-parameter error correlation enables to distinguish between: 1) well constrained parameters (*Lage* and *SLA* for evergreens/ *Lage* and *Csenes* for deciduous); 2) well constrained parameters with a risk of equifinality (*gslope, Kroot, LAImax, Topt* and *Vcmax*); and 3) poorly constrained parameters (*Vj/Vc*, *Klai*, *Tauleaf* and *Wlim;* Table 1). In the following analyses, trait co-variations have to be interpreted in respect to confidence intervals (posterior error) in parameter estimates.

Co-variation between parameters

We analyzed cross-site correlations between optimized parameters in relation to expected trait relationships. The co-variation between all parameters is illustrated in Figure S4.2 (Appendix

S4). For more clarity and considering the large number of parameters, we only describe here the relationships involving four parameters related to phenology (*SLA*, *Lage*) and photosynthesis (*Vcmax*, *gslope*). All relationships are provided in Table S4.6 (Appendix S4).

We observed a negative correlation between *SLA* and *Lage* for all PFTs (r=-0.63; Table 3) as well as for evergreens (r=-0.67) and broadleaves PFTs (r=-0.53), separately. The slope of the emerging relationship between *LMA* (*1/SLA*) and *Lage* (1.91; 1.63-2.24 95% confidence interval; $p < 0.05$) for all PFTs was close to the observed slope from field observations (1.71; 1.62-1.82; Wright *et al.* 2004). Results highlighted other co-variations between *Lage* and *Vcmax* (r=-0.59 overall PFTs), *gslope* and *Lage* (r=-0.7 for broadleaves), *LAImax* and *SLA* (r=0.6 for needleleaves), and between *SLA* and *Vcmax* (r=-0.55 for evergreens). Here again, the slope between *Lage* and *Vcmax* emerging for broadleaves PFTs (-1.69) was close to observations (-1.13; Xu *et al.* 2017).

No relationships were reported between *gslope* and *Lage* or between *glsope* and *SLA*, but a trade-off between the stomatal conductance (*gs*) and *Lage* was observed experimentally (Reich *et al.*, 1992; Poorter & Bongers, 2006), as well as a positive correlation between *gs* and *SLA* (Poorter & Bongers, 2006). The optimizations showed opposite relationships between *gslope* and *SLA* depending on the PFT: a positive significant correlation was obtained for deciduous PFTs and a negative significant correlation for evergreens and grasses (Table 3).

The positive relationship between *SLA* and *LAImax* emerging from optimized parameters for coniferous PFTs was consistent with the positive correlation between *LAI* and *SLA* reported by Pierce *et al.* (1994) for coniferous forests. Finally, a negative correlation between *SLA* and *Vcmax* has been observed experimentally for two gymnosperms species (Niinemets *et al.*, 2007), confirming the negative relationships found in our study for needleleaves. Despite the

387 equifinality risk between *gslope* and the soil water stress *Wlim* in Figure 2, the positive
388 correlation observed for broadleaves (r=0.7) and evergreens (r=0.52) was comparable to
389 observations from independent data compiled by Lin *et al.* (2015).

390 Other significant correlations from the optimized parameters (Table S4.6, Figure S4.2;
391 Appendix S4) could not be verified against observations because of the correlation of errors
392 observed in Figure 2 or because of the scarcity of ecological data preventing us to conclude
393 about the true nature of those correlations, as for example between *gslope* and *Vcmax.*

394

395 *Variation of trait-related parameters with climate*

396 We analyzed correlations between parameters and climate variables (Table 4, Figure S5.4;
397 Appendix S5). As for co-variations between parameters, we only described here those
398 implying *SLA, Lage, Vcmax* and *gslope*. All relationships are listed in Table S5.7 and more
399 detailed analysis are available in Appendix S5.

400 We found a strong negative correlation between leaf lifespan (*Lage*) and temperatures (*MAT,*
401 *TMIN;* r=-0.78/-0.65; Figure 3a) for evergreen PFTs. This correlation was independently
402 reported at global scale (Wright *et al.,* 2005; van Ommen Kloeke *et al.,* 2012) and confirmed
403 by Reich *et al.* (2014) who showed higher needle longevity with cold temperatures for boreal
404 species. However, the observed positive correlation between *Lage* and *MAT* at the global
405 scale for deciduous PFTs (Wright *et al.*, 2005; van Ommen Kloeke *et al.*, 2012) was not
406 found specifically for deciduous systems in our study. Nevertheless, a positive correlation was
407 observed for C_3 grasses and broadleaves (including deciduous). We also found a strong
408 negative correlation between *Lage* and the mean annual precipitations (*MAP)* for evergreens
409 PFTs (r = -0.65), consistent with field data (van Ommen Kloeke *et al.*, 2012). In addition, a

410 negative correlation between *Lage* and incident shortwave radiation (*SW*) for evergreens was
411 obtained, consistent with field observations (Poorter & Bongers, 2006).
412
413 Regarding *SLA,* we found opposite sensitivities to *MAT* for evergreen (r=0.65) and deciduous
414 forests (r=-0.55). This result is consistent with independent leaf-scale data showing a positive
415 correlation between *SLA* and *MAT* for evergreen species (Figure 3b) and a negative
416 correlation for deciduous ones (Wright *et al.* 2005). The model calibration also resulted in a
417 positive correlation between the relative precipitation (*RELP*; Table 2) and *SLA* for deciduous
418 trees (r = 0.60; Figure 3c). Regarding the positive correlations obtained between *SLA* with
419 *Kroot* or *gslope* (Table 3), it suggests that *SLA* is highly sensitive to water stress for deciduous
420 trees. For evergreens, positive correlation between *SLA* and precipitation also emerges when
421 considering the length of the growing season (*MAPgs,* r = 0.57; Table 4); which is consistent
422 with trait data (Wright *et al.*, 2005). For evergreens, *SLA* was positively correlated to *SW*
423 (r=0.53), a relationship observed by Givnish *et al.* (2004) and Poorter & Bongers (2006).
424
425 In their meta-analysis of stomatal conductance parameters from observations of several PFTs,
426 Lin *et al.* (2015) showed that the slope of the stomatal conductance is positively correlated to
427 the mean air temperature over the growing period and to soil moisture stress. Here, our results
428 show the same correlation between *gslope* and soil moisture during the growing season
429 (r=0.71; Figure 3d) and relative precipitation (r=0.66) for deciduous or broadleaved PFTs. On
430 the contrary, we find that *gslope* is negatively correlated with mean annual precipitation for
431 C_3 grasses (r=-0.59), and with shortwave radiation for broadleaved PFTs (r=-0.63). Medlyn *et*
432 *al.* (2011) suggested that *gslope* is proportional to the photosynthesis compensation point for
433 CO_2, and consequently to growth temperatures of the plant (Bernacchi *et al.*, 2001). This

assumption is supported by the data from Lin *et al.* (2015). In our study, the relationship between *gslope* and temperature was not supported.

