



HAL
open science

Carbon and water fluxes of the boreal evergreen needleleaf forest biome constrained by assimilating ecosystem carbonyl sulfide flux observations

Camille Abadie, Fabienne Maignan, Marine Remaud, Kukka-maaria Kohonen,
Wu Sun, Linda Kooijmans, Timo Vesala, Ulli Seibt, Nina Raoult, Vladislav
Bastrikov, et al.

► To cite this version:

Camille Abadie, Fabienne Maignan, Marine Remaud, Kukka-maaria Kohonen, Wu Sun, et al.. Carbon and water fluxes of the boreal evergreen needleleaf forest biome constrained by assimilating ecosystem carbonyl sulfide flux observations. *Journal of Geophysical Research: Biogeosciences*, 2023, 128 (7), pp.e2023JG007407. 10.1029/2023JG007407 . hal-04142530

HAL Id: hal-04142530

<https://hal.science/hal-04142530>

Submitted on 20 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



RESEARCH ARTICLE

10.1029/2023JG007407

Key Points:

- Jointly assimilating ecosystem carbonyl sulfide (COS) flux and gross primary production data improves both simulated plant carbon uptake and transpiration
- Assimilating ecosystem COS flux does not increase gross primary production (GPP) over boreal evergreen needleleaf forests, in contrast with previous inversion studies
- COS flux observations help to identify misrepresentation of the sensitivity of GPP to droughts in models

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

C. Abadie,
camille.abadie.research@gmail.com

Citation:

Abadie, C., Maignan, F., Remaud, M., Kohonen, K.-M., Sun, W., Kooijmans, L., et al. (2023). Carbon and water fluxes of the boreal evergreen needleleaf forest biome constrained by assimilating ecosystem carbonyl sulfide flux observations. *Journal of Geophysical Research: Biogeosciences*, 128, e2023JG007407. <https://doi.org/10.1029/2023JG007407>

Received 23 JAN 2023
 Accepted 15 JUN 2023

Carbon and Water Fluxes of the Boreal Evergreen Needleleaf Forest Biome Constrained by Assimilating Ecosystem Carbonyl Sulfide Flux Observations

Camille Abadie¹ , Fabienne Maignan¹ , Marine Remaud¹, Kukka-Maaria Kohonen^{2,3} , Wu Sun⁴ , Linda Kooijmans⁵, Timo Vesala^{2,6} , Ulli Seibt⁷, Nina Raoult¹ , Vladislav Bastrikov¹ , Sauveur Belviso¹ , and Philippe Peylin¹

¹Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France, ²Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland, ³Department of Environmental Systems Science, Institute of Agricultural Sciences, ETH Zurich, Zurich, Switzerland, ⁴Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA, ⁵Meteorology and Air Quality, Wageningen University and Research Centre, Wageningen, The Netherlands, ⁶Institute for Atmospheric and Earth System Research/Forest Sciences, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland, ⁷Department of Atmospheric & Oceanic Sciences, UCLA, Los Angeles, CA, USA

Abstract Gross primary production (GPP) by boreal forests is highly sensitive to environmental changes. However, GPP simulated by land surface models (LSMs) remains highly uncertain due to the lack of direct photosynthesis observations at large scales. Carbonyl sulfide (COS) has emerged as a promising proxy to improve the representation of GPP in LSMs. Because COS is absorbed by vegetation following the same diffusion pathway as CO₂ during photosynthesis and not emitted back to the atmosphere, incorporating a mechanistic representation of vegetation COS uptake in LSMs allows using COS observations to refine GPP representation. Here, we perform ecosystem COS flux and GPP data assimilations to constrain the COS- and GPP-related parameters in the ORCHIDEE LSM for boreal evergreen needleleaf forests (BorENF). Assimilating ecosystem COS fluxes at Hyttiälä forest increases the simulated net ecosystem COS uptake by 14%. This increase largely results from changes in the internal conductance to COS, highlighting the need to improve the representation of COS internal diffusion and consumption. Moreover, joint assimilation of ecosystem COS flux and GPP at Hyttiälä improves the simulated latent heat flux, contrary to the GPP-only data assimilation, which fails to do so. Finally, we scaled this assimilation framework up to the boreal region and find that the joint assimilation of COS at Hyttiälä and GPP fluxes at 10 BorENF sites increases the modeled vegetation COS uptake up to 18%, but not GPP. Therefore, this study encourages the use of COS flux observations to inform GPP and latent heat flux representations in LSMs.

Plain Language Summary Carbon uptake by boreal forests is highly sensitive to environmental changes. There is large uncertainty about how much carbon dioxide (CO₂) boreal forests absorb through photosynthesis, as represented by land surface models. Carbonyl sulfide (COS), a trace gas that tracks photosynthesis, can help improve the representation of simulated plant CO₂ uptake because COS and CO₂ share a common pathway during leaf uptake. Using a mechanistic model of biospheric COS processes implemented in the ORCHIDEE land surface model, we assimilated ecosystem COS flux and plant CO₂ uptake measured at Hyttiälä boreal forest. We find that this joint assimilation improves the simulated plant CO₂ uptake, as well as transpiration because of the strong link between COS, CO₂ and H₂O fluxes through stomatal diffusion. Scaling up this assimilation framework to evergreen needleleaf boreal forests, we find that assimilating ecosystem COS flux and plant CO₂ uptake data increases the vegetation COS uptake for this biome, but not plant CO₂ uptake. Our results imply that COS has the potential to constrain both plant carbon uptake and transpiration in land surface models, which should be further investigated, especially during drought events.

1. Introduction

Boreal forests absorb a significant amount of atmospheric CO₂ through gross primary production (GPP), representing about 20% of the global GPP (Jung et al., 2017; Tramontana et al., 2016). However, direct observations of GPP over the whole boreal region are not available as photosynthesis cannot be measured at scales larger than the leaf scale. At the ecosystem scale, photosynthesis—that is GPP—can be inferred from partitioning eddy

© 2023. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

covariance observations of net ecosystem exchange into respiration components and GPP. However, flux partitioning methods rely on important assumptions about the relationship between fluxes and their environmental drivers and are impacted by various sources of uncertainty (Tramontana et al., 2020). At regional to global scales, land surface models (LSMs) can simulate GPP from process representations, but the lack of direct GPP measurements makes it challenging to evaluate and improve the representation of GPP in LSMs (Anav et al., 2015). In addition, because boreal forests are highly sensitive to environmental changes, rapid ongoing changes in this biome impact gas exchange and lead to large uncertainties in GPP estimates simulated by LSMs (Bonan, 2008; Fisher et al., 2018) or obtained from data-driven methods (Goetz et al., 2005).

To address the uncertainty in GPP, several optical proxies have been used to infer GPP estimates or to improve GPP representation in LSMs. Vegetation indices (VIs) inform GPP seasonality by tracking seasonal changes in vegetation greenness (Shen et al., 2014). However, the small variations in vegetation greenness over the season for evergreen boreal forests limit the effectiveness of such VIs in tracking photosynthetic activity. Changes in snow cover also affect the monitoring of seasonal variations in greenness and the estimation of GPP using VIs (Beck et al., 2006; Böttcher et al., 2014; Delbart et al., 2005). Solar-induced fluorescence (SIF), radiation in the red or far-red bands emitted from illuminated chlorophylls, can be retrieved from satellite observations and strongly correlates with GPP at large spatial and temporal scales (Frankenberg et al., 2011). However, SIF is also affected by snow cover and previous studies have reported differences in phenology between GPP derived from SIF and that from VIs (Chang et al., 2019; Li et al., 2018), especially under water stress (Wang et al., 2019).

Vegetation carbon uptake is also coupled with water loss, both controlled by stomatal diffusion. Plant transpiration is a key process for ecosystem functioning, sharing common environmental drivers with GPP. Transpiration also faces similar challenges as GPP as it is not well constrained by observations (Schlesinger & Jasechko, 2014), limiting the ability of LSMs to accurately represent its spatial and temporal dynamics (Mencuccini et al., 2019). Transpiration measurements are performed at the leaf or branch scale, it is difficult to upscale them to the ecosystem level (Jarvis, 1995). At the ecosystem scale, evapotranspiration can be measured using flux tower remote sensing methods (Wang & Dickinson, 2012), but it includes not only plant transpiration but also evaporation from the soil and other surfaces due to water interception, or snow sublimation.

Carbonyl sulfide (COS) has emerged as a promising tracer to track GPP (Campbell et al., 2008; Montzka et al., 2007; Sandoval-Soto et al., 2005; Seibt et al., 2010; Stimler et al., 2010), providing complementary information to existing optical proxies of GPP. COS, an atmospheric trace gas, is absorbed by vegetation following the same diffusion pathway as CO₂ during photosynthesis. Inside the leaves, COS is presumed to be totally hydrolyzed by the carbonic anhydrase (CA), an enzyme also involved in CO₂ fixation during photosynthesis (DiMario et al., 2016), and is normally not emitted back to the atmosphere. Therefore, the main advantage of COS lies in the fact that it allows GPP to be estimated independently of CO₂ measurements. In addition, COS helps to constrain stomatal diffusion, which determines the coupling between CO₂ uptake and water loss (Berkelhammer et al., 2020; Kooijmans et al., 2017; Wehr et al., 2017).

Previous studies have used COS to infer stomatal conductance, to investigate stomatal control on GPP, and to explore the physiological links between plant COS uptake, GPP, and transpiration (Berkelhammer et al., 2014; Wehr et al., 2017; Wohlfahrt et al., 2012). Alternatively, some studies have linked vegetation COS uptake to GPP using the leaf relative uptake (LRU) approach, based either on empirical ratios between COS and CO₂ deposition rates in plants (Asaf et al., 2013; Kooijmans et al., 2019), or more recently based on stomatal conductance theories as a function of humidity, temperature, light, or CO₂ concentration (Kohonen, Dewar, et al., 2022; Sun et al., 2022). Process-based representations of COS uptake by plants have also been implemented in LSMs (Berry et al., 2013; Kooijmans et al., 2021; Maignan et al., 2021). These mechanistic models simulate the dynamics of plant COS uptake and of the conductances involved in stomatal diffusion, and provide new global estimates of the vegetation COS sink.

However, the representation of COS exchanges between the atmosphere and land ecosystems rely on parameterizations that are still highly empirical and supported by limited measurements. Indeed, multiyear observations of ecosystem COS fluxes are only available at two sites, Harvard Forest in the United States with 2 years (2012–2013) of measurements and Hyytiälä Forest in Finland with 5 years (2013–2017) of measurements. So far, two studies have used COS flux measurements to directly constrain parameters related to COS and GPP in LSMs (Chen et al., 2023; Cho et al., 2023). Chen et al. (2023) focused on optimizing two parameters in the Boreal Ecosystem Productivity Simulator (BEPS) LSM at these two sites, one parameter related to the carboxylation rate

of the Rubisco enzyme, which is also involved in vegetation COS uptake representation, and another parameter specific to the vegetation COS model. The impacts of other physiological parameters on the simulated ecosystem COS flux and GPP remain largely unexplored.

While vegetation uptake is the largest COS sink, several sinks and sources also contribute to the global COS budget (Whelan et al., 2018). Soils can absorb COS because soil microorganisms contain CA and other COS hydrolases (Masaki et al., 2021), or emit COS due to microbial or abiotic COS production (Whelan et al., 2018). The ocean is a significant source of COS through direct emissions or indirect emissions via dimethyl sulfide (DMS) and carbon disulfide (CS₂) (Lennartz et al., 2017, 2020). Anthropogenic activities are also a major source of COS (Zumkehr et al., 2018), and biomass burning contributes to COS surface emissions (Stinecipher et al., 2019). In the atmosphere, COS can be destroyed through oxidation by OH radicals in the lower troposphere or photolysis (Whelan et al., 2018).

Currently, a major challenge of using COS as a global-scale GPP proxy is the imbalance of the global COS budget, as recently highlighted in the intercomparison of atmospheric COS transport models conducted by Remaud et al. (2023). A previous inversion study by Ma et al. (2021) has suggested the likely presence of a missing COS source in the tropics, and an underestimated COS sink in northern high latitudes. Another atmospheric inversion of COS and CO₂ concentrations carried out by Remaud et al. (2022) supports these findings for COS while also showing an underestimation of GPP simulated by the ORCHIDEE LSM in the high latitudes. Moreover, Vesala et al. (2022) used the longest ecosystem COS flux measurements (5 years) made at the Hyytiälä forest, Finland, to develop an empirical parametrization of vegetation COS fluxes based on environmental drivers. This parametrization was scaled up to boreal evergreen needleleaf forest biome using the SiB4 LSM (Haynes et al., 2020), leading to an increase in COS uptake in the high latitudes, compared to the one computed using the mechanistic COS model implemented in SiB3 by Berry et al. (2013).

In this context, the goal of this study is to evaluate the potential of COS to directly constrain the representation of GPP in the ORCHIDEE LSM using a data assimilation framework. This work addresses the following questions:

1. To what extent can biospheric COS flux measurements help to constrain stomatal diffusion of CO₂ and COS, and as a result, GPP?
2. What is the impact of assimilating biospheric COS fluxes on the simulated latent heat flux (LE)?
3. How does assimilating ecosystem COS flux observations at one site along with GPP data from multiple sites impact the simulated COS, GPP, and LE fluxes over the entire biome of boreal evergreen needleleaf forests?

Here, we optimize the GPP- and COS-related parameters by assimilating GPP and ecosystem COS fluxes at the Hyytiälä forest to evaluate the impact on the simulated vegetation COS and CO₂ uptakes, and LE. In particular, we focus on the constraint provided by these flux data assimilations during a severe drought event in 2006. Then, we perform multi-site assimilations to assess the changes in GPP, LE, and vegetation COS fluxes for the boreal evergreen needleleaf forest biome. Finally, we discuss necessary improvements in COS flux representation in LSMs, as highlighted in this study, and the implications of assimilating ecosystem COS fluxes to constrain GPP compared to using the empirical LRU approach.

2. Methods

The general workflow of this study is represented in Figure 1 illustrating the main steps carried out. First, sensitivity analyses over the variables of interest were conducted following prior simulations in ORCHIDEE. This led to selecting the model parameters to include in the optimizations, which were performed either at the site scale or considering multiple sites for the assimilation. Finally, the optimized simulations following these two optimization procedures are evaluated against eddy covariance data and global evaluation products. These different steps are detailed in the following sections.

2.1. Observation Data Sets

2.1.1. Description of the Studied Sites

We selected 16 boreal evergreen needleleaf forest sites from the FLUXNET network. All sites are located in North America and Europe between 40° and 68°N, with each continent featuring eight sites, as illustrated in

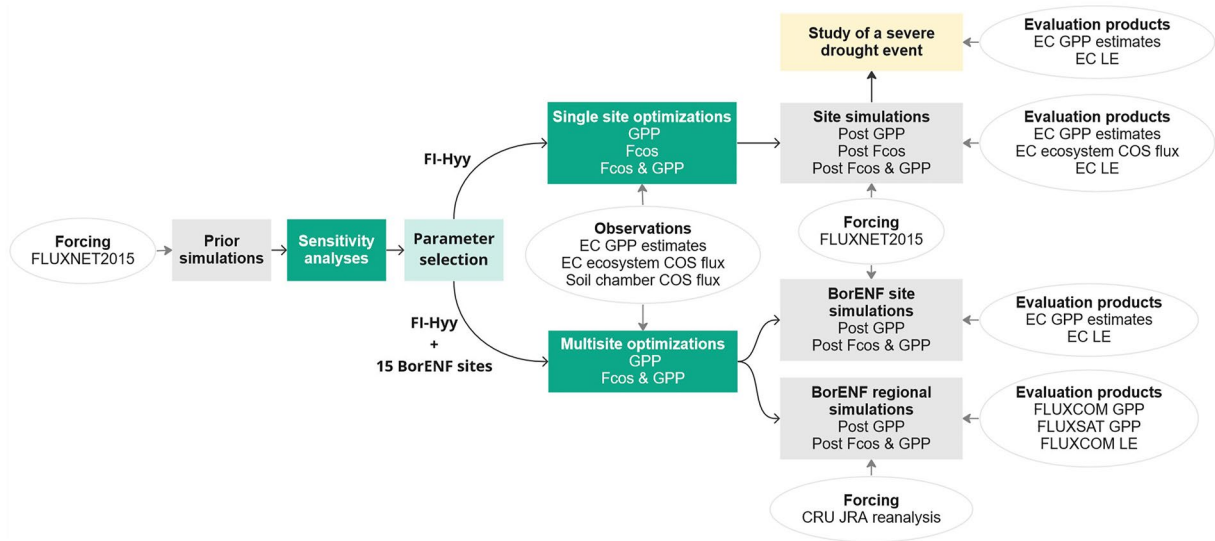


Figure 1. Overview of the workflow, showing the input data and main steps of this study. Actions are represented in boxes, whereas input data used as forcing, observations, and evaluation products, are in circles. EC, eddy covariance; FI-Hyy, Hyttiälä forest in Finland; BorENF, Boreal evergreen needleleaf forest; GPP, gross primary production; Fcos, ecosystem COS flux; LE, latent heat flux.