Finally, *Vcmax* is mostly sensitive to temperature and light for broadleaved PFTs, with a negative correlation observed with *MAT* (r=-0.52) and *SW* (r=-0.54). This result contradicts previous observations by Ali *et al.* (2015), who suggested a positive correlation between *Vcmax* and seasonal temperature and light variations.

Discussion

Uncertainties and shortcomings of the approach

This section provides an overview of possible shortcomings of our approach that may explain some residual mismatch between model and observations. Several factors can impact the optimized value of the parameters, potentially aliasing the observed relationships: 1) flux measurements errors and errors in ecosystem respiration estimates used to derive gap-filled GPP; 2) optimization protocol/setup errors; and 3) model systematic errors deriving from absent or poorly represented processes in the model.

First, we restricted our analysis to GPP. This flux is not directly measured but estimated from NEE measured using the eddy-covariance method with an estimate of ecosystem respiration determined using empirical models (Reichstein *et al.*, 2005), and thus can be biased by several factors (see Appendix S3 for a list of these factors). We chose GPP over a combination of NEE and latent heat or evapotranspiration fluxes, which has often been used to optimize ORCHIDEE (Kuppel *et al.*, 2012; Bacour *et al.*, 2015; Peylin *et al.*, 2016), because it implies the optimization of more parameters related to soil, respiration and energy budget, and therefore increases the risk of equifinality. To reduce the uncertainties, it is

458 necessary to lower the correlation of errors between parameters by assimilating
459 complementary biophysical variables. For example, assimilating both GPP and LAI estimates
460 at the site level could improve the evaluation of parameters such as *SLA* or *Lage*, and
461 consequently improve the estimation of photosynthesis parameters.

462 Second, the Bayesian framework is based on the assumption that the model/observation errors
463 are random and that the model structure is "true". Any bias of model structure is expected to
464 be aliased onto the estimated parameters (MacBean *et al.*, 2016) and might therefore impact
465 the retrieved correlations. For instance, missing processes would be compensated during the
466 optimization by adjusting parameters (e.g. light attenuation, vertical distribution of LAI,
467 diffuse light, horizontal light distribution in the stand) to non optimal values. Also, while
468 traits are usually measured at the leaf level, our approach rather focuses on traits at the canopy
469 level (given the structure of ORCHIDEE and the assumed exponential attenuation of light and
470 LAI from top to bottom of canopy (Krinner *et al.* 2005; Table S1.2), and the assimilation of
471 GPP data). As an additional test, we conducted the above analyses using multi-year instead of
472 single-year observations in order to add more constraints on parameters (see Figure S4.3 &
473 S5.5). The same relationships were found as with single-year observations, thus strengthening
474 our conclusions, showing that spatial correlations are observed even when taking into account
475 a possible temporal variability of traits.

476 Finally, a wrong representation of species and the lack of representation of representation of
477 traits variability within a community in ORCHIDEE can affect simulated processes, which
478 will ultimately impact the estimated parameter values (see Appendix S3 for a discussion on
479 initial site conditions). Especially, the C_3 grass PFT represents diverse grasslands, with
480 different species, ecophysiology (Adams *et al.*, 2016) and management practices (Merbold *et*
481 *al.*, 2014). This results in an increased variability and a high range of estimated plant
482 functional traits (Figure S3.1). A refinement of the PFT definition may improve the

483 robustness of optimizations (for instance by separating natural or semi-managed biomes), or
484 by distinguishing genera or major species (Peaucelle *et al.*, 2016).

485 In order to decrease the impact of uncertainty in PFT composition and reduce the correlation
486 errors between parameters, the use of concomitant observations of traits and carbon fluxes at
487 the FLUXNET sites would enable a) to constrain known parameters and b) to validate
488 optimized traits. However, functional trait observations at FLUXNET sites as well as a
489 precise description of species composition are not yet systematic (Musavi *et al.*, 2015, 2016).

490

491 *Ecological consistency of trait relationships*

492 The optimization of model parameters managed to reproduce many known ecological
493 properties. The optimized parameters consistently matched the well-known relationships
494 resulting from the leaf economic spectrum theory (LES, Reich *et al.*, 1999; Wright *et al.*,
495 2004). Particularly our results align with the trait theory that long lived canopies are
496 metabolically less active and are consistent with the LES empirical evidence that plants invest
497 either in structure or photosynthesis (Liu *et al.*, 2010; Reich, 2014).

498 Our results also reproduced several observed trait-climate relationships at the PFT level.
499 Globally, evergreen PFT parameters showed a strong dependency on mean annual
500 temperature and radiation, while parameters for deciduous PFTs exhibited a strong sensitivity
501 to precipitation and soil moisture over the growing season (Figure S5.4). As postulated by
502 Reich (2014), climate exerts a control on the average leaf characteristics at the community
503 level. The observed relationships obtained at the PFT level might reflect, not only differences
504 in plant response to climate, but also differences in plant community composition (Shi *et al.*,
505 2015). These results suggest that both the development of acclimation processes and trait-

based approaches are needed in TBMs if we seek to capture the effect of biogeography on ecosystem characteristics (Lu *et al.*, 2017; Fisher *et al.*, 2018).

Finally, while the results clearly highlight that photosynthesis and phenological mechanisms implemented in ORCHIDEE are robust enough to reproduce known behaviors of several vegetation species, belowground processes still appear poorly represented, which resulted in weakly constrained parameters and trait co-variations inconsistent with literature. These discrepancies are primarily due to a lack in eco-physiological knowledge reflecting the actual difficulty to study belowground ecological processes. The rooting system uses model-specific parameters (*Kroot*) that are hardly comparable to measured functional traits.

Concluding remarks and recommendations

The approach presented in this study is a new and effective way to validate the processes implemented in TBMs, to better define vegetation response to climate (Liang *et al.*, others, 2018), and could help improving existing data assimilation frameworks (Kaminski *et al.*, 2013; LeBauer *et al.*, 2013; Arsenault *et al.*, 2018) by bringing ecological constraints. The availability of continuous observations from eddy-covariance flux measurements gives a unique opportunity to resolve the different components of the short and long-term variability of traits through this approach.

Our results show that optimized leaf-related parameters align with plant trait theory, and highlight the need to implement acclimation processes and trait-based approaches in models instead of using constant parameters to reduce uncertainties in spatio-temporal patterns of the modeled carbon fluxes. A first step would be to assess the behavior of the model at the global scale when trait-climate relationships characterized in this study are explicitly implemented.

In parallel, relationships highlighted in this study may help to develop or validate new methods to simulate plant acclimation. Used in a prognostic way, this approach could enable to study correlations at the canopy scale and to assess the behavior of trait-related parameters that are difficult to observe experimentally.

Several known ecological properties, observed at the site/leaf scale, emerged from model-data assimilation. However, quantitative comparisons with observations were possible only for two of them, *SLA* and *Lage*, which are also the two most studied traits. This is mainly because TBMs use model-specific parameters that cannot be directly compared to standard trait observations, but also because concomitant observations of functional traits, both in space and time, are scarce in the literature. A recommendation to the TBM community would be to make use of parameters (and processes) that can be related directly to observations in order to unit vegetation model and functional traits (for instance the use of the Specific Root Length for belowground processes).

We argue that co-located systematic and standardized trait observations (starting with key traits related to phenology -*SLA, LAI*-, photosynthesis -*Vcmax, Jmax, Topt*-, water transport -*gs*- and allocation -*Carbon:Nitrogen ratio, shoot/root*-; (Law *et al.*, 2008) along with biometric data are needed at the FLUXNET sites or within other environmental observation networks such as ICOS (Integrated Carbon Observation System) or NEON (National Ecological Observatory Network) if we seek to distinguish temporal and spatial components of trait variability across biomes and climates. The creation of a FLUXNET trait database could improve our comprehension of trait acclimation and help us to disentangle the differences observed at regional and local scales, to improve the up-scaling of processes from the leaf to the canopy/ecosystem level and to properly calibrate/validate ecosystem models.

Supporting information

Appendix S1: Description of PFTs, model parameters and equations

Table S1.1: List of plant functional types.

Table S1.2: List of main equations involving optimized parameters.

Table S1.3: Default parameter value and range allowed by the optimization algorithm.

Appendix S2: List of FLUXNET sites used for the analyses (xlsx file).

Appendix S3: Optimization performances

Figure S3.1: Distribution of optimized parameter values.

Table S3.4: Mean *a priori* and *a posteriori* RMSE between observations and simulations.

Table S3.5: Mean parameter uncertainty reduction between *prior* and *posterior* simulations.

Appendix S4: Relationships between traits

Figure S4.2: Correlation matrices between traits optimized against site-year GPP.

Figure S4.3: Correlation matrices between traits optimized against site GPP.

Table S4.6: Extended Table 3 with all relationships

Appendix S5: Relationships between traits and climate

Figure S5.4: Correlation matrices between traits and environmental variables optimized against site-year GPP

Figure S5.5: Correlation matrices between traits and environmental variables optimized against site GPP

Table S5.7: Extended Table 4 with all relationships

Data accessibility

All FluxNet data can be downloaded at: https://fluxnet.fluxdata.org

Information about the ORCHIDEE model, source code and contact: http://orchidee.ipsl.fr/

Information about the data assimilation system ORCHIDAS: https://orchidas.lsce.ipsl.fr/

References

Adams, M.A., Turnbull, T.L., Sprent, J.I. & Buchmann, N. (2016) Legumes are different: Leaf nitrogen, photosynthesis, and water use efficiency. *Proceedings of the National Academy of Sciences*, **113**, 4098–4103.

Ali, A.A., Xu, C., Rogers, A., McDowell, N.G., Medlyn, B.E., Fisher, R.A., Wullschleger, S.D., Reich, P.B., Vrugt, J.A., Bauerle, W.L. & others (2015) Global-scale environmental control of plant photosynthetic capacity. *Ecological Applications*, **25**, 2349–2365.

Arsenault, K.R., Kumar, S.V., Geiger, J.V., Wang, S., Kemp, E., Mocko, D.M., Beaudoing, H.K., Getirana, A., Navari, M., Li, B. & others (2018) The Land surface Data Toolkit (LDT v7. 2)–a data fusion environment for land data assimilation systems. *Geoscientific Model Development*, **11**, 3605–3621.

Atkin, O.K., Bloomfield, K.J., Reich, P.B., Tjoelker, M.G., Asner, G.P., Bonal, D., Bönisch, G., Bradford, M.G., Cernusak, L.A., Cosio, E.G., Creek, D., Crous, K.Y., Domingues, T.F., Dukes, J.S., Egerton, J.J.G., Evans, J.R., Farquhar, G.D., Fyllas, N.M., Gauthier, P.P.G., Gloor, E., Gimeno, T.E., Griffin, K.L., Guerrieri, R., Heskel, M.A., Huntingford, C., Ishida, F.Y., Kattge, J., Lambers, H., Liddell, M.J., Lloyd, J., Lusk, C.H., Martin, R.E., Maksimov, A.P., Maximov, T.C., Malhi, Y., Medlyn, B.E., Meir, P., Mercado, L.M., Mirotchnick, N., Ng, D., Niinemets, Ü., O'Sullivan, O.S., Phillips, O.L., Poorter, L., Poot, P., Prentice, I.C., Salinas, N., Rowland, L.M., Ryan, M.G., Sitch, S., Slot, M., Smith, N.G., Turnbull, M.H., VanderWel, M.C., Valladares, F., Veneklaas, E.J., Weerasinghe, L.K., Wirth, C., Wright, I.J., Wythers, K.R., Xiang, J., Xiang, S. & Zaragoza-Castells, J. (2015) Global variability in leaf respiration in relation to climate, plant functional types and leaf traits. *New Phytologist*, **206**, 614–636.

Azevedo, G. & Marenco, R. (2012) Growth and physiological changes in saplings of Minquartia guianensis and Swietenia macrophylla during acclimation to full sunlight. *Photosynthetica*, **50**, 86–94.

Bacour, C., Peylin, P., MacBean, N., Rayner, P., Delage, F., Chevallier, F., Weiss, M., Demarty, J., Santaren, D., Baret, F. & others (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.

Bernacchi, C.J., Singsaas, E.L., Pimentel, C., Portis Jr, A.R. & Long, S.P. (2001) Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell & Environment*, **24**, 253–259.

Bjorkman, A.D., Myers-Smith, I.H., Elmendorf, S.C., Normand, S., Thomas, H.J., Alatalo, J.M., Alexander, H., Anadon-Rosell, A., Angers-Blondin, S., Bai, Y. & others (2018) Tundra Trait Team: A database of plant traits spanning the tundra biome. *Global Ecology and Biogeography*, **27**, 1402–1411.

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**, 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**, 1190–1208.

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

Cernusak, L.A., Hutley, L.B., Beringer, J., Holtum, J.A. & Turner, B.L. (2011) Photosynthetic physiology of eucalypts along a sub-continental rainfall gradient in northern Australia. *Agricultural and Forest Meteorology*, **151**, 1462–1470.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P. & others (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137**, 553–597.

Deng, X., Ye, W.-H., Feng, H.-L., Yang, Q.-H., Cao, H.-L., Xu, K.-Y. & Zhang, Y. (2004) Gas exchange characteristics of the invasive species Mikania micrantha and its indigenous congener M. cordata (Asteraceae) in South China. *Botanical Bulletin of Academia Sinica*, **45**.

Domingues, T.F., Meir, P., Feldpausch, T.R., Saiz, G., Veenendaal, E.M., Schrodt, F., Bird, M., Djagbletey, G., Hien, F., Compaore, H., Diallo, A., Grace, J. & Lloyd, J. (2010) Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant, Cell & Environment*, **33**, 959–980.

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

Fisher, J.B., Badgley, G. & Blyth, E. (2012) Global nutrient limitation in terrestrial vegetation. *Global Biogeochemical Cycles*, **26**, n/a–n/a.