Figure 2. Mean annual air temperatures range from -1.4°C to 6.9°C , and mean annual precipitation ranges between 149 and 1,440 mm across these sites. Two sites are located at a high elevation: Davos (Switzerland) at 1,639 m and Niwot Ridge (United States) at 3,050 m. These 16 sites were selected because the normalized root mean square deviation (nRMSD) between the FLUXNET GPP estimates and the GPP simulated in ORCHIDEE is lower than 25%. Sites for which the nRMSD exceeds this threshold may have been impacted by processes not represented in the ORCHIDEE version used in this study, such as fires or clear-cuts, and are therefore not selected for data assimilation. A description of the sites is provided in Table S1 in Supporting Information S1.

Among these 16 sites, COS measurements were only carried out at Hyttiälä forest, Finland (FI-Hyy; 61.845°N , 24.288°E). FI-Hyy is a coniferous forest planted in 1962 (Suni et al., 2003) dominated by Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*). The climate is boreal with an annual mean temperature of 4.5°C and an annual mean precipitation of 632 mm. The soil type is described as Haplic Podzol (Sun et al., 2018).

2.1.2. Site Measurements of COS, GPP, and LE Fluxes

We used GPP and LE measurements from the FLUXNET global network (La Thuile: Baldocchi et al., 2001 or FLUXNET2015: Pastorello et al., 2020), which are available at a half-hourly time step for the 16 selected sites. LE and net ecosystem exchange (NEE) are measured using the eddy covariance (EC) method. Then, GPP is retrieved from NEE measurements based on a nighttime partitioning method (Reichstein et al., 2005) using a variable friction velocity (“U-star”) threshold for each year (VUT). This means that daytime respiration is first estimated with a respiration model parameterized with nighttime NEE data, and GPP is then obtained as

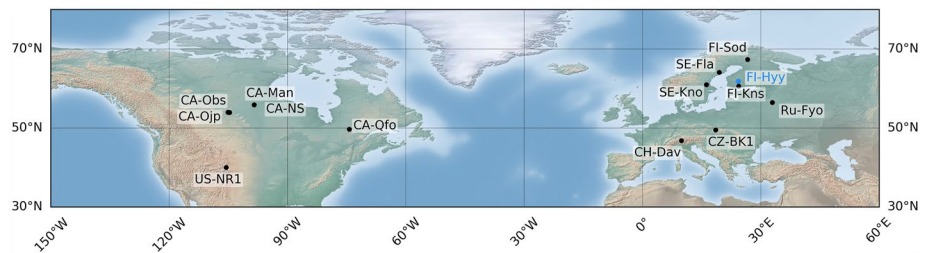


Figure 2. Location of the 16 studied sites. The CA-NS label includes CA-NS1, CA-NS2, and CA-NS5. FI-Hyy is represented in blue as it is the only site with COS observations, while the other sites are represented in black. The background map corresponds to the “shadedrelief” map (<http://www.shadedrelief.com>) from the matplotlib basemap toolkit (https://matplotlib.org/basemap/api/basemap_api.html).

the difference between respiration and NEE (see Pastorello et al., 2020 for details). No significant difference was found between FLUXNET GPP retrieved from the nighttime or daytime (Lasslop et al., 2010) partitioning method at FI-Hyy (not shown).

For COS observations, EC measurements were carried out at FI-Hyy from 2013 to 2017 (Vesala et al., 2022), along with soil chamber measurements in 2015 (Sun et al., 2018) and branch chamber measurements in 2017 (Kooijmans et al., 2019). EC fluxes were measured at 23 m height using an Aerodyne quantum cascade laser spectrometer (QCLS, Aerodyne Research, Billerica, MA, USA). EC data were processed following the recommendations by Kohonen et al. (2020) regarding quality-check and gap-filling, resulting in a half-hourly EC flux data set. EC observations are available from April to November in 2013, March to September in 2014, July to October in 2015, April to November in 2016, and January to August in 2017, with a lack of data during winter-time. Soil COS measurements were conducted from July–November 2015 using two automated soil chambers connected to another Aerodyne QCLS, the same model used for EC observations (Sun et al., 2018). Understory herbs and bryophytes were removed prior to performing chamber measurements to eliminate the influences of plant COS uptake on observed soil COS fluxes.

2.1.3. Characterization of a Drought Event at Hyytiälä

We focused on a drought event that occurred at FI-Hyy to investigate the potential of COS to constrain GPP and LE under these specific stress conditions. FI-Hyy did not undergo any drought event between 2013 and 2017 when the EC COS measurements were carried out (Vesala et al., 2022). However, a severe drought was reported in summer 2006, causing large damage in southern Finland (Gao et al., 2017). The intensity of this drought was characterized by the soil moisture index (SMI), which is defined as the difference between observed volumetric soil moisture and volumetric soil moisture at wilting point, divided by the difference between volumetric soil moisture at field capacity and at wilting point. Gao et al. (2017) found very dry conditions (SMI <0.20) for 37 consecutive days (from 23 July to 28 August) with the most severe part of the drought between 1 and 17 August (SMI <0.15).

2.1.4. GPP and LE Global Observation Products

We evaluated the simulated fluxes over the whole boreal evergreen needleleaf forest region with several global GPP and LE products. First, we used global GPP data products from the FLUXCOM version RS (Jung et al., 2019, 2020) and FLUXSAT version 2.0 (Joiner et al., 2018) databases, which are produced by applying different machine-learning upscaling methods to FLUXNET EC measurements and remote sensing data. FLUXCOM GPP is available between 2001 and 2015, whereas FLUXSAT GPP is available from 2000 to 2020. The FLUXCOM database also provides global estimates of LE derived from the same approach as for GPP. In addition, we considered a second global product for LE from the Global Land Evaporation Amsterdam Model (GLEAM), which separately computes different components of evapotranspiration (ET) from satellite data (Martens et al., 2017) from 1980 to 2021. For all global products, we used monthly average fluxes at a 0.5° spatial resolution to match the temporal and spatial resolution of our global simulated fluxes.

2.2. Model Descriptions

2.2.1. The ORCHIDEE Land Surface Model

The ORCHIDEE land surface model (LSM) is developed at the Institut Pierre Simon Laplace (IPSL). Here, we used the version involved in the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Boucher et al., 2020; Cheruy et al., 2020). ORCHIDEE solves the water, carbon and energy budget between land surfaces and the atmosphere (initially described in Krinner et al., 2005). Fast processes such as photosynthesis, hydrology and energy balance are computed every 30 min, while slow processes such as carbon allocation and phenology are run at a daily timestep. The different vegetation types are grouped into plant functional types (PFTs) with similar characteristics in terms of structure, bioclimatic range, leaf phenology, and the photosynthetic pathway. We ran ORCHIDEE distinguishing among 14 classes of PFTs and a class representing bare soil. Each model grid cell is associated with fractions of PFTs prescribed using yearly varying PFT maps derived from the ESA Climate Change Initiative (CCI) land cover products (Poulter et al., 2015).

In ORCHIDEE, photosynthesis is computed following the approach of Yin and Struik (2009) based on Farquhar et al. (1980) for C3 plants and Collatz et al. (1991) for C4 plants. Soil moisture limitations on stomatal diffusion

and photosynthesis are controlled by a stress factor, which varies linearly from 0 at the wilting point (maximum stress) to 1 when soil moisture is close to the field capacity (no stress). This stress factor regulates stomatal conductance by controlling the minimum stomatal conductance (when irradiance approaches zero) and the factor that describes the effect of the leaf-to-air vapor pressure deficit (VPD) (Yin & Struik, 2009). This stress factor also regulates the mesophyll conductance, Rubisco carboxylation, and leaf day respiration.

The stomatal conductance (g_s) is defined following Yin and Struik (2009),

$$g_s = g_0 + \frac{A + R_d}{C_i - C_{i^*}} \cdot f_{\text{VPD}} \quad (1)$$

with g_0 the residual stomatal conductance (m s^{-1}), A the CO_2 assimilation (minimum between the Rubisco-limited rate and electron transport-limited rate) ($\mu\text{mol m}^{-2} \text{s}^{-1}$), R_d the day respiration ($\mu\text{mol m}^{-2} \text{s}^{-1}$), C_i the intercellular CO_2 partial pressure ($\mu\text{mol m}^{-2} \text{s}^{-1}$), C_{i^*} the base CO_2 compensation point in the absence of R_d ($\mu\text{mol m}^{-2} \text{s}^{-1}$), and f_{VPD} the function describing the effect of VPD, defined as,

$$f_{\text{VPD}} = \frac{1}{[1/(A1 - B1 \cdot \text{VPD}) - 1]} \quad (2)$$

where $A1$ and $B1$ are empirical factors (Table S2 in Supporting Information S1).

A global soil map based on the Food and Agriculture Organization of the United Nations/United States Department of Agriculture (FAO/USDA) texture classification describes the distribution of soil textures in 12 classes (Reynolds et al., 2000). The soil texture in each grid cell determines soil properties such as soil porosity, wilting point, and field capacity.

The ORCHIDEE LSM can be run both at the site level or at the global scale. The site simulations were forced by local micro-meteorological measurements from the FLUXNET network (Pastorello et al., 2020). For global simulations, we used the 0.5° 6-hourly CRU JRA reanalysis (University of East Anglia Climatic Research Unit–Japanese Reanalysis; Friedlingstein et al., 2020). Because simulated vegetation COS uptake depends on COS concentrations in the atmospheric boundary layer (Abadie et al., 2022; Kooijmans et al., 2021), we prescribed near-surface COS and CO_2 concentrations using 3-hourly simulated concentrations extracted from the first vertical level (33 m above sea or ground level) of the Laboratoire de Météorologie Dynamique (LMDz) atmospheric transport model output. The atmospheric COS and CO_2 concentrations were obtained using optimized COS and CO_2 surface fluxes inferred from atmospheric inverse modeling as described in Remaud et al. (2022).

2.2.2. ORCHIDEE Simulations

We first ran site simulations at the 16 selected boreal evergreen needleleaf forest sites. This vegetation type is represented by a dedicated PFT (BorENF) in ORCHIDEE. For each site, a “spin-up” phase was performed to reach an equilibrium state at which all carbon pools are stable and the net biome production oscillates around zero in the absence of any disturbances. In ORCHIDEE, around 340 years are usually needed to reach this equilibrium state as the convergence is accelerated using a pseudo-analytical iterative estimation for the soil carbon pools (Lardy et al., 2011). We carried out the spin-up phase at each site by cycling over the available years (Table S1 in Supporting Information S1) in the FLUXNET meteorological forcing file for about 340 years and using a constant atmospheric CO_2 concentration of 312 ppm (corresponding to the year 1950). Then, we performed transient simulations for about 60 years, by cycling over the years in the FLUXNET forcing files, to introduce disturbances related to climate change, land use change, and increasing CO_2 atmospheric concentrations. Following the transient phase, we ran the site simulations over the period available in the FLUXNET forcing file (Table S1 in Supporting Information S1).

For the simulations over the boreal region, we selected a latitudinal range between 40° and 80°N . We followed the same simulation protocol as for site simulations with preceding spin-up and transient phases. The 340-year spin-up phase was performed by cycling over 10 years (1931–1940) of CRU JRA reanalysis forcing files. Then, we carried out the transient phase cycling over the same 10 years in the forcing files where disturbances were introduced from 1940 to 2000, followed by a global simulation from 2000 to 2019.

For site and regional simulations, we evaluated the simulated GPP and biospheric COS fluxes against the corresponding observations using the root mean square deviation (RMSD) and the normalized root mean square

deviation (nRMSD), for which the normalization denominator is defined as the maximum value minus the minimum value of the observations.

2.2.3. Biosphere COS Exchange Models

Mechanistic representations of vegetation and soil COS fluxes have been previously implemented in ORCHIDEE (Abadie et al., 2022; Maignan et al., 2021). The vegetation COS model is based on Berry et al. (2013) and describes COS uptake by plant as a one-way diffusion from the atmosphere to the leaf interior where COS is irreversibly hydrolyzed by CA (Equation 3).

$$F_{\text{COS,veg}} = [\text{COS}] \cdot \left[\frac{1}{g_{b,\text{COS}}} + \frac{1}{g_{s,\text{COS}}} + \frac{1}{g_{i,\text{COS}}} \right]^{-1} \quad (3)$$

with $F_{\text{COS,veg}}$ the vegetation COS uptake flux ($\text{pmol m}^{-2} \text{s}^{-1}$), $[\text{COS}]$ the atmospheric COS concentration (ppt), $g_{b,\text{COS}}$, $g_{s,\text{COS}}$, and $g_{i,\text{COS}}$ the laminar boundary layer, stomatal, and internal COS conductances ($\text{mol m}^{-2} \text{s}^{-1}$), respectively. The internal conductance $g_{i,\text{COS}}$ includes both the COS diffusion through the mesophyll and the COS consumption by CA as a first-order reaction. The latter two were assumed to scale with the photosynthetic capacity (V_{max}) of the Rubisco enzyme ($\mu\text{mol m}^{-2} \text{s}^{-1}$) (Badger & Dean Price, 1994; Berry et al., 2013). Therefore, $g_{i,\text{COS}}$ is expressed as proportional to V_{max} ,

$$g_{i,\text{COS}} = \alpha \cdot V_{\text{max}} \quad (4)$$

where α is a parameter, the value of which depends on the photosynthetic pathway. Typical values of α are $0.0012 \text{ mol } \mu\text{mol}^{-1}$ for C3 and $0.013 \text{ mol } \mu\text{mol}^{-1}$ for C4 plants (Berry et al., 2013). Because COS is assumed to be totally hydrolyzed by CA, COS internal concentration is zero at the terminus.

The soil COS model in ORCHIDEE is based on Ogée et al. (2016), with a distinction between oxic soils and anoxic soils. Anoxic soils such as wetlands, rice paddies, or salt marshes, are represented as a COS source only, while oxic soils can both emit and absorb COS. The oxic soil COS model resolves COS diffusion into the soil matrix, abiotic and biotic COS hydrolysis, and COS production, resulting in the following formulation following Ogée et al. (2016),

$$F_{\text{COS,oxic soil}} = \sqrt{kB\theta D} \cdot \left([\text{COS}] - \frac{z_1^2 P}{D} \left(1 - \exp\left(-\frac{z_{\text{max}}}{z_1}\right) \right) \right) \quad (5)$$

where $F_{\text{COS,oxic soil}}$ is the COS flux from oxic soils ($\text{mol m}^{-2} \text{s}^{-1}$), k is the first-order COS consumption rate constant within the soil (s^{-1}), B is the solubility of COS in water ($\text{m}^3 \text{ water m}^{-3} \text{ air}$), θ is the soil volumetric water content ($\text{m}^3 \text{ water m}^{-3} \text{ soil}$), D is the total effective COS diffusivity ($\text{m}^2 \text{ s}^{-1}$), P is the COS production term expressed as a function of soil temperature ($\text{mol m}^{-2} \text{ s}^{-1}$), $z_1 = D/kB$ (m), and z_{max} is the maximum soil depth (m).

In ORCHIDEE, the grid cells corresponding to anoxic soil are represented using a map of wetlands from Tootchi et al. (2019). Then the anoxic soil COS flux is expressed as a function of soil temperature following Ogée et al. (2016),

$$F_{\text{COS,anoxic soil}} = P_{\text{ref}} \cdot z_{\text{max}} \cdot Q_{10}^{\left(\frac{T-T_{\text{ref}}}{10}\right)} \quad (6)$$

where $F_{\text{COS,anoxic soil}}$ is the COS flux from anoxic soils ($\text{mol m}^{-2} \text{s}^{-1}$), P_{ref} is the reference production term ($\text{mol m}^{-2} \text{s}^{-1}$), T_{ref} is a reference soil temperature (K) and Q_{10} is the multiplicative factor of the production rate for a 10°C increase in soil temperature (unitless).

A more detailed description of the vegetation and soil COS models implemented in ORCHIDEE can be found in Maignan et al. (2021) and Abadie et al. (2022), respectively.

2.2.4. The ORCHIDEE Data Assimilation System

The ORCHIDEE Data Assimilation System (ORCHIDAS) was designed to optimize ORCHIDEE parameters related to carbon, water, and energy processes. ORCHIDAS has been frequently used and described in detail in previous studies focusing on the assimilation of EC flux data (Bastrikov et al., 2018; Kuppel et al., 2014; MacBean et al., 2022; Mahmud et al., 2021; Peylin et al., 2016). We used ORCHIDAS to find the parameter

values that give the best fit between the observations and the corresponding ORCHIDEE outputs. Assuming that the observations, parameters, and model outputs follow a Gaussian distribution, the optimized parameters are obtained through the minimization of the following cost function $J(x)$ (Tarantola, 2005),

$$J(x) = \frac{1}{2} \cdot [(H(x) - y)^T \cdot R^{-1} \cdot (H(x) - y) + (x + x_b)^T \cdot B^{-1} \cdot (x + x_b)] \quad (7)$$

with x_b the a priori vector of parameters, x the optimized vectors of parameters, y the observations, $H(x)$ the corresponding model outputs, and R and B the prior error covariance matrices for the observations (including measurement and model errors) and the parameters, respectively. R and B are diagonal in this study since the error covariances are difficult to access and hence neglected. The errors (i.e., variances) occupy the diagonal elements in each matrix. The observation errors in R are defined as the mean squared differences between the prior model and the observations, following a classical approach used in the studies listed above where the model error dominates the overall observation error. The parameter uncertainty in B is defined as 15% of the parameter physical range of variation.