Fisher, R.A., Koven, C.D., Anderegg, W.R., Christoffersen, B.O., Dietze, M.C., Farrior, C.E., Holm, J.A., Hurtt, G.C., Knox, R.G., Lawrence, P.J. & others (2018) Vegetation demographics in Earth System Models: A review of progress and priorities. *Global change biology*, **24**, 35–54.

Van Genuchten, M.T. (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil science society of America journal*, **44**, 892–898.

Givnish, T.J., Montgomery, R.A. & Goldstein, G. (2004) Adaptive radiation of photosynthetic physiology in the Hawaiian lobeliads: light regimes, static light responses, and whole-plant compensation points. *American Journal of Botany*, **91**, 228–246.

Ibrom, A., Jarvis, P.G., Clement, R., Morgenstern, K., Oltchev, A., Medlyn, B.E., Wang, Y.P., Wingate, L., Moncrieff, J.B. & Gravenhorst, G. (2006) A Comparative Analysis of Simulated and Observed Photosynthetic CO2 Uptake in Two Coniferous Forest Canopies. *Tree Physiology*, **26**, 845–864.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R. & others (2013) The BETHY/JSBACH carbon cycle data assimilation system: Experiences and challenges. *Journal of Geophysical Research: Biogeosciences*, **118**, 1414–1426.

Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P.B., Wright, I.J., Cornelissen, J.H.C., Violle, C., Harrison, S.P., Van BODEGOM, P.M., Reichstein, M., Enquist, B.J., Soudzilovskaia, N.A., Ackerly, D.D., Anand, M., Atkin, O., Bahn, M., Baker, T.R., Baldocchi, D., Bekker, R., Blanco, C.C., Blonder, B., Bond, W.J., Bradstock, R., Bunker, D.E., Casanoves, F., Cavender-Bares, J., Chambers, J.Q., Chapin Iii, F.S., Chave, J., Coomes, D., Cornwell, W.K., Craine, J.M., Dobrin, B.H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W.F., Fang, J., Fernández-Méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G.T., Fyllas, N.M., Gallagher, R.V., Green, W.A., Gutierrez, A.G., Hickler, T., Higgins, S.I., Hodgson, J.G., Jalili, A., Jansen, S., Joly, C.A., Kerkhoff, A.J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J.M.H., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T.D., Leishman, M., Lens, F., Lenz, T., Lewis, S.L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M.D., Manning, P., Massad, T., Medlyn, B.E., Messier, J., Moles, A.T., Müller, S.C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V.G., Onoda, Y., Ordoñez, J., Overbeck, G., Ozinga, W.A., Patiño, S., Paula, S., Pausas, J.G., Peñuelas, J., Phillips, O.L., Pillar, V., Poorter, H., Poorter, L., Poschlod, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch, S., Reu, B., Sack, L., Salgado-Negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A., Sosinski, E., Soussana, J.-F., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram, M., Weiher, E., White, M., White, S., Wright, S.J., Yguel, B., Zaehle, S., Zanne, A.E. & Wirth, C. (2011) TRY – a global database of plant traits. *Global Change Biology*, **17**, 2905–2935.

Kattge, J., Knorr, W., Raddatz, T. & Wirth, C. (2009) Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology*, **15**, 976–991.

Krinner, G., Viovy, N., Noblet-Ducoudré, N. de, Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. & Prentice, I.C. (2005) A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, **19**, 33 PP.

Kroner, Y. & Way, D.A. (2016) Carbon fluxes acclimate more strongly to elevated growth temperatures than to elevated CO2 concentrations in a northern conifer. *Global Change Biology*, **22**, 2913–2928.
Kuppel, S. (2012) *Assimilation de mesures de flux turbulents d'eau et de carbone dans un modèle de la biosphère continentale*, Versailles-St Quentin en Yvelines.
Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F. & Richardson, A. (2012) Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*, **9**, 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**, 2581–2597.
Lardy, R., Bellocchi, G. & Soussana, J.-F. (2011) A new method to determine soil organic carbon equilibrium. *Environmental Modelling & Software*, **26**, 1759–1763.
Law, B.E., Arkebauer, T., Campbell, J.L., Chen, J., Sun, O., Schwartz, M., van Ingen, C. & Verma, S. (2008) Terrestrial carbon observations: Protocols for vegetation sampling and data submission. *FAO, Rome*.
LeBauer, D.S., Wang, D., Richter, K.T., Davidson, C.C. & Dietze, M.C. (2013) Facilitating feedbacks between field measurements and ecosystem models. *Ecological Monographs*, **83**, 133–154.
Legendre, P. (2014) *lmodel2: Model II Regression. R package version 1.7-2*, Available at: ht tp://CRAN. R-project. org/package= lmodel2 (accessed 2 March 2015).
Liang, J., Xia, J., Shi, Z., Jiang, L., Ma, S., Lu, X., Mauritz, M., Natali, S.M., Pegoraro, E. & Penton, C.R., others (2018) Biotic responses buffer warming-induced soil organic carbon loss in Arctic tundra. *Global change biology*.
Lin, Y.-S., Medlyn, B.E., Duursma, R.A., Prentice, I.C., Wang, H., Baig, S., Eamus, D., de Dios, V.R., Mitchell, P., Ellsworth, D.S. & others (2015) Optimal stomatal behaviour around the world. *Nature Climate Change*, **5**, 459–464.
Liu, G., Freschet, G.T., Pan, X., Cornelissen, J.H.C., Li, Y. & Dong, M. (2010) Coordinated variation in leaf and root traits across multiple spatial scales in Chinese semi-arid and arid ecosystems. *New Phytologist*, **188**, 543–553.
Luo, Y., Jiang, L., Niu, S. & Zhou, X. (2017) Nonlinear responses of land ecosystems to variation in precipitation. *New Phytologist*, **214**, 5–7.
Lu, X., Wang, Y.-P., Wright, I.J., Reich, P.B., Shi, Z. & Dai, Y. (2017) Incorporation of plant traits in a land surface model helps explain the global biogeographical distribution of major forest functional types. *Global Ecology and Biogeography*, **26**, 304–317.
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**, 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**, 3569–3588.
Maire, V., Wright, I.J., Prentice, I.C., Batjes, N.H., Bhaskar, R., Bodegom, P.M., Cornwell, W.K., Ellsworth, D., Niinemets, Ü., Ordonez, A. & others (2015) Global effects of soil and climate on leaf photosynthetic traits and rates. *Global Ecology and Biogeography*, **24**, 706–717.
Medlyn, B.E., Duursma, R.A., Eamus, D., Ellsworth, D.S., Prentice, I.C., Barton, C.V.M., Crous, K.Y., De Angelis, P., Freeman, M. & Wingate, L. (2011) Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, **17**, 2134–2144.
Meir, P., Levy, P.E., Grace, J. & Jarvis, P.G. (2007) Photosynthetic parameters from two contrasting woody vegetation types in West Africa. *Plant Ecology*, **192**, 277–287.
Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D. & Buchmann, N. (2014) Greenhouse gas budget (CO2, CH4 and N2O) of intensively managed grassland following restoration. *Global change biology*, **20**, 1913–1928.
Musavi, T., Mahecha, M.D., Migliavacca, M., Reichstein, M., van de Weg, M.J., van Bodegom, P.M., Bahn, M., Wirth, C., Reich, P.B., Schrodt, F. & others (2015) The imprint of plants on ecosystem functioning: A data-driven approach. *International Journal of Applied Earth Observation and Geoinformation*, **43**, 119–131.

Musavi, T., Migliavacca, M., van de Weg, M.J., Kattge, J., Wohlfahrt, G., van Bodegom, P.M., Reichstein, M., Bahn, M., Carrara, A., Domingues, T.F., Gavazzi, M., Gianelle, D., Gimeno, C., Granier, A., Gruening, C., Havránková, K., Herbst, M., Hrynkiw, C., Kalhori, A., Kaminski, T., Klumpp, K., Kolari, P., Longdoz, B., Minerbi, S., Montagnani, L., Moors, E., Oechel, W.C., Reich, P.B., Rohatyn, S., Rossi, A., Rotenberg, E., Varlagin, A., Wilkinson, M., Wirth, C. & Mahecha, M.D. (2016) Potential and limitations of inferring ecosystem photosynthetic capacity from leaf functional traits. *Ecology and Evolution*, **6**, 7352–7366.

Nachtergaele, F.O., Velthuizen, H. van, Verelst, L., Wiberg, D., Batjes, N.H., Dijkshoorn, J.A., Engelen, V.W. van, Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Teixeira, E. & Shi, X. (2012) Harmonized World Soil Database (version 1.2). 42.

Nascimento, H.C. & Marenco, R.A. (2013) Mesophyll conductance variations in response to diurnal environmental factors in Myrcia paivae and Minquartia guianensis in Central Amazonia. *Photosynthetica*, **51**, 457–464.

Niinemets, Ü., Lukjanova, A., Turnbull, M.H. & Sparrow, A.D. (2007) Plasticity in mesophyll volume fraction modulates light-acclimation in needle photosynthesis in two pines. *Tree Physiology*, **27**, 1137–1151.

Niinemets, Ü., Oja, V. & Kull, O. (1999) Shape of leaf photosynthetic electron transport versus temperature response curve is not constant along canopy light gradients in temperate deciduous trees. *Plant, Cell & Environment*, **22**, 1497–1513.

Van Ommen Kloeke, A., Douma, J., Ordonez, J., Reich, P. & Van Bodegom, P. (2012) Global quantification of contrasting leaf life span strategies for deciduous and evergreen species in response to environmental conditions. *Global Ecology and Biogeography*, **21**, 224–235.

Ordoñez, J.C., Van Bodegom, P.M., Witte, J.-P.M., Wright, I.J., Reich, P.B. & Aerts, R. (2009) A global study of relationships between leaf traits, climate and soil measures of nutrient fertility. *Global Ecology and Biogeography*, **18**, 137–149.

Pavlick, R., Drewry, D.T., Bohn, K., Reu, B. & Kleidon, A. (2013) The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences*, **10**, 4137–4177.

Peaucelle, M., Bellassen, V., Ciais, P., Peñuelas, J. & Viovy, N. (2016) A new approach to optimal discretization of plant functional types in a process-based ecosystem model with forest management: a case study for temperate conifers. *Global Ecology and Biogeography*, n/a–n/a.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P.J., Kuppel, S., Koffi, E.N., Kane, A., Maignan, F., Chevallier, F., Ciais, P. & Prunet, P. (2016) A new step-wise Carbon Cycle Data Assimilation System using multiple data streams to constrain the simulated land surface carbon cycle. *Geoscientific Model Development Discussions*, **2016**, 1–52.

Pierce, L.L., Running, S.W. & Walker, J. (1994) Regional-scale relationships of leaf area index to specific leaf area and leaf nitrogen content. *Ecological Applications*, 313–321.

Poorter, L. & Bongers, F. (2006) Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology*, **87**, 1733–1743.

Prentice, I.C., Dong, N., Gleason, S.M., Maire, V. & Wright, I.J. (2014) Balancing the costs of carbon gain and water transport: testing a new theoretical framework for plant functional ecology. *Ecology Letters*, **17**, 82–91.

R Core Team, D. (2016) *R: A Language and Environment for Statistical Computing*, Vienna, Austria.

Reich, P.B. (2014) The world-wide "fast–slow" plant economics spectrum: a traits manifesto. *Journal of Ecology*, **102**, 275–301.

Reich, P.B., Ellsworth, D.S., Walters, M.B., Vose, J.M., Gresham, C., Volin, J.C. & Bowman, W.D. (1999) Generality of leaf trait relationships: a test across six biomes. *Ecology*, **80**, 1955–1969.

Reich, P.B., Rich, R.L., Lu, X., Wang, Y.-P. & Oleksyn, J. (2014) Biogeographic variation in evergreen conifer needle longevity and impacts on boreal forest carbon cycle projections. *Proceedings of the National Academy of Sciences*, **111**, 13703–13708.

Reich, P.B., Sendall, K.M., Stefanski, A., Wei, X., Rich, R.L. & Montgomery, R.A. (2016) Boreal and temperate trees show strong acclimation of respiration to warming. *Nature*, **531**, 633–636.

Reich, P.B., Walters, M.B. & Ellsworth, D.S. (1992) Leaf Life-Span in Relation to Leaf, Plant, and Stand Characteristics among Diverse Ecosystems. *Ecological Monographs*, **62**, 365–392.

Reich, P.B., Wright, I.J. & Lusk, C.H. (2007) Predicting leaf physiology from simple plant and climate attributes: a global GLOPNET analysis. *Ecological Applications*, **17**, 1982–1988.

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

Reu, B., Zaehle, S., Proulx, R., Bohn, K., Kleidon, A., Pavlick, R. & Schmidtlein, S. (2011) The role of plant functional trade-offs for biodiversity changes and biome shifts under scenarios of global climatic change. *Biogeosciences*, **8**, 1255–1266.

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**, 7137.

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**, 2013.

Schürmann, G., Kaminski, T., Köstler, C., Carvalhais, N., Vo\s sbeck, M., Kattge, J., Giering, R., Rödenbeck, C., Heimann, M. & Zaehle, S. (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**, 2999–3026.

Shi, Z., Sherry, R., Xu, X., Hararuk, O., Souza, L., Jiang, L., Xia, J., Liang, J. & Luo, Y. (2015) Evidence for long-term shift in plant community composition under decadal experimental warming. *Journal of Ecology*, **103**, 1131–1140.

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

Thum, T., MacBean, N., Peylin, P., Bacour, C., Santaren, D., Longdoz, B., Loustau, D. & Ciais, P. (2017) The potential benefit of using forest biomass data in addition to carbon and water flux measurements to constrain ecosystem model parameters: Case studies at two temperate forest sites. *Agricultural and Forest Meteorology*, **234-235**, 48 – 65.