We conducted the optimizations with the genetic algorithm (GA) method (Goldberg, 1989; Haupt & Haupt, 2004). This global search method allows us to obtain a combination of optimized parameters without risking getting stuck in a local minimum of the cost function J . Using a population of 32, the algorithm was run for 25 iterations, which was sufficient for the optimization to converge.

Then, at the minimum of the cost function J , the posterior uncertainties can be approximated assuming linearity of the model around the solution and Gaussian prior errors. The reduction in posterior uncertainty for each parameter after optimization can be computed as the difference between the prior and posterior parameter uncertainties, divided by the prior parameter uncertainty. This can inform which parameters are the most well constrained by observations during the optimization.

2.3. Optimization Protocol

2.3.1. Sensitivity Analysis for Parameter Selection

We conducted sensitivity analyses (SAs) prior to performing optimizations to identify the parameters to which simulated GPP and ecosystem COS fluxes are the most sensitive (Figure 1). This enables us to focus only on the key parameters during the optimization, reducing computational cost and the risk of overfitting. We selected the Morris method for SA, which is a time efficient qualitative method that ranks the parameters by importance (Campolongo et al., 2007; Morris, 1991). We considered a large number of parameters for SA, including parameters in the COS flux models and those related to photosynthesis, phenology, conductances, albedo, snow and soil thermal properties, and soil hydrology. A complete list of all parameters can be found in Table S2 in Supporting Information S1.

For soil COS fluxes, we performed an SA at FI-Hyy in 2015 following the method described in Abadie et al. (2022). Then, we conducted SA for simulated GPP and vegetation COS fluxes over one year for each of the 16 BorENF forest sites. Although ecosystem COS flux measurements are only available at FI-Hyy, here, we aimed to identify the key parameters for vegetation COS uptake in the BorENF PFT considering all the studied sites. Indeed, all selected sites correspond to the same vegetation type in ORCHIDEE (BorENF), although soil types vary across sites. The parameters to which simulated soil COS fluxes, vegetation COS uptake, and GPP fluxes are the most sensitive are shown in Figure S1 in Supporting Information S1.

2.3.2. Single Site COS and GPP Optimization Scenarios

In this study, we investigated the potential of COS flux measurements for improving the simulated GPP and LE fluxes in an LSM using the ORCHIDAS optimization framework. Because COS and CO₂ follow a common diffusion pathway during plant uptake, we expect that assimilating vegetation COS flux measurements will affect parameters that also control GPP and LE. However, whereas partitioned GPP data are available at all study sites, measurements of vegetation COS fluxes, including EC fluxes and soil fluxes, are only available at FI-Hyy. Therefore, at this site, we conducted the optimization of ecosystem COS fluxes in two steps.

First, we only assimilated soil COS fluxes using the soil chamber measurements collected in 2015 as previously done in Abadie et al. (2022). Following the results of the first SA experiment, we selected 5 parameters for soil

Table 1
Default Values and Ranges of Variation of the Parameters Included in the Optimization

Parameter name	Description	Default value [range of variation]	Unit	Optimized variables		
				Ecosystem COS flux	GPP	Ecosystem COS flux and GPP
α	Proportionality constant used in the calculation of the internal conductance to COS	0.001200 [0.000720, 0.001680]	$\text{mol } \mu\text{mol}^{-1}$	x		x
$\text{acclim}_{J_{\max,a}}$	Intercept of the linear regression representing the acclimation to temperature of the Entropy term for J_{\max} (the maximum value of the electron transport rate under saturated light) following Kattge and Knorr (2007)	659.7 [494.8, 824.6]	$\text{J K}^{-1} \text{mol}^{-1}$	x	x	x
$\text{acclim}_{V_{\text{cmax},a}}$	Intercept of the linear regression representing the acclimation to temperature of the Entropy term for V_{cmax} following Kattge and Knorr (2007)	668.39 [501.3, 835.5]	$\text{J K}^{-1} \text{mol}^{-1}$	x	x	x
SLA	Specific leaf area, involved in the calculation of leaf biomass	0.00926 [0.00695, 0.01157]	$\text{m}^2 \text{gC}^{-1}$	x	x	x
$V_{\text{cmax}_{25}}$	Maximum rate of Rubisco activity-limited carboxylation at 25°C	45. [34., 56.]	$\mu\text{mol m}^{-2} \text{s}^{-1}$	x	x	x
Leaf_age_crit	Critical leaf age, involved in the calculation of leaf photosynthetic efficiency, and in the calculation of leaf turnover as a function of long term temperature	910. [683., 1138.]	days	x	x	x
A1	Empirical factor involved in the calculation of the function describing the effect of leaf-to-air vapor difference on the stomatal conductance	0.85 [0.8, 0.9]	-	x		x
$g_{\text{b,ref}}$	Leaf bulk boundary layer conductance	0.04 [0.03, 0.05]	m s^{-1}	x	x	x
$T_{\text{photo}_{\min}}$	Minimum photosynthesis temperature	-4. [-5., -3.]	°C	x	x	x
LAI_{\max}	Maximum projected leaf area index, involved in the allocation of carbon, and in the calculation of leaf biomass	4.5 [3.4, 5.6]	$\text{m}^2 \text{m}^{-2}$		x	x

COS flux optimization: F_{CA} , α_{soil} , β_{soil} , the van Genuchten water retention curve coefficient n (Van_Genuchten_{*n*}), and the saturated volumetric water content (θ_{sat}) (Table S2 in Supporting Information S1). F_{CA} , α_{soil} , and β_{soil} are parameters of the soil COS model, with F_{CA} representing the soil microbial community that contains the CA enzyme and can consume COS, while α_{soil} and β_{soil} are involved in the production rate expressed as a function of soil temperature. The van Genuchten water retention curve coefficient n and the saturated volumetric water content both describe soil hydrology, which determines soil water content and COS diffusivity in the soil column.

Then, to optimize parameters related to vegetation COS fluxes, we re-ran ORCHIDEE simulations at FI-Hyy using the optimized set of parameters for soil COS fluxes. The use of optimized soil parameters in the second optimization step enabled us to focus on the parameters to which vegetation COS fluxes are sensitive. Starting from this optimized soil COS flux simulation, we assimilated EC-measured ecosystem COS fluxes over the 5 years of available data (2013–2017). Based on vegetation COS flux SA, we considered the 9 most important parameters for the optimization of ecosystem COS fluxes. The default values and ranges of variation for the parameters included in each optimization are given in Table 1. We defined the physical range of variations based on expert and physical knowledge for each parameter. Note that for the α parameter, we considered a large range of variation corresponding to $\pm 40\%$ of the prior value as its value is not very well constrained by the linear regression applied in Berry et al. (2013) and as was shown based on site observations by Kooijmans et al. (2021).

We compared the data assimilation of ecosystem COS flux to a standard optimization in which we assimilated only GPP data at FI-Hyy over the last 5 years of available FLUXNET GPP estimates (2010–2014) (Figure 1).

Finally, to investigate the additional constraint provided by COS observations on top of that provided by GPP, we performed an optimization with both GPP and ecosystem COS fluxes (Figure 1). This optimization was carried out by assimilating GPP for 5 years (2010–2014) with COS data also for 5 years (2013–2017). This joint assimilation aims at finding a combination of optimized parameters that minimizes the misfit for both GPP and ecosystem COS fluxes even if the assimilation periods differ. Because there are data gaps in the ecosystem COS

Table 2

Summary of the 3 Multi-Site Optimization Scenarios, With the Weight Given to the GPP and Ecosystem COS Flux Data in the Assimilation for Each Scenario

	Post GPP	Post GPP & $F_{\text{COS}}^{1/2}$	Post GPP & $F_{\text{COS}}^{1/3}$
GPP (10 BorENF sites)	1	$1/2$	$2/3$
Ecosystem COS fluxes (FI-Hyy)	0	$1/2$	$1/3$

Note. The period considered for the assimilation and evaluation at each site are presented in Table S1 in Supporting Information S1.

flux measurements, we applied a weighting factor to the GPP term in the cost function (Equation 7) in order to give the same weight to COS and GPP in the optimization. This weight was computed as the ratio of the number of ecosystem COS flux observations and the number of GPP estimates over the 5-year period considered for each. In this work we chose not to assimilate LE observations as evapotranspiration includes not only plant transpiration, on which COS offers a constraint through the gas diffusion pathway, but also snow sublimation, water interception by and evaporation from the canopy or ground vegetation, and soil evaporation, none of which is physiologically linked to COS uptake through stomata. Note that at this site, bare soil evaporation is marginal as the soil is mainly covered by mosses.

For the data assimilation experiments, data were averaged at a daily time step and smoothed over a 7-day running mean. A few net ecosystem COS emissions are found in the EC observations, especially in April, August, and September 2014. These net emissions may be due to data noise, or possibly reflect some soil emission episodes at high temperatures at the end of the summer. However, in ORCHIDEE, the annual mean simulated soil COS flux is about $-3 \text{ pmol m}^{-2} \text{ s}^{-1}$ with no strong seasonal variations, which is in line with the average soil COS fluxes measured in the two soil chambers in 2015 (-2.8 ± 1.0 and $-2.5 \pm 1.2 \text{ pmol m}^{-2} \text{ s}^{-1}$, Sun et al., 2018). Therefore, we filtered the EC COS fluxes to remove the data for which the mean daily COS fluxes was higher than $-3 \text{ pmol m}^{-2} \text{ s}^{-1}$, as such values cannot be reached at FI-Hyy by the COS models implemented in ORCHIDEE.

2.3.3. Multi-Site COS and GPP Optimization Scenarios

To evaluate the additional constraint on GPP and LE from COS across the entire BorENF biome, we also performed multi-site optimizations using only GPP data or combining GPP and COS flux data (Figure 1). In a multi-site optimization, one common set of parameters is obtained by simultaneously optimizing over all sites. Ten of the 16 BorENF sites presented in Section 2.1.1 were selected because at least 4 years of GPP and LE measurements are available (Table S1 in Supporting Information S1). Among these 4 years, we used 3 years for data assimilation and the last year was used for evaluating the optimization with a full independent seasonal cycle.

We performed a first multi-site optimization assimilating only GPP data at these 10 sites (“Post GPP”). Then, in a second multi-site optimization, we assimilated ecosystem COS measurements at FI-Hyy in addition to GPP at the 10 BorENF sites. In this scenario, we aimed at including as much information from COS observations as from GPP estimates in the assimilation. However, as we only have COS observations at one site for this multi-site optimization, we applied a weighting factor to the COS term in the cost function (Equation 7) so that COS observations would have the same weight as GPP observations at the 10 sites combined (scenario “Post GPP & $F_{\text{COS}}^{1/2}$ ”). Then, we tested another scenario in which we arbitrarily downweighted the importance of COS data by adjusting the multiplier before the COS cost term to $1/3$ (scenario “Post GPP & $F_{\text{COS}}^{1/3}$ ”; Table 2) because COS measurements are fewer than GPP data and are available at only one site. In both cases, because we intended to use the FI-Hyy site as an additional site to constrain the COS fluxes and COS-related parameters only, we did not use it for GPP data assimilation or the evaluation of optimizations results.

For each of the multi-site optimizations, we considered the same set of parameters as for the single site optimization at FI-Hyy for GPP only or joint GPP and COS assimilation (Table 1). In addition to the 10 sites used in the multi-site optimizations, we used 5 other BorENF sites for which 1 year of GPP and LE measurements is available to further evaluate the impact of optimizations on model performance at independent sites.

3. Results

3.1. Impact of COS and GPP Assimilations on Carbon and Water Fluxes at Hyytiälä Forest, Finland

3.1.1. Constraint on COS and GPP Seasonal Cycles

First, we assessed the impacts of COS flux-only, GPP-only, or a joint COS flux-GPP assimilation on ecosystem COS fluxes and GPP at FI-Hyy. Figure 3 shows the change in the simulated mean seasonal cycle for the ecosystem COS flux over the 5 years of assimilation (Figure 3a) and for GPP (Figure 3b) over 8 independent years.

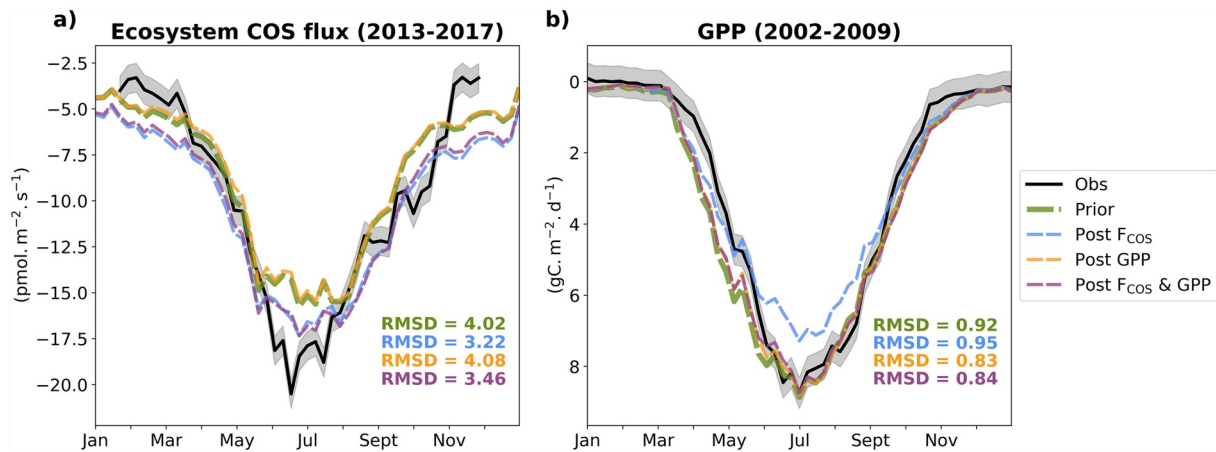


Figure 3. Mean seasonal cycles of the weekly ecosystem COS flux ($\text{pmol m}^{-2} \text{s}^{-1}$) (a) and GPP ($\text{gC m}^{-2} \text{d}^{-1}$) (b) at FI-Hyy. The weekly averages are computed considering both nighttime and daytime data for the eddy covariance observations and simulated fluxes. Note that the y-axis has been inverted for GPP (with largest positive values on the bottom, and decreasing moving up) to visually aid the comparison with vegetation COS uptake. The observations (“Obs”) are represented in black with the shaded area representing the standard error of the mean, computed as the standard deviation divided by the square root of the number of observations. The ORCHIDEE simulation prior to the assimilation is in green (“Prior”), and the ORCHIDEE simulations after optimization are represented in blue for the ecosystem COS flux assimilation only (“Post COS”), in orange for the GPP assimilation only (“Post GPP”), and in purple for the assimilation of both the ecosystem COS flux and GPP (“Post F_{COS} & GPP”). The mean seasonal cycles of GPP are averaged over 8 years (2002–2009), and no data during this period were assimilated (the period considered for assimilation is 2010–2014) in order to independently evaluate the posterior seasonal cycles of GPP. Note that GPP seasonal cycles for the Prior, Post GPP, and Post F_{COS} & GPP simulations track each other closely during the fall decrease. For COS, the mean seasonal cycle is computed over the available period for the observations (2013–2017), the same period as for COS data assimilation. The RMSD values computed over the weekly fluxes are given for the prior and the post-optimization simulations in corresponding colors.

The two optimizations that assimilate COS flux observations (“Post F_{COS}” and “Post F_{COS} & GPP”) lead to a similar mean seasonal cycle with a higher net ecosystem COS uptake compared to the prior simulation (mean ecosystem COS flux of $-10 \text{ pmol m}^{-2} \text{ s}^{-1}$ over the period against $-8.8 \text{ pmol m}^{-2} \text{ s}^{-1}$ for the prior). This increase in COS uptake reduces the RMSD by 20% and 14% for the COS-only assimilation and the joint assimilation of F_{COS} and GPP, respectively. The mean seasonal amplitude (difference between the annual maximum and minimum of weekly mean net uptake) of ecosystem COS flux is still underestimated compared to the observations (observed amplitude of $17.2 \text{ pmol m}^{-2} \text{ s}^{-1}$), but is slightly improved by the optimizations that assimilate COS flux observations (amplitudes near $12.5 \text{ pmol m}^{-2} \text{ s}^{-1}$ against $11.7 \text{ pmol m}^{-2} \text{ s}^{-1}$ after the GPP-only assimilation and $11.9 \text{ pmol m}^{-2} \text{ s}^{-1}$ for the prior).