Verheijen, L.M., Aerts, R., Brovkin, V., Cavender-Bares, J., Cornelissen, J.H.C., Kattge, J. & van Bodegom, P.M. (2015) Inclusion of ecologically based trait variation in plant functional types reduces the projected land carbon sink in an earth system model. *Global Change Biology*, **21**, 3074–3086.

Violle, C., Choler, P., Borgy, B., Garnier, E., Amiaud, B., Debarros, G., Diquelou, S., Gachet, S., Jolivet, C., Kattge, J. & others (2015) Vegetation ecology meets ecosystem science: permanent grasslands as a functional biogeography case study. *Science of The Total Environment*, **534**, 43–51.

Vuichard, N. & Papale, D. (2015) Filling the gaps in meteorological continuous data measured at FLUXNET sites with ERA-Interim reanalysis. *Earth System Science Data*, **7**, 157.

Van de Weg, M.J., Meir, P., Grace, J. & Ramos, G.D. (2012) Photosynthetic parameters, dark respiration and leaf traits in the canopy of a Peruvian tropical montane cloud forest. *Oecologia*, **168**, 23–34.

Wright, I.J., Reich, P.B., Cornelissen, J.H., Falster, D.S., Groom, P.K., Hikosaka, K., Lee, W., Lusk, C.H., Niinemets, Ü., Oleksyn, J. & others (2005) Modulation of leaf economic traits and trait relationships by climate. *Global Ecology and Biogeography*, **14**, 411–421.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. & Villar, R. (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.

Wu, J., Jansson, P.-E., van der Linden, L., Pilegaard, K., Beier, C. & Ibrom, A. (2013) Modelling the decadal trend of ecosystem carbon fluxes demonstrates the important role of functional changes in a temperate deciduous forest. *Ecological modelling*, **260**, 50–61.

Xia, J.Y., Luo, Y.Q., Wang, Y.-P., Weng, E.S. & Hararuk, O. (2012) A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, **5**, 1259–1271.

852 Xu, X., Medvigy, D., Joseph Wright, S., Kitajima, K., Wu, J., Albert, L.P., Martins, G.A., Saleska,
853 S.R. & Pacala, S.W. (2017) Variations of leaf longevity in tropical moist forests predicted by a
854 trait-driven carbon optimality model. *Ecology letters*, **20**, 1097–1106.
855

Table 1: Description of the 14 optimized parameters and associated processes. All the parameters are common to each PFT. Kbm and bbdate are scaling factor added in the model to improve the optimization of the seasonal cycle of the GPP but are not analyzed in the study. (See Table S1.2; Appendix 1 for the detailed equations involving each parameter).

Parameter	**Description**	**Processes involved**
SLA	Specific leaf area (m^2 gC^{-1})	Photosynthesis, Phenology, Allocation
Lage	Leaf lifespan (days)	Photosynthesis, Phenology
Vcmax	Maximal carboxylation rate limited by CO_2 (µmol m^{-2} s^{-1})	Photosynthesis
Vj/Vc	Ratio between the maximal carboxylation rate limited by light and Vcmax	Photosynthesis
Topt	Optimal temperature of the photosynthesis (°C)	Photosynthesis
gslope	Slope of the Ball-Berry relationship for the stomatal conductance	Photosynthesis, energy budget
LAImax	Maximal leaf area index	Photosynthesis, Phenology, Allocation
Klai	Minimal fraction of LAImax for the use of carbohydrate reserves	Allocation
bbdate	Budburst date (day of the year)	Phenology
tauleaf	Period after budburst during which the use of carbohydrates is allowed	Allocation
Csenes	Temperature for leaf senescence (used only for deciduous)	Phenology
Kbm	Multiplicative factor for the initial leaf biomass (used only for evergreens)	Phenology, Allocation
Kroot	Exponential factor describing the root profile and depth	Water budget, Photosynthesis
Wlim	Minimal threshold at which the photosynthesis becomes limited by water availability	Photosynthesis

Table 2: Description of bio-climatic variables calculated at each site and for each year.

Variable	Description	Units
LAT	Latitude	°N
MAT	Mean annual temperature	Celsius
TMAX	Mean temperature of the warmest month of the year	Celsius
TMIN	Mean temperature of the coldest month of the year	Celsius
TVAR	Temperature difference between TMAX and TMIN	Celsius
DTR	Yearly average of diurnal temperature range	Celsius
MAP	Mean annual precipitation	mm y^{-1}
REH	Mean annual relative humidity	%
SW	Mean annual downward shortwaves radiation (visible and near IR)	W m^{-2}
PDRY	The driest quarter of the year is determined (to the nearest week), and the total precipitation over this period is calculated.	mm y^{-1}
RELP	PDRY divided by MAP	fraction
SHUM	Yearly averaged soil humidity	fraction
GSL	MATgs, DTRgs, SWgs, MAPgs, REHgs and SHUMgs are the same variables averaged during the growing season of the plant	-

Table 3 : Relationships between trait-related parameters. For some relationships, values are log-transformed (x). For each relationship is the number of sites, as well as the correlation coefficient (r, blue when negative; red when positive). Only relationships with an absolute and significant (p-value<0.05) correlation coefficient higher than 0.5 are listed for the different groups of PFT : all, broadleaves (bro ; TroEB, TemEB, TDB, BDB), needleleaves (need ; TEN, BEN), evergreens (ever ; TroEB, TemEB, TEN, BEN), deciduous (dec ; TDB, BDB) and C3 grasses (gra). Note that evergreens include needleleaves and that broadleaves include deciduous. The type of relationship is given for each trait: 0= verified with ecological observations; 1=partially verified on similar data or 3= different from observations. When available, the reference for verification is given. Well constrained parameters are in bold, parameters with a risk of equifinality are normal, poorly constrained parameters are in italics. Refer to Table 1 for the description of each parameter.

Parameters		**r**	**PFT**	**log**	**SMA slope**	**n sites**	**references**	**Type**
Lage	**SLA**	-0.67	ever	x	-1.39	49	(Reich *et al.*, 1999); (Wright *et al.*, 2004)	0
		-0.53	bro	x	-3.47	37		0
		-0.63	All	x	-1.92	98		0
Lage	Vcmax	-0.90	Bro	x	-1.69	37	(Xu *et al.*, 2017)	0
		-0.65	Dec		-2.15	23		0
		-0.59	All	x	-3.13	98		0
gslope	**Lage**	-0.70	Bro	x	-0.74	37	(Reich *et al.*, 1992) (Poorter & Bongers, 2006)	1
		-0.57	Grass		0.00	26		1
gslope	**SLA**	-0.62	Ever		-534.01	49	(Poorter & Bongers, 2006)	3
		0.52	Dec		418.99	23		1
		-0.51	Grass		-235.65	26		3
LAImax	**SLA**	0.60	Need		422.11	35	(Pierce *et al.*, 1994)	1
SLA	Vcmax	-0.55	Ever	x	-1.28	49	(Niinemets *et al.*, 2007)	1
		-0.53	Need	x	-0.75	35		1
gslope	*Wlim*	0.70	Bro	x	1.61	37	(Lin *et al.*, 2015)	3
		0.52	Ever	x	1.47	49		3

Table 4: Relationships between trait-related parameters and climate variables. For some relationships, traits values are log-transformed (x). For each relationship is given the correlation coefficient (r). Only relationships with an absolute (and significant p-value<0.05) correlation coefficient higher than 0.5 are listed for the different groups of PFT : all, broadleaves (bro ; TroEB, TemEB, TDB, BDB), needleleaves (need ; TEN, BEN), evergreens (ever ; TroEB, TemEB, TEN, BEN), deciduous (dec ; TDB, BDB) and C3 grasses (gra). The type of relationship is given for each trait: 0= verified with ecological observations; 1=partially verified on similar data; 2= not verified or, 3=different from observations. When available, the reference for verification is given. Well constrained parameters are in bold, parameters with a risk of equifinality are normal. Refers to Table 1 and 2 for the description of each parameter and climate variables respectively.

Trait	Climat	r	PFT	log	SMA slope	Référence	Type
Lage	LAT	0.59	ever		24.90	(Reich *et al.*, 2014)	0
		-0.56	bro		-13.44	-	2
	MAP	0.66	grass		1.14	-	2
		-0.65	need	x	-0.66	(van Ommen Kloeke *et al.*, 2012)	0
	MAT	-0.78	ever	x	-16.95	(Wright *et al.*, 2005; van Ommen Kloeke *et al.*, 2012; Reich *et al.*, 2014)	0
		-0.62	need	x	-17.93		0
		0.54	grass	x	107.81	-	2
		0.53	bro	x	30.14	-	2
	SW	-0.53	ever	x	-1.84	(Poorter & Bongers, 2006)	1
		0.52	bro	x	3.85	-	2
	TMIN	-0.65	ever	x	-30.99	(Wright *et al.*, 2005; van Ommen Kloeke *et al.*, 2012; Reich *et al.*, 2014)	1
SLA	MAP	0.54	need	x	0.37	(Wright *et al.*, 2005)	0
	MAPgs	0.57	ever	x	0.47		0
	MAT	0.65	ever	x	12.16	(Wright *et al.*, 2005)	0
	MATgs	-0.63	bro	x	-0.86		0
		-0.55	dec	x	-0.96		0
	RELP	0.60	dec	x	0.25	-	2
		0.59	bro		0.08	-	2
	SW	0.53	ever		0.00	(Givnish *et al.*, 2004; Poorter & Bongers, 2006; Reich *et al.*, 2014)	1
gslope	MAP	-0.59	grass	x	-1.12	-	2
	PDRY	0.58	dec		0.02	(Lin *et al.*, 2015)	1
	REH	0.64	dec		19.24		1
	RELP	0.66	bro		42.67		1
		0.58	dec		29.05		1
	SHUMgs	0.71	dec		20.53		1
	SW	-0.63	bro		-0.10	-	2
	SWgs	-0.55	dec		-0.08	-	2
Vcmax	MAT	-0.52	bro		-4.77	(Ali *et al.*, 2015)	3
	RELP	0.60	bro		511.72	-	2
	SW	-0.54	bro		-1.15	(Ali *et al.*, 2015)	3

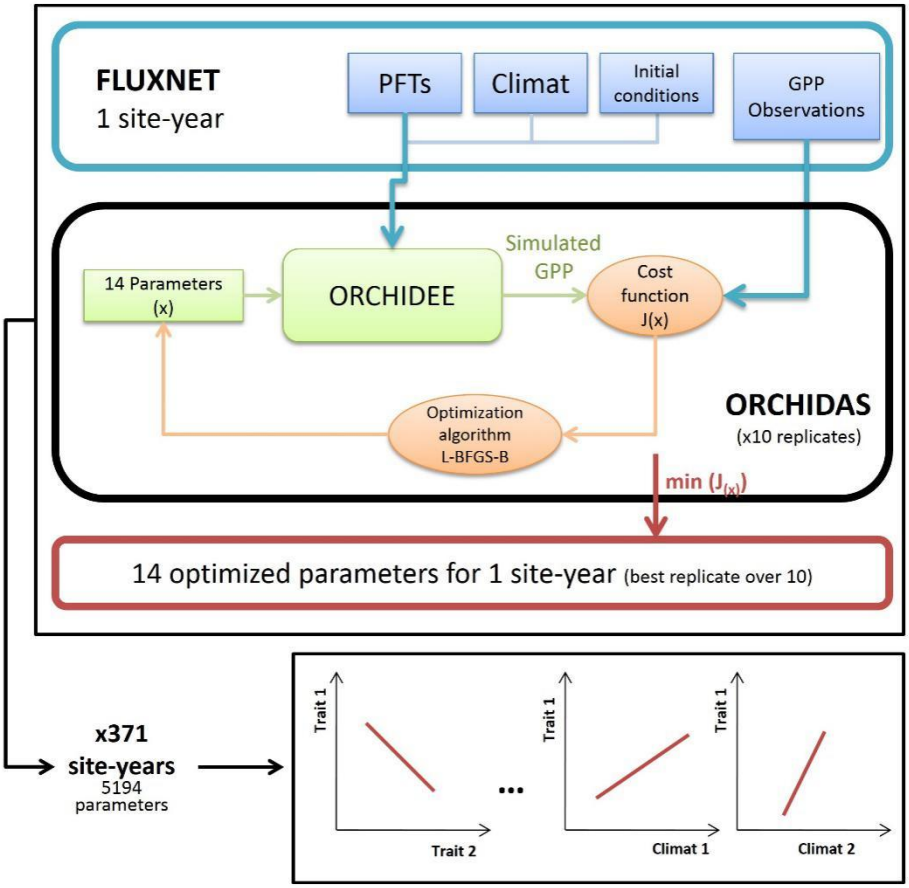

Figure 1: Schematic representation of the modeling protocol followed in this study. For each FLUXNET site-year (blue), the model ORCHIDEE (green) was calibrated with the data assimilation system ORCHIDAS (red) in order to reproduce GPP observations. The ORCHIDAS system uses a gradient-based approach (L-BFGS-B) to reduce the cost function J(x). For each site-year, 14 parameters (listed in Table 1) were optimized 10 times with different initial values. The best calibration, i.e. leading to the minimum value of J(x), was retained. This procedure was repeated for each site-year, resulting in 371 sets of 14 independently optimized parameters. Finally, correlations between optimized parameters and climate were explored using standardized major axis regressions.