As expected, assimilating GPP observations only has little impact on ecosystem COS fluxes seasonal cycle. Concerning the GPP seasonal cycle, the prior simulation is already in good agreement with the estimates from flux partitioning (RMSD = $0.92 \text{ gC m}^{-2} \text{ d}^{-1}$) at FI-Hyy. However, assimilating GPP observations only or GPP and COS observations together further improves the RMSD, with a reduction of about 10% over 8 independent years. Note that for these two optimization scenarios, a similar reduction in RMSD (8%) is found over the period considered for GPP assimilation (2010–2014). On the other hand, assimilating only COS data leads to a degradation of the seasonal cycle with an underestimation of GPP seasonal amplitude of about $2 \text{ gC m}^{-2} \text{ d}^{-1}$ compared to the observations.

When including COS data in the assimilations, the maximum net COS uptake in 2015 and 2016 is similar to the observed one ($-19.4 \text{ pmol m}^{-2} \text{ s}^{-1}$ in 2015, $-21.9 \text{ pmol m}^{-2} \text{ s}^{-1}$ in 2016) while it was underestimated in the prior simulation ($-17.6 \text{ pmol m}^{-2} \text{ s}^{-1}$ in 2015, $-19.5 \text{ pmol m}^{-2} \text{ s}^{-1}$ in 2016) (Figure S2 in Supporting Information S1). The years 2014 and 2015 show the lowest ecosystem COS uptakes over the observation period, which have been related to a high VPD in July 2014 and a low soil water content in April 2015 (Vesala et al., 2022). On the contrary, the two optimizations assimilating COS observations do not capture the strongest ecosystem COS uptakes observed in 2013 and 2017, for which the model still underestimates the seasonal amplitude.

3.1.2. Constraint on Leaf Gas Exchange Parameters

To understand how the optimizations impact the modeled COS and GPP fluxes at FI-Hyy, we studied the changes in the simulated stomatal and internal conductance to COS (Figure 4). As the boundary layer conductance is an

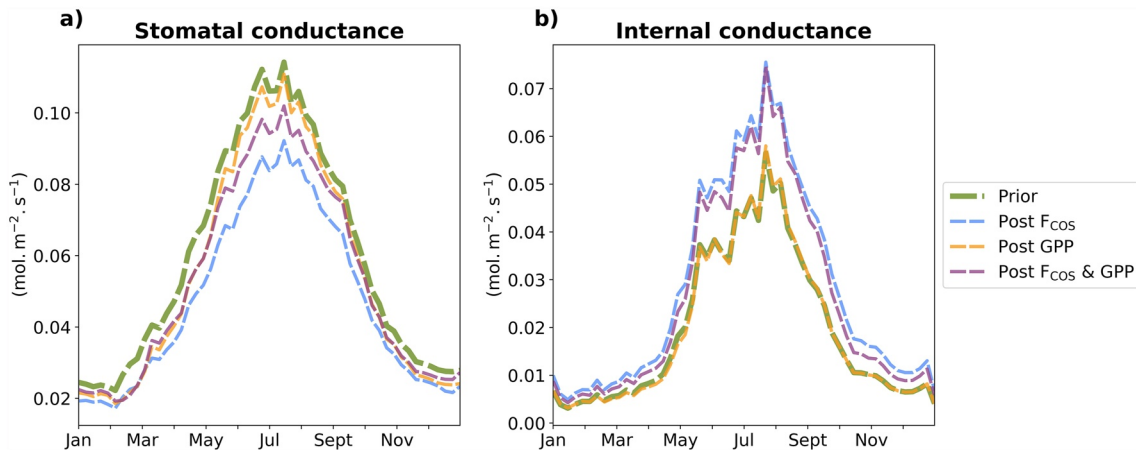


Figure 4. Mean seasonal cycles of the weekly modeled stomatal (a), and internal (b) conductances to COS ($\text{mol m}^{-2} \text{s}^{-1}$) at FI-Hyy between 2013 and 2017. Note that the posterior seasonal cycles of the internal conductance in the “Prior” and “Post F_{COS} ” simulations overlap with each other.

order of magnitude higher than the stomatal and the internal conductances, we do not focus on this conductance as it is not the main limiting factor for gas diffusion.

The assimilation of COS only or both COS and GPP observations leads to a decrease in the stomatal conductance to COS compared to the prior simulation (Figure 4a). This decrease is driven by a lower value of the A1 parameter after optimization (Figure S3 in Supporting Information S1), which leads to lower stomatal conductance under the same VPD conditions (Equation 2). Because the ratio between stomatal conductance to CO_2 and that to COS is a constant of 1.21 (Seibt et al., 2010), stomatal conductance to CO_2 is proportionally affected by the change in A1. Similarly, assimilating only GPP observations also reduces stomatal conductance, though to a lesser extent. Therefore, assimilating COS instead of or in combination with GPP data provides a stronger constraint on stomatal diffusion.

Finally, the two optimizations that assimilate COS data both increase the internal conductance to COS compared to the prior (Figure 4b). The increase in $V_{\text{cmax}_{25}}$ and α parameters (Figure S3 in Supporting Information S1) explains this higher internal conductance as it relates to V_{cmax} by the multiplicative factor α (Equation 4).

In addition to analyzing the changes in parameter values after optimization, evaluating the reduction in uncertainty informs on which parameters are the most constrained after optimization (see Section 2.2.4). Across all three optimizations, the parameters showing a high percentage of reduction in uncertainties are $V_{\text{cmax}_{25}}$ (66%–85% reduction) and $\text{acclim}_{V_{\text{cmax},a}}$ (the intercept of the linear regression representing the acclimation to temperature of the entropy term for V_{cmax}) (95%–98% reduction). The uncertainty on α is reduced by 27% in the COS-only data assimilation and 82% in the joint assimilation of both COS and GPP data.

Then, the uncertainty in Leaf_age_crit (critical leaf age) also shows a strong reduction across all optimizations (49%–85% reduction), as well as $g_{\text{b,ref}}$ (the leaf bulk boundary layer conductance) for the assimilations including COS data (75%–88% reduction). Finally, the lowest reduction in parameter uncertainties are found for $T_{\text{photo_min}}$ (the minimum photosynthesis temperature) and LAI_{max} (the maximum LAI), both of which experience less than 1% reduction in uncertainty.

3.1.3. Impact on LE and WUE

Considering the crucial role of stomatal in controlling in the coupled plant carbon and water fluxes, we evaluated the impact of optimizing the GPP- and COS-related parameters on LE (Figure 5a). While assimilating only GPP data has little impact on the simulated LE mean seasonal cycle, the two optimizations assimilating COS observations reduce the LE seasonal amplitude. This reduction is caused by a decrease in stomatal conductance after the optimizations, which results from the COS constraint on stomatal diffusion, as noted previously (Figure 4a). The optimizations also reduce the RMSD in LE by about 20% (from 10.22 W m^{-2} for the prior to 8.22 W m^{-2} and 8.48 W m^{-2} for “Post F_{COS} ” and “Post $F_{\text{COS}} \& \text{GPP}$,” respectively). Note that when focusing on the summer months (from May to August), the RMSD is reduced by about 35% (9.11 W m^{-2} for “Post $F_{\text{COS}} \& \text{GPP}$ ”) to 45% (8.37

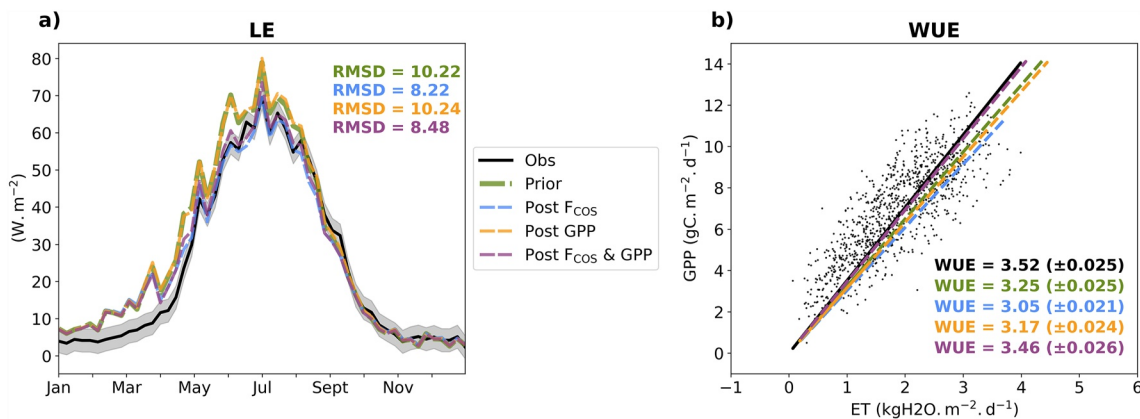


Figure 5. Mean seasonal cycle of the weekly observed and modeled LE ($W m^{-2}$) over 2002–2009 (a) and summer (May to August) WUE ($gC kgH_2O^{-1}$) (b) at FI-Hyy. For LE, the RMSD values computed over the weekly fluxes are given for the prior and the post-optimization simulations in the corresponding color (see legend). WUE ($gC kgH_2O^{-1}$) is computed as the slope of the linear regression between GPP ($gC m^2 d^{-1}$) and ET ($kgH_2O m^2 d^{-1}$), presented in the corresponding color with the associated standard error. The black dots in the scatter plot represent daily averages for the eddy covariance data.

$W m^{-2}$ for “Post F_{cos} ”) compared to the prior simulation RMSD ($12.39 W m^{-2}$). However, the simulated LE is still overestimated from January to April, with a mean bias of $6.24 W m^{-2}$. As the simulated plant transpiration is zero in winter and increases in April with the start of the growing season, the overestimation in LE at the beginning of the year cannot be attributed to the transpiration flux but to other components of LE (see Section 4.1). Overall, because no LE data were assimilated, the post-optimization results highlight the potential of using COS flux observations to constrain LE.

Finally, we examined the coupling between GPP and ET as indicated in the ecosystem water use efficiency (WUE), the rate of carbon uptake per unit of water loss. We computed the WUE at FI-Hyy as the coefficient of the linear regression between GPP and ET (Figure 5b). Model estimates of WUE in the prior simulation and the simulation that assimilates only GPP data are $3.25 (\pm 0.025) gC kgH_2O^{-1}$ and $3.17 (\pm 0.024) gC kgH_2O^{-1}$, respectively, which are below the observationally derived WUE of $3.52 (\pm 0.025) gC kgH_2O^{-1}$. This underestimation can be related to the higher LE in summer, which is not well captured by these simulations (Figure 5a). Then, while assimilating only COS observations further degrades the WUE ($3.05 (\pm 0.025) gC kgH_2O^{-1}$), the assimilation of COS and GPP together leads to a WUE close to the observationally inferred value of $3.46 (\pm 0.026) gC kgH_2O^{-1}$.

3.1.4. Focus on a Severe Drought Event at Hyytiälä

We focused on the period characterized by very dry conditions at FI-Hyy (SMI < 0.20 from 23 July to 28 August 2006, see Section 2.1.3) to study the impact of our three optimization scenarios on the simulated GPP, LE and vegetation COS fluxes during this severe drought event.

Figure 6a shows the response of the simulated and FLUXNET GPP during the drought event at FI-Hyy in 2006. Just before the beginning of the most severe phase of the drought, the FLUXNET GPP depicts a sharp decrease of about $5 gC m^2 d^{-1}$, but stabilizes after DOY 218 under high soil moisture stress (SMI < 0.15). After assimilating COS observations, the optimizations lead to a decrease in the simulated GPP compared to the prior simulation, reducing the mismatch with the observations during the most severe part of the drought. However, the decrease in the simulated GPP following the progression of the drought is less abrupt than that in the observations. The simulated GPP reaches its minimum just before the end of the most severe phase of the drought (DOY 228), 10 days after when the FLUXNET GPP reaches its minimum. After the most severe phase of the drought, observed GPP partially recovers (DOY 231). However, all simulations strongly underestimate GPP in this recovery phase (see Section 4.1).

Contrary to the variations of GPP during this drought event, the observed variations in LE are better captured by the modeled fluxes with a strong decrease in LE from the second day of the very dry conditions (DOY 206) to the end of the most severe drought period (Figure 6b). Note that the decrease in observed LE is stronger than that simulated in ORCHIDEE. However, in the two scenarios with COS data assimilation, the decrease in LE is

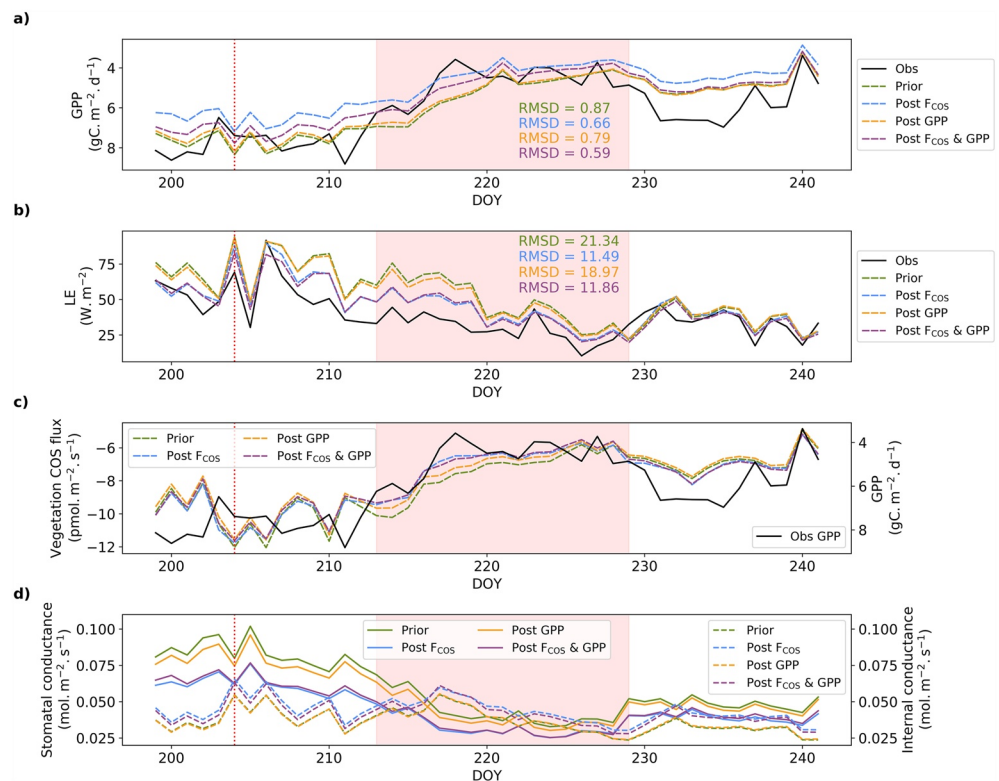


Figure 6. Daily averages of observed and modeled GPP ($\text{gC m}^{-2} \text{d}^{-1}$) (a) and LE (W m^{-2}) (b) at FI-Hyy during a severe drought event in 2006. Note that the y-axis has been inverted for GPP (with largest positive values on the bottom, and decreasing moving up) to facilitate the comparison with vegetation COS uptake. The comparison between the modeled vegetation COS fluxes ($\text{pmol m}^{-2} \text{s}^{-1}$) and the observed GPP is presented (c), as well as the simulated stomatal and internal conductances (right vertical axis) to COS (d). The red vertical line indicates the start of the drought event (DOY 204) after which SMI was lower than 0.20. The 17 consecutive days (1–17 August, DOY 213–229) with an observed SMI lower than 0.15 are indicated by the shaded area. The daily RMSD values are given for the prior and the post optimization simulations over these 17 days of severe drought.

higher than in the prior simulation or the simulation that only assimilates GPP data, especially in the two weeks of LE decrease (DOY 206–220).

Simulated vegetation COS uptake closely follows the decreasing trend in the simulated GPP until the end of the most intense phase of the drought across all assimilation scenarios (Figure 6c). However, for vegetation COS uptake, the impact of the different assimilation scenarios is mainly concentrated at the beginning of the most severe part of the drought. Indeed, when assimilating COS, the decrease in vegetation uptake during this period is stronger than the decrease in the GPP-only assimilation or that in the prior simulation. This faster reduction in vegetation COS uptake better tracks the response of the observed GPP to water stress than does the simulated GPP.