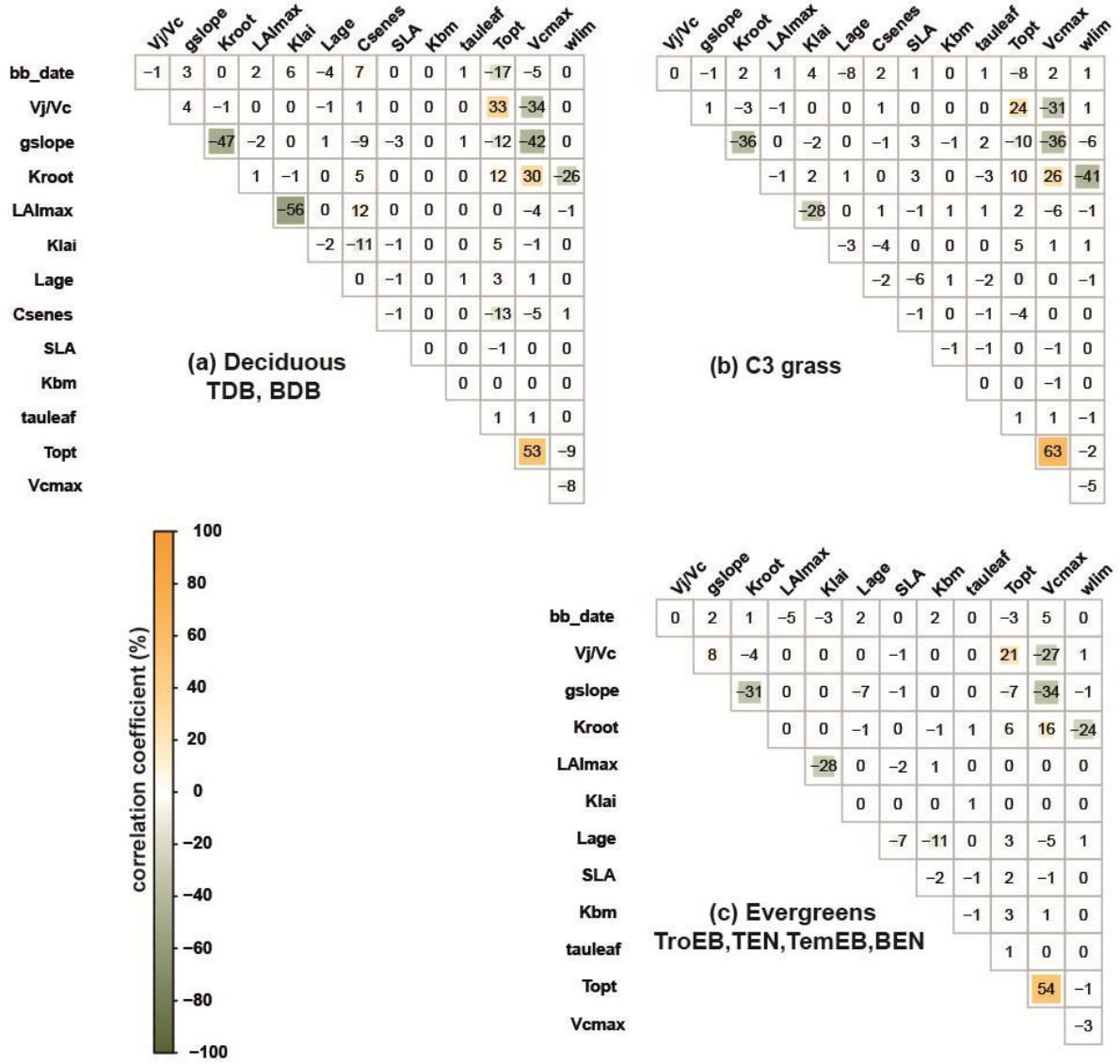

Figure 2: Error correlation between optimized parameters (derived from the A matrix) averaged over deciduous trees, evergreen trees, and C3 grass. The color scale gives the error correlation coefficient. For more clarity, the coefficient is indicated in % in each matrix cell. The description of each parameter is listed in Table 1.

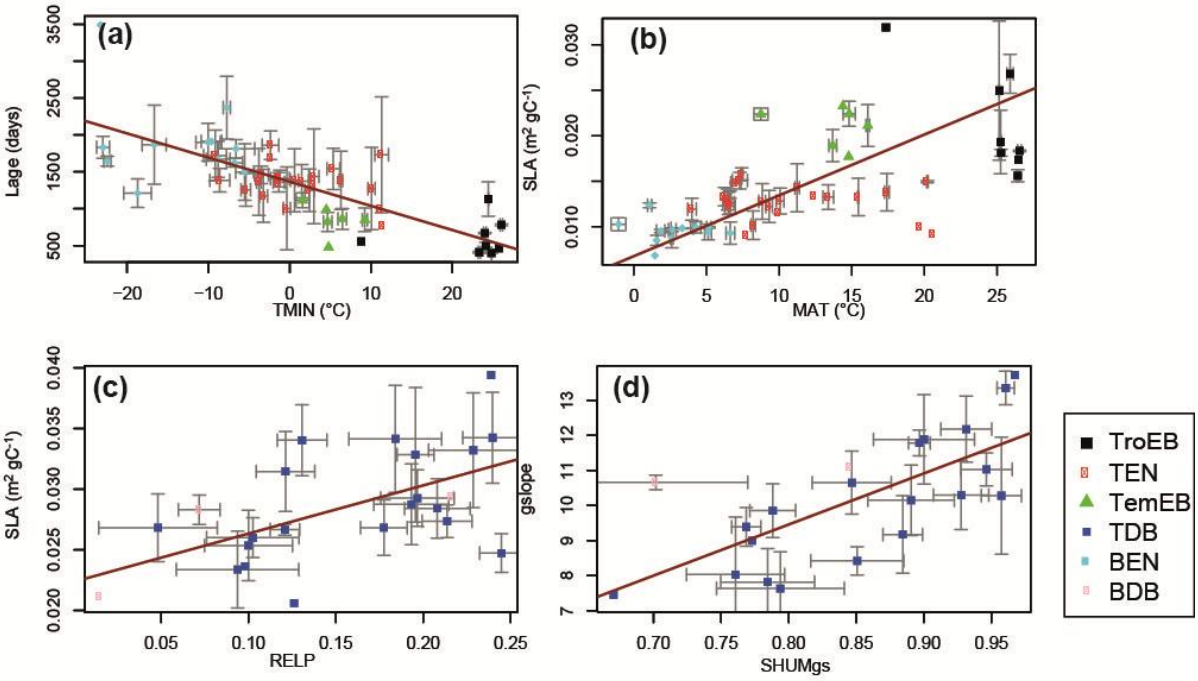

Figure 3: Four examples of co-variations obtained between optimized parameters (Table 1) and environmental conditions (Table 2) of the sites for PFTs TroEB (black square), TEN (red square), TemEB (green triangle), TDB (blue square), BEN (cyan dots) and BDB (pink dots). Each point represents the mean optimized parameter (environmental variable) value for one site while the error bars represent the inter-annual variability (no bars means only one year of measurement). The red line represents the slope of the standardized major axis regression.