Finally, we studied the variations in the simulated stomatal and internal conductance to COS during this drought event for the three optimization scenarios (Figure 6d). As previously seen for the mean seasonal cycles (Figure 4), the assimilations of COS data reduce the simulated stomatal conductance and increase the simulated internal conductance to COS. While the internal conductance to COS does not show a sharp decreasing trend during the drought event, the stomatal conductance is halved from the beginning of the very dry conditions to the end of the most intense phase of the drought. Therefore, during the period with particularly high soil moisture stress ($\text{SMI} < 0.15$), the stomatal conductance shows lower values than the internal conductance and becomes the most limiting factor for vegetation COS uptake. Note that these conductance responses are found in ORCHIDEE simulations, though there are no observations to validate these responses during this drought event. This change in the limiting conductance is more pronounced when assimilating COS observations, because the difference

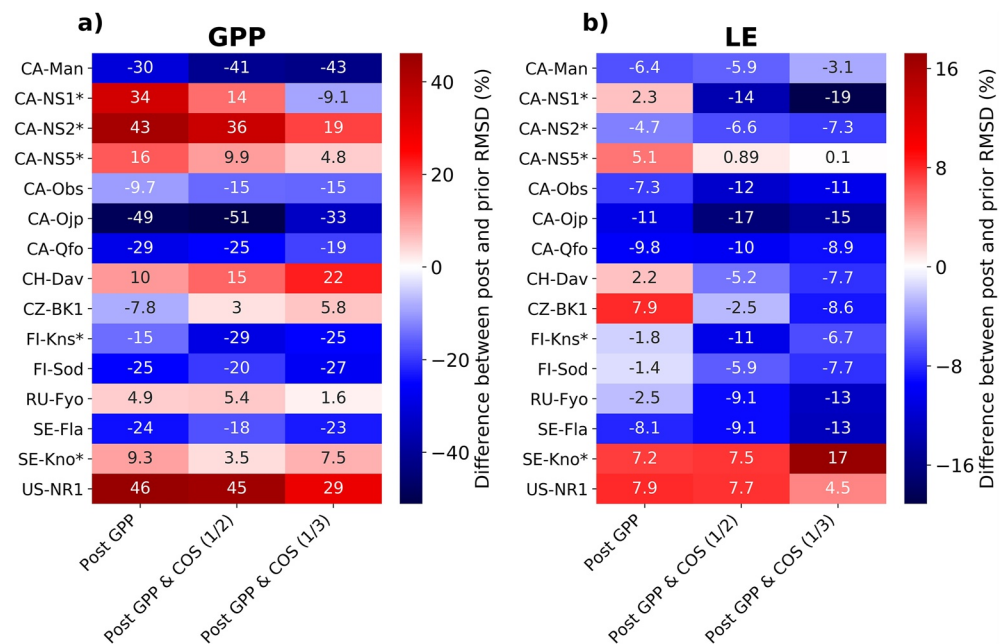


Figure 7. Relative difference (%) in RMSDs of weekly GPP (a) and LE (b) between prior simulations and after multi-site optimizations at each of the 15 validation sites. A negative difference (in blue) means a decrease in RMSD after multi-site optimization. In all multi-site optimization scenarios, GPP data from 10 BorENF sites (not marked with a star), each with an observational record of at least 4 years, were assimilated. For each site in this group, three out of the 4 years of data were assimilated, and the rest was used for evaluation. In addition, 5 other BorENF sites (marked with a star), each of which provides 1 year of GPP data, were used for evaluation but not data assimilation.

between the internal conductance and the stomatal conductance is higher in these optimization scenarios than in the prior simulation or the optimization that assimilates only GPP data.

3.2. Evaluation of COS and GPP Multi-Site Assimilations at Several Boreal Evergreen Needleleaf Forest Sites

After evaluating the potential of COS to constrain GPP and LE at FI-Hyy, we investigated the additional constraint provided by COS measurements at FI-Hyy on multiple BorENF sites, as compared to the constraint provided solely by GPP data at the same sites. Figure 7 presents the improvement or degradation of RMSD in the 3 multi-site optimization scenarios compared to the prior RMSD at each of the 15 validation sites, of which 10 were used for the assimilations (see Section 2.3.3).

First, we find that the multi-site optimization that assimilates only GPP data at the 10 assimilation sites (“Post GPP”) deteriorates the simulated GPP at 7 of the 15 validation sites (CA-NS1, CA-NS2, CA-NS5, CH-Dav, RU-Fyo, SE-Kno, and US-NR1). At these sites, prior simulations underestimate GPP compared to the observations (Figure S4 in Supporting Information S1).

Assimilating COS observations at FI-Hyy in addition to GPP at 10 BorENF sites (“Post GPP & $F_{\text{COS}}^{1/2}$ ” and “Post GPP & $F_{\text{COS}}^{1/3}$ ”) also increases the RMSD of GPP at 5 sites (CA-NS1, CA-NS2, CA-NS5, SE-Kno, and US-NR1), but the degradation of GPP performance is not as severe as after GPP-only data assimilation (“Post GPP”). This is mainly because assimilating both COS and GPP data leads to a stronger seasonal amplitude at these 5 sites compared to assimilating only GPP data.

For sites at which the prior GPP is underestimated, different changes in parameter values may explain the milder degradation of GPP performance for the optimizations that assimilated COS data than for the optimization that assimilates only GPP data. Contrary to assimilating both COS and GPP data, the optimization that assimilates only GPP data reduces three GPP-related parameters: $\text{acclim}_{\text{max,a}}$, specific leaf area (SLA), and LAI_{max} (Figure S5 in Supporting Information S1; see Table 1 for parameter definitions). In particular, because the ratio between LAI_{max} and SLA determines the maximum leaf biomass, a decrease in this ratio after assimilating only GPP data

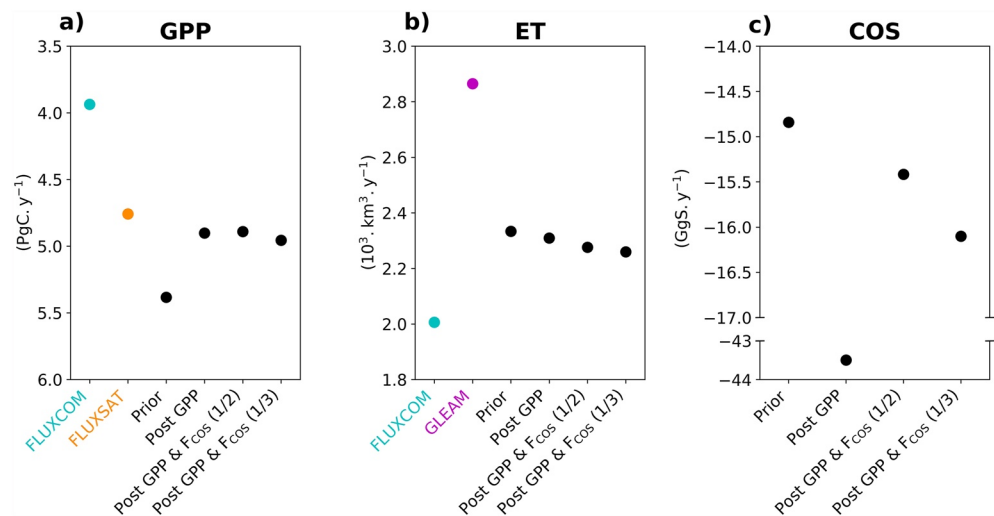


Figure 8. GPP (a), ET (b), and vegetation COS fluxes (c) in the prior simulation and each of the multi-site optimization scenarios computed over land grid cells that have a fractional coverage of BorENF higher than 40%, averaged between 2006 and 2015. For GPP and ET, the fluxes from the evaluation products (FLUXCOM, FLUXSAT, GLEAM) were shown in colors. Note that the y-axis has been inverted for GPP (with largest positive values on the bottom, and decreasing moving up) so that both GPP and vegetation COS fluxes that represent net uptake are in the same direction.

leads to a lower maximum leaf biomass. In contrast, this ratio increases after the two optimizations that assimilate COS data, leading to a higher maximum leaf biomass.

At the 8 sites where GPP performance was improved after the GPP-only multi-site optimization, the prior simulations overestimate GPP. Note that similar to the single-site optimization at FI-Hyy (Figure S3 in Supporting Information S1), the multi-site optimizations that assimilate COS data also lead to a decrease in A1 (Figure S5 in Supporting Information S1), which reduces stomatal conductance at the same VPD (Equation 2). Then, all sites considered, the mean improvement of GPP after optimization, as measured by the relative difference in RMSD, is about 2% when assimilating only GPP data, and 4% and 7% when assimilating both COS and GPP data with a weight of 1/2 or 1/3, respectively.

Similar patterns are found for LE which experiences performance degradation at 6 sites (CA-NS1, CA-NS5, CH-Dav, CZ-BK1, SE-Kno, and US-NR1) after assimilating only GPP data (“Post GPP”). In contrast, assimilating COS and GPP data (“Post GPP & F_{COS} 1/2” and “Post GPP & F_{COS} 1/3”) mitigates the performance degradation or even reduces the RMSD compared to the prior (except SE-Kno). The prior simulations overestimate LE compared to the observations at most sites, and assimilating COS data in addition to GPP data helps to reduce this overestimation (Figure S6 in Supporting Information S1). Overall, assimilating only GPP data has little impact on the mean relative change in RMSD (1%), while a joint assimilation of GPP and COS data reduces the mean RMSD by 6% considering a weight of 1/2 for COS observations and by 7% with a weight of 1/3 for COS observations.

For both GPP and LE, assimilating GPP and COS observations using a weight of 1/3 for the latter yields post-optimization GPP and LE that are closer to the observations than using a weight of 1/2 for COS observations in the assimilation. This can be due to the larger decrease in A1 and increase in V_{cmax25} for the joint assimilation with a weight of 1/3 for COS data compared to the assimilation using a weight of 1/2 for COS data. At US-NR1, the model struggles to represent both GPP and LE and the optimizations fail to improve the simulated fluxes. This site is located at a high altitude (3,050 m, Table S1 in Supporting Information S1) and this specific condition may have contributed to the inaccurate representation of GPP and LE.

3.3. COS and GPP Multi-Site Assimilation Upscaling to Boreal High Latitudes

The three optimized sets of parameters for BorENF from the multi-site assimilation scenarios were used to run simulations over the whole boreal region (40°–80°N) (Figure S7 in Supporting Information S1). Figure 8 presents

the changes in GPP, ET, and COS fluxes over land grid cells that have a fractional coverage of BorENF higher than 40%.

In ORCHIDEE, the prior simulation overestimates GPP compared to FLUXCOM and FLUXSAT products over the studied areas. The three optimizations lead to similar reductions of 9% in GPP compared to this prior simulation, and help to reduce the mismatch with both evaluation products. This decrease in GPP mainly occurs between the beginning of the growing season and the end of July (Figure S8 in Supporting Information S1). GPP estimates from all three optimizations are close to the FLUXSAT GPP estimate (4.76 PgC yr^{-1}), but exceed that of FLUXCOM (3.94 PgC yr^{-1}).

The prior estimate of ET is between those of FLUXCOM ($2.01 \cdot 10^3 \text{ km}^3 \text{ yr}^{-1}$) and GLEAM ($2.86 \cdot 10^3 \text{ km}^3 \text{ yr}^{-1}$) and is only slightly impacted by the optimizations. Note that over land grid cells that have a fractional coverage of BorENF higher than 40%, the GLEAM ET is 30% higher than FLUXCOM ET, but the spatial distribution of ET also differs between these products (Figure S9 in Supporting Information S1).

The two optimizations that assimilate COS data lead to a stronger vegetation COS sink with an increase of 5% (“Post F_{COS} & GPP $\frac{1}{2}$ ”) and 9% (“Post F_{COS} & GPP $\frac{1}{3}$ ”) compared to the prior. Surprisingly, the optimization that assimilates only GPP data leads to a substantial increase in vegetation COS uptake to $43.50 \text{ GgS yr}^{-1}$ over the whole study area, which is about 3 times the prior estimate.

Finally, spatial distribution of the vegetation COS uptake, GPP, and LE does not differ significantly between the three optimization scenarios and the prior (not shown).

We also investigated how the post-optimization changes in GPP and ET affect WUE over the BorENF biome (Figure S10 in Supporting Information S1). ORCHIDEE overestimates WUE in the prior simulation ($2.32 \pm 0.008 \text{ gC kgH}_2\text{O}^{-1}$) compared with FLUXCOM WUE ($2.0 \pm 0.007 \text{ gC kgH}_2\text{O}^{-1}$). All optimizations reduce WUE as well as the difference in WUE between ORCHIDEE simulations and FLUXCOM. The largest reduction in WUE occurs for the optimization that assimilates only GPP data ($2.14 \pm 0.009 \text{ gC kgH}_2\text{O}^{-1}$ and 8% reduction compared with the prior estimate).

4. Discussion

4.1. COS Data Assimilation Informs Biospheric Processes Represented in ORCHIDEE

Because vegetation COS uptake is more sensitive to stomatal conductance than is GPP, assimilating ecosystem COS flux data in models provides a more robust constraint on parameters that govern stomatal diffusion than does assimilating GPP data. Indeed, we find that the simulated GPP is weakly sensitive to the conductance-related parameters according to a multi-site sensitivity analysis (Figure S1 in Supporting Information S1). This could be due to infrequent high VPD conditions over the BorENF biome, the influence of which is represented by the parameter A1 (Equation 1). Therefore, for this biome, GPP is more sensitive to other photosynthesis-related parameters, particularly those that determine the light reactions of photosynthesis (e.g., $\text{acclim}_{\text{Jmax,a}}$).

In addition to constraining stomatal conductance, COS data assimilation also highlights the need to improve the model-represented response of GPP to drought, especially during the most severe phase of the drought and the recovery phase that follows. In ORCHIDEE, the response of the simulated GPP to soil moisture stress depends on a soil water stress factor, which linearly varies between 0 at the wilting point and 1 for soil moisture close to the field capacity (Section 2.2.1). The inability of the model to capture the observed decrease in GPP at the beginning of the most severe phase of the drought at FI-Hyy, even after optimization (Figure 6), indicates that uncertainty in the response of GPP to drought is dominated by structural uncertainty associated with the functional form rather than parametric uncertainty. That is to say, a linear response to soil moisture stress cannot accurately represent drought impacts on GPP and COS fluxes.

Moreover, a comparison between the surface soil moisture measured at FI-Hyy and that simulated in ORCHIDEE highlights an underestimation of the simulated soil moisture at the onset of drought recovery (Figure S11 in Supporting Information S1). The lower simulated soil moisture translates to a stronger water stress on GPP, which explains the slower recovery of the simulated GPP than in the observations. In addition, the highest volumetric soil moisture values simulated in ORCHIDEE during the recovery phase (18%–20%) are associated with a GPP

of about $4.5 \text{ gC m}^2 \text{ d}^{-1}$, which is lower than the FLUXNET GPP of $6.6 \text{ gC m}^2 \text{ d}^{-1}$ at similar soil moisture values (Figure S12 in Supporting Information S1).

Despite the potential of COS to constrain plant carbon and water exchanges, a joint assimilation of COS and GPP data seems necessary to improve the seasonal cycle of COS fluxes without risking degrading GPP at the same time (Figure 3). Indeed, the degradation in GPP seasonal cycle after assimilating only COS data can be explained by error compensation between $V_{\text{cmax}_{25}}$ and α . When both $V_{\text{cmax}_{25}}$ and α need to be increased, the optimization that assimilates only COS observations increases α more than $V_{\text{cmax}_{25}}$ ($V_{\text{cmax}_{25}} = 46.6 \mu\text{mol m}^2 \text{ s}^{-1}$ and $\alpha = 0.0016$, Figure S3 in Supporting Information S1), which reduces errors in COS uptake at the expense of GPP performance. In contrast, the optimization that assimilates both COS and GPP observations increases $V_{\text{cmax}_{25}}$ more than α ($V_{\text{cmax}_{25}} = 54 \mu\text{mol m}^2 \text{ s}^{-1}$ and $\alpha = 0.0014$), which reduces errors in both COS uptake and GPP. Therefore, the stronger increase in α needed to reproduce observed vegetation COS uptake will not affect GPP, whereas the weaker increase in $V_{\text{cmax}_{25}}$ needed to reproduce GPP cannot compensate for the decrease in the stomatal conductance resulting from the decrease in A_1 (Figure S3 in Supporting Information S1). The joint data assimilation is a delicate balancing act between the two competing needs.

Multi-site GPP-only assimilation leads to an unconstrained and unrealistically large increase in vegetation COS sink over the boreal forest biome (Figure 8). This large increase in vegetation COS uptake seems to be driven by increases in the leaf boundary layer conductance ($g_{\text{b,ref}}$) and in the parameter that represents thermal acclimation of V_{cmax} ($\text{acclim}_{V_{\text{cmax},a}}$), because such increases are absent in the optimization that assimilates both GPP and COS data (Figure S5 in Supporting Information S1). This is because the simulated vegetation COS uptake is particularly sensitive to $\text{acclim}_{V_{\text{cmax},a}}$ which affects the internal conductance to COS through V_{cmax} according to the Berry et al. (2013) parameterization (Figure S1 in Supporting Information S1). Thus, this overly sensitive behavior of vegetation COS uptake highlights a need for a mechanistic representation of the internal conductance to COS beyond what an empirical scaling factor α can offer.

Finally, this study highlights other deficiencies in the representation of some processes in ORCHIDEE, as illustrated by the overestimation of the simulated LE from January to April at FI-Hyy (Figure 5), which cannot be corrected by data assimilation. The culprit seems to be snow sublimation, which shows strong peaks early in the year for several years (Figure S13 in Supporting Information S1). These peaks coincide with those of the simulated LE, such as the peaks in 2005 or 2008 between January and March. Therefore, it is likely that the overestimation of the simulated LE at this time of the year originates from the snow sublimation component, which cannot be mitigated by assimilating GPP and COS flux observations. Note that the earlier onset of the simulated GPP compared to the FLUXNET GPP could also result from a misrepresentation in the snow processes (Figure 3).

4.2. BorENF COS and GPP Fluxes and Related LRU

All studies focusing on the BorENF PFT find an increase in vegetation COS uptake after optimization (Figure 9), which is in line with a suspected high-latitude missing sink of COS proposed in recent inversion studies (Hu et al., 2021; Kuai et al., 2022; Ma et al., 2021; Remaud et al., 2022). Remaud et al. (2022) optimized the COS and CO_2 gross fluxes simulated in ORCHIDEE against atmospheric concentrations from the NOAA/ESRL/GML and provided prior and posterior values for GPP and vegetation COS uptake for each PFT. While their optimization led to an increase of 20% in the vegetation COS uptake for BorENF, this study found a lower increase up to 10% (“Post F_{COS} & GPP $\frac{1}{3}$ ”). It is to be noted that their prior vegetation COS fluxes were computed using the LRU approach with the LRU values from Whelan et al. (2018), which gives a prior vegetation COS uptake for BorENF 80% larger than the prior in this study (32.3 GgS yr^{-1} against $17.78 \text{ GgS yr}^{-1}$). In their inversion framework, the errors from all boreal PFT fluxes (BorENF, Boreal Broadleaf Summergreen, and Boreal Needleleaf Deciduous) were correlated, which does not allow to strictly optimize the fluxes per PFT separately. On the contrary, the prior GPP simulated in their ORCHIDEE version (5.8 PgC yr^{-1}) is 10% lower than the one computed in the version used in this study (6.36 PgC yr^{-1}). In another inversion study, Hu et al. (2021) derived the GPP over the North American Arctic and boreal region using atmospheric COS concentrations from the NOAA/ESRL and the LRU approach. Over this region, the atmospheric COS inversion led to a vegetation COS uptake 40% higher than the prior simulated in the SiB4 LSM. Finally, to investigate BorENF contribution to the missing COS sink from a bottom-up approach, Vesala et al. (2022) developed a parametric model based on environmental drivers (photo-synthetically active radiation, vapor pressure deficit, air temperature, and leaf area index) to simulate vegetation

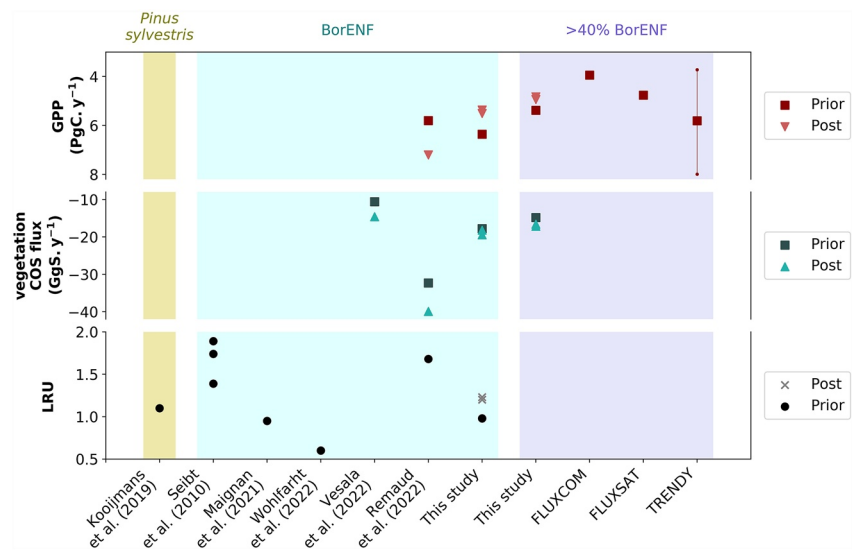


Figure 9. Synthesis of LRU values and vegetation COS and GPP fluxes from different studies focusing on boreal biomes. Note that the y-axis has been inverted for GPP (with largest positive values on the bottom, and decreasing moving up) so that both GPP and vegetation COS fluxes representing uptakes are in the same direction. Also note that each study follows a different approach to estimate the changes in COS or GPP fluxes relative to the reference values (see the referenced papers for more details). Each background color represents a specific vegetation type or studied area: Scots pine (*Pinus sylvestris*) tree (light yellow); BorENF PFT only (light blue); all land grid cells that have a fractional coverage of BorENF higher than 40% (in purple). For the comparison over land grid cells that have a fractional coverage of BorENF higher than 40% (in purple), we added the GPP estimates for the FLUXCOM and FLUXSAT evaluation products, and the mean GPP from the TRENDY-V10 model ensemble with the minimum and maximum GPP. Over this region, contrary to Figure 8, the COS and GPP fluxes for this study were computed using the post optimization values for α and A1 parameters from the multi-site optimizations for all C₃ PFTs.

COS fluxes that was used in SiB4 to scale up the COS fluxes to the region covered with BorENF. This parameterization also led to an increase in COS uptake for BorENF from 10.6 GgS yr⁻¹ for the prior SiB4 simulation to 14.6 GgS yr⁻¹. In this study, we found a vegetation COS sink for BorENF from the two optimizations including COS assimilation (average of 18.8 GgS yr⁻¹) that is higher than the one computed in SiB4 with the parametric model.

Unlike inversion studies where a linear relationship is used to link the COS fluxes and GPP, directly optimizing the model parameters does not imply that vegetation COS fluxes and GPP evolve in the same direction. While Remaud et al. (2022) and this study agree on an increase of vegetation COS uptake for BorENF, the changes in GPP differ. The optimization performed in Remaud et al. (2022) increased the GPP considerably for BorENF (from 5.8 PgC yr⁻¹ to 7.2 PgC yr⁻¹), which would disagree with the estimates from the FLUXCOM and FLUXSAT products (3.94 PgC yr⁻¹ and 4.76 PgC yr⁻¹ respectively over land grid cells that have a fractional coverage of BorENF higher than 40%). Then, Kuai et al. (2022) also used atmospheric COS observations to optimize the vegetation COS uptake but using aircraft profiles over Alaska (CARVE). The vegetation COS uptake was increased by 25% over the Northern high latitudes (40°–90°N), as well as GPP given that vegetation COS flux and GPP are linked with a linear relationship. It should be noted that all these inversion studies use atmospheric COS concentrations over northern North America to constrain the whole boreal region, while the optimization of ecosystem COS fluxes is driven by FI-Hyy in this study.

Estimating LRU is critical for inversion studies that rely on this empirical approach to link vegetation COS uptake and GPP. However, temporal variations in LRU in response to changes in environmental drivers such as light and VPD (Kohonen, Dewar, et al., 2022; Kooijmans et al., 2019; Sun et al., 2022) are not accounted for in some of these studies (Hu et al., 2021; Remaud et al., 2022). Following our optimizations, we computed the seasonal variations of optimal LRU values of BorENF for each scenario (Figure S14 in Supporting Information S1), as described in Maignan et al. (2021). The LRU seasonal cycles show that monthly LRU can vary by 0.3–0.7 within a year with lower values in spring and autumn. Then, the resulting mean annual LRU is increased by 20% with the two assimilations including COS data (LRU of 1.20 for “Post F_{COS} & GPP ½” and 1.23 for “Post F_{COS} &

GPP $\frac{1}{3}$ ”) compared to the prior (0.98). The assimilation including only GPP data led to the highest LRU of 3.24, a value not reported in Figure 9 because based on an inconsistent vegetation COS uptake after optimization. The LRU values from the optimized fluxes including GPP and COS assimilations are in the low range of the reported LRU values with a median of 1.68 for C3 plants (Whelan et al., 2018), but close to the value obtained in Kooijmans et al. (2019) at FI-Hyy for Scots pine (*Pinus sylvestris*) under light-saturated conditions (1.1). Indeed, the lowest LRU values have been estimated for boreal ecosystems (Maignan et al., 2021; Seibt et al., 2010). This was recently supported by Wohlfahrt et al. (2023) who provided new LRU estimates for each biome based on plant optimality principles and found LRU values around 0.5 in high latitudes. However, this LRU data set is computed for light-saturated conditions (photosynthetically active radiations higher than $1,000 \mu\text{mol m}^{-2} \text{s}^{-1}$), which can explain its low values as LRU increases with low photosynthetically active radiations (Kooijmans et al., 2019). Therefore, considering a sunlit leaf at the top of the canopy also induces a scale bias between the two LRU estimates as ORCHIDEE computes LRU integrated over the canopy, and not at the leaf scale. It is to be noted that their estimate for BorENF (around 0.6) is about half of the LRU values found after our optimizations including COS data. Therefore, while branch chamber measurements enable to study the LRU dynamics for a given species and determine a LRU value under light-saturated conditions (Kooijmans et al., 2019), a strong uncertainty remains when considering a constant LRU in inversion frameworks for BorENF after several studies focusing on this biome, with values ranging from 0.6 to 1.89 (Figure 9).

Finally, as BorENF is only one of the biomes in the high latitudes where the missing COS sink was identified, we performed a scaling up of the two optimizations including COS assimilations to the whole boreal region (40° – 80° N). We ran ORCHIDEE using the post optimization values for α and A1 parameters from the multi-site optimizations (“Post F_{COS} & GPP $\frac{1}{2}$ ” and “Post F_{COS} & GPP $\frac{1}{3}$ ”) for all C₃ PFTs as the values of these two parameters were initially defined according to the photosynthetic pathways (C₃ or C₄ plants). While the α parameter is specific to the vegetation COS uptake model, the A1 parameter is involved in the computation of the stomatal conductance, which impacts COS exchanges but also GPP and LE. Considering the higher post optimization value for α and the decrease in A1 (Figure S5 in Supporting Information S1) leads to an increase of 18% in vegetation COS uptake (93.9 GgS yr^{-1} for both assimilation scenarios) over the whole boreal region compared to the prior simulation ($79.75 \text{ GgS yr}^{-1}$) (Table S3 in Supporting Information S1). On the other hand, changing only the value of A1 has a negligible impact on GPP and LE fluxes over the boreal region.

4.3. Remaining Challenges for Ecosystem COS Flux Modeling

The inability of the simulated ecosystem COS flux to reproduce the observed seasonal amplitude after optimization can highlight model deficiencies specific to the biospheric COS flux representation in ORCHIDEE. In particular, the implementation of the internal conductance to COS as a function of the Rubisco maximum carboxylation rate (V_{cmax}) instead of directly representing the activity of the CA enzyme could lead to an inadequate response of vegetation COS uptake to environmental factors. Indeed, while photosynthetic activity depends on both light and temperature, the activity of CA is not light dependent (Protoschill-Krebs et al., 1996). Moreover, Cho et al. (2023) recently proposed a new function to describe the temperature response of CA that accounts for a temperature optimum specific to this enzyme. Following this approach, the internal conductance to COS is still expressed as a function of the α parameter and the $V_{\text{cmax}_{25}}$ of the Rubisco, but includes the temperature response function of CA instead of the one of Rubisco. They found that the temperature response of CA and Rubisco differs, with a lower optimum temperature for CA at FI-Hyy (22°C against 45°C for Rubisco). We implemented this new CA temperature response in ORCHIDEE and optimized the ecosystem COS flux at FI-Hyy with this updated formulation of $g_{i,\text{COS}}$ (see Text S1 in Supporting Information S1). However, the updated prior simulated ecosystem COS uptake was strongly overestimated compared to the observations (Figure S15a in Supporting Information S1). Then, the COS assimilations led to a higher RMSD than the one obtained after the optimization with the formulation from Berry et al. (2013) for $g_{i,\text{COS}}$ (RMSD of $3.6 \text{ pmol m}^{-2} \text{ s}^{-1}$ against $3.2 \text{ pmol m}^{-2} \text{ s}^{-1}$), further degrading the ecosystem COS flux seasonal amplitude. Therefore, in ORCHIDEE, the implementation of a temperature response specific to CA is not sufficient to improve the representation of vegetation COS fluxes. Furthermore, assimilating ecosystem COS flux only, with this CA temperature response, also leads to a stronger degradation of the GPP seasonal amplitude (Figure S15b in Supporting Information S1). A next step in the vegetation COS uptake model development could be to define a formulation for COS consumption by CA that is independent from the Rubisco activity for CO₂ assimilation. This would imply to distinguish between the mesophyll conductance and CA activity in $g_{i,\text{COS}}$. Following a different approach, Davidson et al. (2022) recently highlighted

the potential of sulfur isotopes to better constrain the internal conductance to COS by providing information on COS internal concentration inside the leaves.

In the current formulations of $g_{i,COS}$ (Berry et al., 2013; Cho et al., 2023), the α parameter is critical for vegetation COS uptake and optimizing its value is necessary as this parameter, but also its temporal and spatial variability and drivers, are not well constrained in the literature (Berry et al., 2013). In this study with the formulation of Berry et al. (2013) for $g_{i,COS}$, the two optimizations including COS data increase the value of α , which participates in the increase of the internal conductance to COS (Figure S3 in Supporting Information S1). Post optimization values are between 0.0014 and 0.0016 mol μmol^{-1} while the default value used for all C_3 plants in Berry et al. (2013) is 0.0012 mol μmol^{-1} . However, in SiB3 where the Berry et al. (2013) model was initially implemented, the value of α was recently revised to 0.0014 mol μmol^{-1} for C_3 plants (Kooijmans et al., 2021). Cho et al. (2023) obtained lower values for α (0.001316 mol μmol^{-1} and 0.001331 mol μmol^{-1} for growth and maturity phenological stages, respectively) at FI-Hyy following their optimization with the CA temperature response in SiB4. Then, performing similar optimizations on sites dominated by different vegetation types could help to refine the definition of this parameter with estimated values specific to a biome type instead of a photosynthetic pathway only (C_3 and C_4 plants). For example, Cho et al. (2023) extended the optimization of α to the Harvard Forest site (United States) in addition to FI-Hyy and found values between 0.001740 mol μmol^{-1} and 0.002224 mol μmol^{-1} depending on the phenological stage.

Despite the insufficiency of the α parameter to adequately represent the internal conductance to COS, its value still needed to be calibrated compared to the default value of 0.0012 mol μmol^{-1} as a major change was made to the vegetation COS model when considering 3-hourly variable atmospheric COS concentrations instead of a constant value of 500 ppt in the previous version of the model (Maignan et al., 2021). Initially α was estimated by fitting a regression between vegetation COS uptake observations (Stimler et al., 2010) and SiB3 simulated fluxes (Berry et al., 2013). When estimating α following this approach, the estimated value of α depends on the atmospheric COS concentration considered to compute the vegetation COS fluxes.

Neglecting the COS drawdown during the growing season would lead to an overestimation of the simulated ecosystem COS uptake, impacting the optimized parameter values when performing data assimilation. It is to be noted that an additional drawdown of COS concentration further down inside the canopy due to leaf COS uptake, or possibly related to understory, ground cover vegetation, or mosses, is not taken into account in ORCHIDEE. To highlight the importance of considering a variable COS concentration in the biospheric COS models, Belviso et al. (2022) estimated the impact of the recent decline in atmospheric COS concentration on the simulated vegetation COS fluxes in the ORCHIDEE LSM. However, at FI-Hyy, Vesala et al. (2022) found a lower univariate correlation between COS fluxes and atmospheric COS mixing ratio than with air or soil temperature or net radiation, but still larger than with VPD. It should be noted that the optimized atmospheric COS concentrations used to compute the COS fluxes are also associated with posterior uncertainties (Remaud et al., 2022).

In addition to the uncertainty on the parameter estimates, the vegetation and soil COS models implemented in ORCHIDEE are missing some processes, which can contribute to the model observation mismatch. For example, non stomatal components such as mosses were found to absorb COS during nighttime (Rastogi et al., 2018; Sun et al., 2018). COS uptake by understory vegetation is also neglected. These missing processes could explain part of the underestimation of the net ecosystem COS uptake simulated during summertime at FI-Hyy (Figure 3). The representation of soil COS exchanges also neglects processes such as the effect of snow cover on COS diffusion for subnival COS uptake or the impact of solar radiations on COS production (Abadie et al., 2022; Kitz et al., 2020; Spielmann et al., 2019).

Then, a strong limitation for assimilating biospheric COS measurements to constrain GPP and LE is the lack of COS flux measurement sites and long time series, with FI-Hyy being the longest available time series of 5 consecutive years. Moreover, the absence of data during wintertime does not enable to fully capture the seasonal cycle amplitude of ecosystem COS fluxes for evergreen biomes. Going further in the assimilation of ecosystem COS fluxes in LSMs would greatly benefit from having a larger network of COS flux measurements, on the model of the FLUXNET network for carbon and energy flux measurements. However, developing such a network is currently restrained by the high instrumental cost of performing COS flux measurements. The NOAA/ESRL offers a better cover for atmospheric COS flask measurements, with about 15 stations measuring COS concentrations since 2000. However, the continental stations are mainly located over North America.

Despite these challenges in ecosystem COS flux modeling, assimilating biospheric COS fluxes has different advantages than assimilating COS concentrations. Indeed, tower based COS flux measurements capture only the COS sinks and sources included in the tower footprint, enabling to mainly focus on the vegetation and soil contributions. Therefore, working at the ecosystem scale enables us to exclude contributions from other components of the COS budget associated with large uncertainties, such as the ocean with a contribution estimated between 200 GgS yr⁻¹ and 1,000 GgS yr⁻¹ (Berry et al., 2013; Kuai et al., 2015; Launois et al., 2015; Lennartz et al., 2017, 2020; Remaud et al., 2022; Suntharalingam et al., 2008). In addition, this can help to identify missing processes that have an impact at the ecosystem scale.

5. Conclusions and Outlooks

We performed a joint assimilation of ecosystem COS fluxes and GPP to directly optimize the parameters involved in the representation of these fluxes for BorENF. The key messages from this study and the arising recommendations for future work are the following:

- Jointly assimilating ecosystem COS flux and GPP data enables us to improve both GPP and LE due to the strong link among COS, CO₂ and H₂O fluxes through stomatal diffusion. This finding supports using COS flux data to obtain new insights into plant transpiration and encourages for future exploration into inferring transpiration from COS measurements. An alternative optimization worth considering could be to assimilate both ecosystem COS flux and LE, which could then be compared to the assimilation of ecosystem COS flux and GPP performed in this study. Exploring different data assimilation strategies may inform the complementarity of different flux measurements in constraining leaf-level parameters that control biosphere-atmosphere exchange of H₂O, CO₂, and COS.
- Despite the fact that COS data assimilation yields an increase in simulated net ecosystem COS uptake to better match the observations at FI-Hyy, the representation of COS internal conductance following Berry et al. (2013) is structurally insufficient to reproduce the ecosystem COS flux seasonal amplitude. This evinces the need to implement a mechanistic representation of a mesophyll diffusion and enzymatic consumption specific to COS in LSMs.
- Ecosystem COS flux assimilation can highlight deficiencies in simulated GPP and LE sensitivity to drought events. Along this line, COS fluxes could be used to evaluate GPP and LE sensitivity to water stress in LSMs, and this potential application should be further investigated, for example, COS flux measurements could help to better capture the response of the simulated WUE that is currently underestimated in the ORCHIDEE LSM (De Pue et al., 2022). Thus, it would be interesting to perform COS data assimilation for ecosystems more exposed to drought than BorENF, using the data collected in Wohlfahrt et al. (2018) and Cochavi et al. (2021) during heat waves, in a Mediterranean pine forest and citrus orchard, respectively.
- The joint assimilation of ecosystem COS flux and GPP leads to an increase in vegetation COS uptake, but not in GPP for BorENF, thereby alleviating the mismatch between the model and the evaluation GPP products. It is to be noted that the assimilation of ecosystem COS flux and GPP data together enables parameter optimization in a way consistent with the shared stomatal control of COS and CO₂ vegetation uptakes, contrary to a COS flux-only or GPP-only data assimilation. These post-optimization changes in COS and GPP fluxes contrast with previous top-down studies assimilating atmospheric COS concentrations, which find increased vegetation COS uptake along with GPP due to an empirical linear relationship (i.e., through LRU) being imposed between the two fluxes. Therefore, the empirical approach to derive GPP from COS may introduce an important source of uncertainty. This also highlights the need for continued work to reconcile top-down and bottom-up COS-derived GPP estimates.
- Considering that COS serves as a mechanistic constraint on stomatal diffusion and helps to inform the carbon assimilation stage of photosynthesis, a complementary approach could be to jointly assimilate COS and solar-induced fluorescence, the latter of which has a global coverage from satellite observations and can inform the light reactions of photosynthesis. This complementary approach is a promising path toward a more comprehensive mechanistic understanding of global photosynthesis.

Data Availability Statement

The eddy covariance COS flux data from Hyttiälä used for ecosystem COS flux assimilation and evaluation in this study are available in Kohonen, Tramontana, and Kolari (2022) via <https://doi.org/10.5281/zenodo.6940750>. The soil COS flux from chamber measurements at Hyttiälä used for assimilation are published in Sun et al. (2017) and available via <https://doi.org/10.5281/zenodo.322936>. The FLUXNET data used for data assimilation and evaluation are available at <https://fluxnet.org> following the Creative Commons (CC-BY 4.0) license (Pastorello et al., 2020). The FLUXCOM GPP and LE data products are freely available from the data portal of Max Planck Institute for Biogeochemistry at <http://fluxcom.org>. The FLUXSAT GPP data product can be downloaded from the NASA Goddard Space Flight Center at <https://avdc.gsfc.nasa.gov>. The GLEAM LE product is provided by the GLEAM project and available at <https://www.gleam.eu>. The CMIP6 version of the ORCHIDEE model including the COS submodel is available upon request to the authors. Figures were made using Python programming language version 3 (Van Rossum and Drake, 2009), available at <https://www.python.org>. Figure 2 uses the “shadedrelief” map as background (<http://www.shadedrelief.com>).

Acknowledgments

The authors are very grateful to everyone who participated in field data collection used in this study. We warmly thank X. Wang-Faivre and C. Bacour for formatting and sharing the global FLUXCOM, GLEAM, and FLUXSAT products. We would also like to thank the editor and the three reviewers, Mary Whelan and two anonymous referees, for their insightful comments and suggestions that have helped us improve the manuscript. This work used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (U.S. Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), Green-Grass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval and Environment Canada and US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California - Berkeley, University of Virginia. The FLUXNET data can be downloaded at <https://fluxnet.org>. Finally, we would like to acknowledge the computer team at LSCE for the computational resources provided. This research has been mainly supported by the 4C project of the European Commission's Horizon 2020 framework programme (Grant 821003). The eddy covariance measurements at Hyttiälä were supported by ICOS-Finland (319871) by the University of Helsinki.

References

- Abadie, C., Maignan, F., Remaud, M., Ogee, J., Campbell, J. E., Whelan, M. E., et al. (2022). Global modelling of soil carbonyl sulfide exchanges. *Biogeosciences*, 19(9), 2427–2463. <https://doi.org/10.5194/bg-19-2427-2022>
- Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., et al. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. *Reviews of Geophysics*, 53(3), 785–818. <https://doi.org/10.1002/2015RG000483>
- Asaf, D., Rotenberg, E., Tatarinov, F., Dicken, U., Montzka, S. A., & Yakir, D. (2013). Ecosystem photosynthesis inferred from measurements of carbonyl sulphide flux. *Nature Geoscience*, 6(3), 186–190. <https://doi.org/10.1038/ngeo1730>
- Badger, M. R., & Dean Price, G. (1994). The role of carbonic anhydrase in photosynthesis. In *Annual Review of Plant Physiology and Plant Molecular Biology* (Vol. 45). Retrieved from www.annualreviews.org
- Baldocchi, D., Falge, L., Gu, L., Olson, R., Hollinger, D., Running, S., et al. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, 82(11), 2415–2434. [https://doi.org/10.1175/1520-0477\(2001\)082<2415:FANTTS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2)
- 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. <https://doi.org/10.5194/gmd-11-4739-2018>
- Beck, P. S. A., Atzberger, C., Høgda, K. A., Johansen, B., & Skidmore, A. K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, 100(3), 321–334. <https://doi.org/10.1016/j.rse.2005.10.021>
- Belviso, S., Remaud, M., Abadie, C., Maignan, F., Ramonet, M., & Peylin, P. (2022). Ongoing decline in the atmospheric COS seasonal cycle amplitude over Western Europe: Implications for surface fluxes. *Atmosphere*, 13(5), 812. <https://doi.org/10.3390/atmos13050812>
- Berkelhammer, M., Alsip, B., Matamala, R., Cook, D., Whelan, M. E., Joo, E., et al. (2020). Seasonal evolution of canopy stomatal conductance for a prairie and maize field in the midwestern United States from continuous carbonyl sulfide fluxes. *Geophysical Research Letters*, 47(6). <https://doi.org/10.1029/2019GL085652>
- Berkelhammer, M., Asaf, D., Still, C., Montzka, S., Noone, D., Gupta, M., et al. (2014). Constraining surface carbon fluxes using in situ measurements of carbonyl sulfide and carbon dioxide. *Global Biogeochemical Cycles*, 28(2), 161–179. <https://doi.org/10.1002/2013GB004644>
- Berry, J., Wolf, A., Campbell, J. E., Baker, I., Blake, N., Blake, D., et al. (2013). A coupled model of the global cycles of carbonyl sulfide and CO₂: A possible new window on the carbon cycle. *Journal of Geophysical Research: Biogeosciences*, 118(2), 842–852. <https://doi.org/10.1002/jgrg.20068>
- Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, 320(5882), 1444–1449. <https://doi.org/10.1126/science.1155121>
- Böttcher, K., Aurela, M., Kervinen, M., Markkanen, T., Mattila, O.-P., Kolari, P., et al. (2014). MODIS time-series-derived indicators for the beginning of the growing season in boreal coniferous forest — A comparison with CO₂ flux measurements and phenological observations in Finland. *Remote Sensing of Environment*, 140, 625–638. <https://doi.org/10.1016/j.rse.2013.09.022>
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, 12(7). <https://doi.org/10.1029/2019MS002010>
- Campbell, J. E., Carmichael, G. R., Chai, T., Mena-Carrasco, M., Tang, Y., Blake, D. R., et al. (2008). Photosynthetic control of atmospheric carbonyl sulfide during the growing season. *Science*, 322(5904), 1085–1088. <https://doi.org/10.1126/science.1164015>
- Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10), 1509–1518. <https://doi.org/10.1016/j.envsoft.2006.10.004>
- Chang, Q., Xiao, X., Jiao, W., Wu, X., Dougherty, R., Wang, J., et al. (2019). Assessing consistency of spring phenology of snow-covered forests as estimated by vegetation indices, gross primary production, and solar-induced chlorophyll fluorescence. *Agricultural and Forest Meteorology*, 275, 305–316. <https://doi.org/10.1016/j.agrformet.2019.06.002>
- Chen, B., Wang, P., Wang, S., Ju, W., Liu, Z., & Zhang, Y. (2023). Simulating canopy carbonyl sulfide uptake of two forest stands through an improved ecosystem model and parameter optimization using an ensemble Kalman filter. *Ecological Modelling*, 475, 110212. <https://doi.org/10.1016/j.ecolmodel.2022.110212>
- Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, É., Gastineau, G., et al. (2020). Improved near-surface continental climate in IPSL-CM6A-LR by combined evolutions of atmospheric and land surface physics. *Journal of Advances in Modeling Earth Systems*, 12(10). <https://doi.org/10.1029/2019MS002005>
- Cho, A., Kooijmans, L. M. J., Kohonen, K.-M., Wehr, R., & Krol, M. C. (2023). Optimizing the carbonic anhydrase temperature response and stomatal conductance of carbonyl sulfide leaf uptake in the simple biosphere model (SiB4). <https://doi.org/10.5194/egusphere-2022-1305>

- Cochavi, A., Amer, M., Stern, R., Tatarinov, F., Migliavacca, M., & Yakir, D. (2021). Differential responses to two heatwave intensities in a Mediterranean citrus orchard are identified by combining measurements of fluorescence and carbonyl sulfide (COS) and CO₂ uptake. *New Phytologist*, 230(4), 1394–1406. <https://doi.org/10.1111/nph.17247>
- Collatz, G. J., Ball, J. T., Grivet, C., & Berry, J. A. (1991). Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: A model that includes a laminar boundary layer. *Agricultural and Forest Meteorology*, 54(2–4), 107–136. [https://doi.org/10.1016/0168-1923\(91\)90002-8](https://doi.org/10.1016/0168-1923(91)90002-8)
- Davidson, C., Amrani, A., & Angert, A. (2022). Carbonyl sulfide sulfur isotope fractionation during uptake by C3 and C4 plants. *Journal of Geophysical Research: Biogeosciences*, 127(10). <https://doi.org/10.1029/2022jg007035>
- Delbart, N., Kergoat, L., le Toan, T., L'hermitte, J., & Picard, G. (2005). Determination of phenological dates in boreal regions using normalized difference water index. Retrieved from <https://hal.ird.fr/ird-00392489>
- de Pue, J., Barrios, J. M., Liu, L., Ciaï, P., Arboleda, A., Hamdi, R., et al. (2022). Local-scale evaluation of the simulated interactions between energy, water and vegetation in ISBA, ORCHIDEE and a diagnostic model. *Biogeosciences*, 19(17), 4361–4386. <https://doi.org/10.5194/bg-19-4361-2022>
- DiMario, R. J., Quebedeaux, J. C., Longstreth, D. J., Dassanayake, M., Hartman, M. M., & Moroney, J. V. (2016). The cytoplasmic carbonic anhydrases βCA2 and βCA4 are required for optimal plant growth at low CO₂. *Plant Physiology*, 171(1), 280–293. <https://doi.org/10.1104/pp.15.01990>
- Farquhar, G. D., von Caemmerer, S. S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO₂ assimilation in leaves of C3 species. *Planta*, 90(1), 78–90. <https://doi.org/10.1007/BF00386231>
- Fisher, J. B., Hayes, D. J., Schwalm, C. R., Huntzinger, D. N., Stofferahn, E., Schaefer, K., et al. (2018). Missing pieces to modeling the Arctic-Boreal puzzle. *Environmental Research Letters*, 13(2), 020202. <https://doi.org/10.1088/1748-9326/aa9d9a>
- 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(17). <https://doi.org/10.1029/2011GL048738>
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., et al. (2020). Global carbon budget 2020. *Earth System Science Data*, 12(4), 3269–3340. <https://doi.org/10.5194/essd-12-3269-2020>
- Gao, Y., Markkanen, T., Aurela, M., Mammarella, I., Thum, T., Tsuruta, A., et al. (2017). Response of water use efficiency to summer drought in a boreal Scots pine forest in Finland. *Biogeosciences*, 14(18), 4409–4422. <https://doi.org/10.5194/bg-14-4409-2017>
- Goetz, S. J., Bunn, A. G., Fiske, G. J., & Houghton, R. A. (2005). Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance. *Proceedings of the National Academy of Sciences*, 102(38), 13521–13525. <https://doi.org/10.1073/pnas.0506179102>
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Publishing Company.
- Haupt, R. L., & Haupt, S. E. (2004). *Practical genetic algorithms*. Wiley. Retrieved from www.copyright.com
- Haynes, K., Baker, I., & Denning, S. (2020). *The simple biosphere model version 4.2: SiB4 technical description*. Mountain Scholar, Colorado State University. [online]. Retrieved from <https://hdl.handle.net/10217/200691>
- Hu, L., Montzka, S. A., Kaushik, A., Andrews, A. E., Sweeney, C., Miller, J., et al. (2021). COS-derived GPP relationships with temperature and light help explain high-latitude atmospheric CO₂ seasonal cycle amplification. <https://doi.org/10.1073/pnas.2103423118/-DCSupplemental>
- Jarvis, P. G. (1995). Scaling processes and problems. *Plant, Cell and Environment*, 18(10), 1079–1089. <https://doi.org/10.1111/j.1365-3040.1995.tb00620.x>
- 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., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, 6(1), 74. <https://doi.org/10.1038/s41597-019-0076-8>
- Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., et al. (2017). Compensatory water effects link yearly global land CO₂ sink changes to temperature. *Nature*, 541(7638), 516–520. <https://doi.org/10.1038/nature20780>
- Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, 17(5), 1343–1365. <https://doi.org/10.5194/bg-17-1343-2020>
- Kattge, J., & Knorr, W. (2007). Temperature acclimation in a biochemical model of photosynthesis: A reanalysis of data from 36 species. *Plant, Cell and Environment*, 30(9), 1176–1190. <https://doi.org/10.1111/j.1365-3040.2007.01690.x>
- Kitz, F., Spielmann, F. M., Hammerle, A., Kolle, O., Migliavacca, M., Moreno, G., et al. (2020). Soil COS exchange: A comparison of three European ecosystems. *Global Biogeochemical Cycles*, 34(4). <https://doi.org/10.1029/2019GB006202>
- Kohonen, K. M., Dewar, R., Tramontana, G., Mauranen, A., Kolari, P., Kooijmans, L. M. J., et al. (2022). Intercomparison of methods to estimate gross primary production based on CO₂ and COS flux measurements. *Biogeosciences*, 19(17), 4067–4088. <https://doi.org/10.5194/bg-19-4067-2022>
- Kohonen, K. M., Kolari, P., Kooijmans, L. M. J., Chen, H., Seibt, U., Sun, W., & Mammarella, I. (2020). Towards standardized processing of eddy covariance flux measurements of carbonyl sulfide. *Atmospheric Measurement Techniques*, 13(7), 3957–3975. <https://doi.org/10.5194/amt-13-3957-2020>
- Kohonen, K. M., Tramontana, G., & Kolari, P. (2022). Dataset for "Intercomparison of methods to estimate gross primary production based on CO₂ and COS flux measurements. [Dataset]. <https://doi.org/10.5281/zenodo.6940750>
- Kooijmans, L. M. J., Cho, A., Ma, J., Kaushik, A., Haynes, K. D., Baker, I., et al. (2021). Evaluation of carbonyl sulfide biosphere exchange in the Simple Biosphere Model (SiB4). *Biogeosciences*, 18(24), 6547–6565. <https://doi.org/10.5194/bg-18-6547-2021>
- Kooijmans, L. M. J., Maseyk, K., Seibt, U., Sun, W., Vesala, T., Mammarella, I., et al. (2017). Canopy uptake dominates nighttime carbonyl sulfide fluxes in a boreal forest. *Atmospheric Chemistry and Physics*, 17(18), 11453–11465. <https://doi.org/10.5194/acp-17-11453-2017>
- Kooijmans, L. M. J., Sun, W., Aalto, J., Erkkilä, K. M., Maseyk, K., Seibt, U., et al. (2019). Influences of light and humidity on carbonyl sulfide-based estimates of photosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7), 2470–2475. <https://doi.org/10.1073/pnas.1807600116>
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogee, 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(Issue 1), 1–33. <https://doi.org/10.1029/2003GB002199>
- Kuai, L., Parazoo, N. C., Shi, M., Miller, C. E., Baker, I., Bloom, A. A., et al. (2022). Quantifying northern high latitude gross primary productivity (GPP) using carbonyl sulfide (OCS). *Global Biogeochemical Cycles*, 36(9). <https://doi.org/10.1029/2021gb007216>
- Kuai, L., Worden, J. R., Campbell, J. E., Kulawik, S. S., Li, K. F., Lee, M., et al. (2015). Estimate of carbonyl sulfide tropical oceanic surface fluxes using aura tropospheric emission spectrometer observations. *Journal of Geophysical Research*, 120(20), 11012–11023. <https://doi.org/10.1002/2015JD023493>

- Kuppel, S., Peylin, P., Maignan, F., Chevallier, F., Kiely, G., Montagnani, L., & Cescatti, A. (2014). Model–data fusion across ecosystems: From multisite optimizations to global simulations. *Geoscientific Model Development*, 7(6), 2581–2597. <https://doi.org/10.5194/gmd-7-2581-2014>
- Lardy, R., Bellocchi, G., & Soussana, J. F. (2011). A new method to determine soil organic carbon equilibrium. *Environmental Modelling & Software*, 26(12), 1759–1763. <https://doi.org/10.1016/j.envsoft.2011.05.016>
- Lasslop, G., Reichstein, M., Papale, D., Richardson, A., Arneth, A., Barr, A., et al. (2010). Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: Critical issues and global evaluation. *Global Change Biology*, 16(1), 187–208. <https://doi.org/10.1111/j.1365-2486.2009.02041.x>
- Launois, T., Peylin, P., Belviso, S., & Poulter, B. (2015). A new model of the global biogeochemical cycle of carbonyl sulfide - Part 2: Use of carbonyl sulfide to constrain gross primary productivity in current vegetation models. *Atmospheric Chemistry and Physics*, 15(16), 9285–9312. <https://doi.org/10.5194/acp-15-9285-2015>
- Lennartz, S., A Marandino, C., von Hobe, M., O Andreae, M., Aranami, K., Atlas, E., et al. (2020). Marine carbonyl sulfide (OCS) and carbon disulfide (CS₂): A compilation of measurements in seawater and the marine boundary layer. *Earth System Science Data*, 12(1), 591–609. <https://doi.org/10.5194/essd-12-591-2020>
- Lennartz, S. T., Marandino, C. A., von Hobe, M., Cortes, P., Quack, B., Simo, R., et al. (2017). Direct oceanic emissions unlikely to account for the missing source of atmospheric carbonyl sulfide. *Atmospheric Chemistry and Physics*, 17(1), 385–402. <https://doi.org/10.5194/acp-17-385-2017>
- Li, X., Xiao, J., He, B., Altaf Arain, M., Beringer, J., Desai, A. R., et al. (2018). Solar-induced chlorophyll fluorescence is strongly correlated with terrestrial photosynthesis for a wide variety of biomes: First global analysis based on OCO-2 and flux tower observations. *Global Change Biology*, 24(9), 3990–4008. <https://doi.org/10.1111/gcb.14297>
- Ma, J., Kooijmans, L. M. J., Cho, A., Montzka, S. A., Glatthor, N., Worden, J. R., et al. (2021). Inverse modelling of carbonyl sulfide: Implementation, evaluation and implications for the global budget. *Atmospheric Chemistry and Physics*, 21(5), 3507–3529. <https://doi.org/10.5194/acp-21-3507-2021>
- MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E. N., Kuppel, S., et al. (2022). Quantifying and reducing uncertainty in global carbon cycle predictions: Lessons and perspectives from 15 years of data assimilation studies with the ORCHIDEE terrestrial biosphere model. *Global Biogeochemical Cycles*, 36(Issue 7). <https://doi.org/10.1029/2021GB007177>
- Mahmud, K., Scott, R. L., Biederman, J. A., Litvak, M. E., Kolb, T., Meyers, T. P., et al. (2021). Optimizing carbon cycle parameters drastically improves terrestrial biosphere model underestimates of dryland mean net CO₂ flux and its inter-annual 2 variability.
- Maignan, F., Abadie, C., Remaud, M., Kooijmans, L. M. J., Kohonen, K. M., Commare, R., et al. (2021). Carbonyl sulfide: Comparing a mechanistic representation of the vegetation uptake in a land surface model and the leaf relative uptake approach. *Biogeosciences*, 18(9), 2917–2955. <https://doi.org/10.5194/bg-18-2917-2021>
- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5), 1903–1925. <https://doi.org/10.5194/gmd-10-1903-2017>
- Masaki, Y., Iizuka, R., Kato, H., Kojima, Y., Ogawa, T., Yoshida, M., et al. (2021). Fungal carbonyl sulfide hydrolase of trichoderma harzianum strain thif08 and its relationship with clade D β-carbonic anhydrases. *Microbes and Environments*, 36(2). <https://doi.org/10.1264/jmsme2.ME20058>
- Mencuccini, M., Manzoni, S., & Christoffersen, B. (2019). Modelling water fluxes in plants: From tissues to biosphere. *New Phytologist*, 222(3), 1207–1222. <https://doi.org/10.1111/nph.15681>
- Montzka, S. A., Calvert, P., Hall, B. D., Elkins, J. W., Conway, T. J., Tans, P. P., & Sweeney, C. S. (2007). On the global distribution, seasonality, and budget of atmospheric carbonyl sulfide (COS) and some similarities to CO₂. *Journal of Geophysical Research*, 112(9), D09302. <https://doi.org/10.1029/2006JD007665>
- Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2), 161–174. <https://doi.org/10.1080/00401706.1991.10484804>
- Ogé, J., Sauze, J., Kesselmeier, J., Genty, B., van Diest, H., Launois, T., & Wingate, L. (2016). A new mechanistic framework to predict OCS fluxes from soils. *Biogeosciences*, 13(8), 2221–2240. <https://doi.org/10.5194/bg-13-2221-2016>
- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y. W., et al. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, 7(1), 225. <https://doi.org/10.1038/s41597-020-0534-3>
- 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>
- 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>
- Protoschill-Krebs, G., Wilhelm, C., & Kesselmeier, J. (1996). Consumption of carbonyl sulfide (COS) by higher plant carbonic anhydrase (CA). *Atmospheric Environment*, 30(Issue 18), 3151–3156. [https://doi.org/10.1016/1352-2310\(96\)00026-x](https://doi.org/10.1016/1352-2310(96)00026-x)
- Rastogi, B., Berkelhammer, M., Wharton, S., Whelan, M. E., Meinzer, F. C., Noone, D., & Still, C. J. (2018). Ecosystem fluxes of carbonyl sulfide in an old-growth forest: Temporal dynamics and responses to diffuse radiation and heat waves. *Biogeosciences*, 15(23), 7127–7139. <https://doi.org/10.5194/bg-15-7127-2018>
- 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>
- Remaud, M., Chevallier, F., Maignan, F., Belviso, S., Berchet, A., Parouffe, A., et al. (2022). Plant gross primary production, plant respiration, and carbonyl sulfide emissions over the globe inferred by atmospheric inverse modeling. *Atmospheric Chemistry and Physics*, 22(4), 2525–2552. <https://doi.org/10.5194/acp-22-2525-2022>
- Remaud, M., Ma, J., Krol, M., Abadie, C., Cartwright, M. P., Patra, P., et al. (2023). Intercomparison of atmospheric carbonyl sulfide (TransCom-COS; part one): Evaluating the impact of transport and emissions on tropospheric variability using ground-based and aircraft data. *Journal of Geophysical Research: Atmospheres*, 128(6), e2022JD037817. <https://doi.org/10.1029/2022JD037817>
- Reynolds, C. A., Jackson, T. J., & Rawls, W. J. (2000). Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions. *Water Resources Research*, 36(12), 3653–3662. <https://doi.org/10.1029/2000WR900130>
- Sandoval-Soto, L., Stanimirov, M., von Hobe, M., Schmitt, V., Valdes, J., Wild, A., & Kesselmeier, J. (2005). Global uptake of carbonyl sulfide (COS) by terrestrial vegetation: Estimates corrected by deposition velocities normalized to the uptake of carbon dioxide (CO₂). *Biogeosciences*, 2, 125–132. <https://doi.org/10.5194/bg-2-125-2005>

- Schlesinger, W. H., & Jasechko, S. (2014). Transpiration in the global water cycle. *Agricultural and Forest Meteorology*, 189–190, 115–117. <https://doi.org/10.1016/j.agrformet.2014.01.011>
- Seibt, U., Kesselmeier, J., Sandoval-Soto, L., Kuhn, U., & Berry, J. A. (2010). A kinetic analysis of leaf uptake of COS and its relation to transpiration, photosynthesis and carbon isotope fractionation. *Biogeosciences*, 7(1), 333–341. <https://doi.org/10.5194/bg-7-333-2010>
- Shen, M., Tang, Y., Desai, A. R., Gough, C., & Chen, J. (2014). Can EVI-derived land-surface phenology be used as a surrogate for phenology of canopy photosynthesis? *International Journal of Remote Sensing*, 35(3), 1162–1174. <https://doi.org/10.1080/01431161.2013.875636>
- Spielmann, F. M., Wohlfahrt, G., Hammerle, A., Kitz, F., Migliavacca, M., Alberti, G., et al. (2019). Gross primary productivity of four European ecosystems constrained by joint CO₂ and COS flux measurements. *Geophysical Research Letters*, 46(10), 5284–5293. <https://doi.org/10.1029/2019GL082006>
- Stimler, K., Montzka, S. A., Berry, J. A., Rudich, Y., & Yakir, D. (2010). Relationships between carbonyl sulfide (COS) and CO₂ during leaf gas exchange. *New Phytologist*, 186(4), 869–878. <https://doi.org/10.1111/j.1469-8137.2010.03218.x>
- Stinecipher, J. R., Cameron-Smith, P. J., Blake, N. J., Kuai, L., Lejeune, B., Mahieu, E., et al. (2019). Biomass burning unlikely to account for missing source of carbonyl sulfide. *Geophysical Research Letters*, 46(24), 14912–14920. <https://doi.org/10.1029/2019GL085567>
- Sun, W., Berry, J. A., Yakir, D., & Seibt, U. (2022). Leaf relative uptake of carbonyl sulfide to CO₂ seen through the lens of stomatal conductance–photosynthesis coupling. *New Phytologist*, 235(5), 1729–1742. <https://doi.org/10.1111/nph.18178>
- Sun, W., Kooijmans, L. M. J., Maseyk, K., Chen, H., Mammarella, I., Vesala, T., et al. (2017). Dataset for "Soil fluxes of carbonyl sulfide (COS), carbon monoxide, and carbon dioxide in a boreal forest in southern Finland. [Dataset]. <https://doi.org/10.5281/zenodo.322936>
- Sun, W., Kooijmans, L. M. J., Maseyk, K., Chen, H., Mammarella, I., Vesala, T., et al. (2018). Soil fluxes of carbonyl sulfide (COS), carbon monoxide, and carbon dioxide in a boreal forest in southern Finland. *Atmospheric Chemistry and Physics*, 18(2), 1363–1378. <https://doi.org/10.5194/acp-18-1363-2018>
- Suni, T., Rinne, J., & Reissell, A. (2003). Long-term measurements of surface fluxes above a Scots pine forest in Hyytiälä, Southern Finland, 1996–2001. 8, 287–301.
- Suntharalingam, P., Kettle, A. J., Montzka, S. M., & Jacob, D. J. (2008). Global 3-D model analysis of the seasonal cycle of atmospheric carbonyl sulfide: Implications for terrestrial vegetation uptake. *Geophysical Research Letters*, 35(19), L19801. <https://doi.org/10.1029/2008GL034332>
- Tarantola, A. (2005). Inverse problem theory and methods for model parameter estimation. *Society for Industrial and Applied Mathematics*. <https://doi.org/10.1137/1.9780898717921>
- Tootchi, A., Jost, A., & Ducharne, A. (2019). Multi-source global wetland maps combining surface water imagery and groundwater constraints. *Earth System Science Data*, 11(1), 189–220. <https://doi.org/10.5194/essd-11-189-2019>
- 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>
- Tramontana, G., Migliavacca, M., Jung, M., Reichstein, M., Keenan, T. F., Camps-Valls, G., et al. (2020). Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks. *Global Change Biology*, 26(9), 5235–5253. <https://doi.org/10.1111/gcb.15203>
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.
- Vesala, T., Kohonen, K. M., Kooijmans, L. M. J., Praplan, A. P., Foltynová, L., Kolari, P., et al. (2022). Long-term fluxes of carbonyl sulfide and their seasonality and interannual variability in a boreal forest. *Atmospheric Chemistry and Physics*, 22(4), 2569–2584. <https://doi.org/10.5194/acp-22-2569-2022>
- Wang, C., Beringer, J., Hutley, L. B., Cleverly, J., Li, J., Liu, Q., & Sun, Y. (2019). Phenology dynamics of dryland ecosystems along the North Australian tropical transect revealed by satellite solar-induced chlorophyll fluorescence. *Geophysical Research Letters*, 46(10), 5294–5302. <https://doi.org/10.1029/2019GL082716>
- Wang, K., & Dickinson, R. E. (2012). A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Reviews of Geophysics*, 50(Issue 2). <https://doi.org/10.1029/2011RG000373>
- Wehr, R., Commane, R., Munger, J. W., Barry Mcmanus, J., Nelson, D. D., Zahniser, M. S., et al. (2017). Dynamics of canopy stomatal conductance, transpiration, and evaporation in a temperate deciduous forest, validated by carbonyl sulfide uptake. *Biogeosciences*, 14(2), 389–401. <https://doi.org/10.5194/bg-14-389-2017>
- Whelan, M., Lennartz, S., Gimeno, T., Wehr, R., Wohlfahrt, G., Wang, Y., et al. (2018). Reviews and syntheses: Carbonyl sulfide as a multi-scale tracer for carbon and water cycles. *European Geosciences Union*, 15(12), 3625–3657. <https://doi.org/10.5194/bg-15-3625-2018>
- Wohlfahrt, G., Brill, F., Hörtnagl, L., Xu, X., Bingemer, H., Hansel, A., & Loreto, F. (2012). Carbonyl sulfide (COS) as a tracer for canopy photosynthesis, transpiration and stomatal conductance: Potential and limitations. *Plant, Cell and Environment*, 35(4), 657–667. <https://doi.org/10.1111/j.1365-3040.2011.02451.x>
- Wohlfahrt, G., Gerdel, K., Migliavacca, M., Rotenberg, E., Tatarinov, F., Müller, J., et al. (2018). Sun-induced fluorescence and gross primary productivity during a heat wave. *Scientific Reports*, 8(1), 14169. <https://doi.org/10.1038/s41598-018-32602-z>
- Wohlfahrt, G., Hammerle, A., Spielmann, F. M., Kitz, F., & Yi, C. (2023). Technical note: Novel estimates of the leaf relative uptake rate of carbonyl sulfide from optimality theory. *Biogeosciences*, 20(3), 589–596. <https://doi.org/10.5194/bg-20-589-2023>
- 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. <https://doi.org/10.1016/j.njas.2009.07.001>
- Zumkehr, A., Hilton, T. W., Whelan, M., Smith, S., Kuai, L., Worden, J., & Campbell, J. E. (2018). Global gridded anthropogenic emissions inventory of carbonyl sulfide. *Atmospheric Environment*, 183, 11–19. <https://doi.org/10.1016/j.atmosenv.2018.03.063>